 doi: 10.1111/j.1365-2524.2010.00962.x

Detailed guidelines for reporting quantitative research in *Health & Social Care in the Community*

Malcolm Campbell MSc PhD
School of Nursing, Midwifery & Social Work, University of Manchester, Manchester, UK

Accepted for publication 18 August 2010

Correspondence
Dr Malcolm Campbell
School of Nursing, Midwifery & Social Work
Jean McFarlane Building
University of Manchester
Manchester, M13 9PL, UK
E-mail: malcolm.campbell@manchester.ac.uk

What is known about this topic

- Common statistical faults recur in manuscripts submitted to *HSCC* reporting quantitative research.
- Even when statistical methods are appropriate, the statistical presentation of results is often poor or does not conform to our house style.
- Shortcomings in the description, application and presentation of statistical methods slow down the peer-review process.

What this paper adds

- This paper gives a detailed explanation of the structure, layout and content for quantitative research manuscripts expected by *HSCC*.
- It warns against inappropriate statistical methods, advises on statistical reporting and suggests ways of presenting tables of results.
- Authors following these guidelines will improve the quality of their manuscripts and facilitate the peer-review process.

Abstract

As Statistical Editor of *Health & Social Care in the Community (HSCC)*, I statistically and editorially review all manuscripts reporting empirical quantitative research before they are accepted for publication. Some common faults in statistical analysis occur from time to time, but the problem that causes most work in statistical review lies in the reporting. By expanding our author guidelines for reporting quantitative research studies in *HSCC*, this article aims to help prospective authors with the planning and writing of their manuscripts and facilitate the peer review process. It gives a more detailed explanation of the structure, layout and content expected for manuscripts reporting the results of empirical quantitative research. The article includes general advice on the reporting of statistical findings, outlines common statistical faults seen in submitted manuscripts, and covers the reporting of commonly-used methods with examples of good practice from *HSCC* and suggested layouts for tables. Guidance on how to submit a paper to *HSCC* is also included. Authors intending to submit quantitative manuscripts to *HSCC* are strongly encouraged to read our author guidelines and this article to help them structure, write and submit their manuscripts. While this article goes into considerable detail, prospective authors may only need to study relevant sections to improve their manuscripts..

Keywords: Quantitative Analysis, Quantitative Research, Statistical Analysis, Statistical Methods, Statistical Reporting, Statistics

Introduction

Health & Social Care in the Community (*HSCC*) is an international interdisciplinary journal. It is read by and welcomes submissions from researchers in a variety of fields, including nursing, social work, physiotherapy, occupational therapy, general practice, health psychology, health economy, primary health-care, health promotion and professional or patient education. As Statistical Editor, I statistically and editorially review all manuscripts reporting substantial empirical quantitative research before they are accepted for publication in the journal.

A number of common statistical errors frequently recur in manuscripts submitted to *HSCC*. Many authors fail to justify their sample size or include dates of data collection, and most are not aware that *HSCC* expects dates of data collection in the Abstract. Few manuscripts submitted to *HSCC* report on randomised studies but those that do often incorrectly compare groups at baseline using statistical tests. More manuscripts report on observational designs that inevitably involve confounding, where the association between an explanatory variable and an outcome variable is affected by the behaviour of other variables. Authors should be aware that analyses involving only pairs of variables may only

provide limited explanations. There is often a lack of appreciation of the impact of missing data on findings, whether due to non-participation during initial recruitment, attrition during a prospective study or the exclusion of incomplete data in an analysis involving several variables where each is required to be non-missing. Continuous variables are often inappropriately dichotomised for analysis because this seemed 'easier' than using the continuous versions, and authors tend to apply Pearson chi-squared tests when one of the variables is ordinal and a chi-squared test for trend could be used instead. Several authors give test results without the context given by descriptive statistics; some report *p*-values without test statistics or report ranges of *p*-values when a glance at papers in *HSCC* over the last 10 years shows that we require both test statistics and actual *p*-values to be reported. Odds ratios are still commonly misinterpreted as 'number of times more likely' and authors tend not to include a practical interpretation of findings especially when the number used for analysis is relatively large or small. Such errors will be addressed later in this article.

In my experience, statistical methods applied in submitted manuscripts are generally correct, but the manuscripts tend to fall down on their general presentation of results. Prospective authors can find our house style for reporting quantitative research from recent copies of the journal and our existing guidelines to authors, as shown on the inside back cover of the journal and online at the journal's website (www.wileyonlinelibrary.com/journal/hsc), but evidence in submitted manuscripts suggests that they do not.

This article expands our previous author guidelines for reporting quantitative research studies in *HSCC* to help prospective authors with the planning and writing of their manuscripts and facilitate the peer-review process. It gives a more detailed explanation of the structure, layout and content expected for manuscripts reporting the results of empirical quantitative research, with guidance on how to submit a manuscript to *HSCC*. It also discusses common statistical faults seen in submitted manuscripts, and includes general advice on the reporting of statistical findings and suggested layouts for tables.

Authors intending to submit quantitative manuscripts to *HSCC* are strongly encouraged to read this article to help them structure, write and submit their manuscripts. While it goes into considerable detail, prospective authors may only need to study relevant sections to improve the reporting of their research.

Guidelines in the literature for reporting quantitative research

There are several established guidelines for reporting empirical quantitative research in the biomedical field (Table 1).

Leading biomedical journals require that submitted manuscripts follow the Uniform Requirements for

Table 1 Published guidelines for reporting quantitative research in biomedical journals

Guideline	Authors	Study design	Link
Statistical reporting in articles for medical journals	Bailar & Mosteller (1988)	All	http://www.people.vcu.edu/~albest/Guidance/guidelines_for_statistical_reporting.htm
Uniform requirements	ICJME (2008)	All	http://www.icmje.org/
CONSORT statements	Schulz *et al.* (2010) Moher *et al.* (2010)*	Parallel-group RCT	http://www.consort-statement.org/
	Campbell *et al.* (2004a)	Cluster RCT	
	Piaggio *et al.* (2006)	Non-inferiority and equivalence RCT	
	Gagnier *et al.* (2006)	Herbal intervention RCT	
	Boutron *et al.* (2008a) Boutron *et al.* (2008b)*	Non-pharmacological intervention RCT	
STARD statement	Bossuyt *et al.* (2003a) Bossuyt *et al.* (2003b)*	Diagnostic accuracy/ agreement	http://www.stard-statement.org/
TREND statement	Des Jarlais *et al.* (2004)	Non-randomised intervention trial	http://www.trend-statement.org/, or http://www.cdc.gov/trendstatement/
STROBE statement	von Elm *et al.* (2007) Vandenbroucke *et al.* (2007)*	Cohort study, case–control study and cross-sectional survey	http://www.strobe-statement.org/

ICJME, International Committee of Medical Journal Editors; RCT, randomised controlled trial.
Many of the guidelines are accessible via the Equator Network website (http://www.equator-network.org/).
*Elaboration and explanation.

Manuscripts Submitted to Biomedical Journals developed by the International Committee of Medical Journal Editors (ICMJE 2008, see http://www.icmje.org/journals.html for a list of supporting journals). The Uniform Requirements cover ethical considerations in the conduct and reporting of research, the editorial and peer-review process, and the general preparation and submission of manuscripts. As regards specific study designs, leading journals require that manuscripts reporting the results of randomised intervention trials should conform to the appropriate CONSORT statement (Table 1), those reporting non-randomised intervention studies should conform to the TREND statement, and those reporting observational designs such as cohort studies, case–control studies and surveys should conform to the STROBE statement.

Acknowledging these guidelines, Lang (2006) gives an excellent summary of the important topics to be addressed in each section of a paper. While the guidelines address the problems of structure and general content, they do not give sufficient guidance on writing style or the presentation of statistical results. Hall (1998) and Peat *et al.* (2002) give advice on scientific writing. Altman *et al.* (2000) give general advice on presenting statistical findings, while Lang & Secic (1997) and Peacock & Kerry (2007) advise on how to present the results of specific analyses. Nichol & Pexman (1999) give advice on presenting findings in tables by statistical method and Nichol & Pexman (2003) give advice on statistical graphics, following the guidelines of the American Psychological Association (2001).

Structure and content of quantitative manuscripts for *HSCC*

Compliance with published guidelines

Quantitative manuscripts submitted to *HSCC* should comply with the Uniform Requirements (ICMJE 2008). Those reporting randomised intervention trials must conform to the appropriate CONSORT statement (e.g. Schulz *et al.* 2010 for parallel-group designs and Boutron *et al.* 2008a for the non-pharmacological interventions commonly encountered in community health and social care). Those reporting non-randomised intervention studies or other observational designs should conform to the TREND statement (Des Jarlais *et al.* 2004) or STROBE statement (von Elm *et al.* 2007) respectively.

Use of English

The language of publication for *HSCC* is English, and authors must write to a good standard that will be understood by a general reader of the journal. Authors should remember that they are writing for an international audience from a range of disciplines; clarity and readability are important, so jargon and abbreviations should be avoided. At the same time, authors need to give sufficient detail to demonstrate the rigour of their study design and analysis so that their conclusions may be interpreted appropriately.

The use of non-discriminatory language is encouraged, and spelling should agree with that in the *Concise Oxford English Dictionary*. Authors for whom English is a second language should have their manuscript professionally edited by an English-speaking person before submission to make sure the English is of high quality. The journal's publishers, Wiley–Blackwell, have supplied a list of independent suppliers of editing services (http://authorservices.wiley.com/bauthor/english_language.asp). All services are arranged by and paid for the authors, and use of these services does not guarantee acceptance or preference for publication.

Title Page

The Title Page for *HSCC* must contain the title of the manuscript, the names and qualifications of the authors, their affiliations and the full postal address, and email address and telephone number of an author to whom correspondence may be addressed. For *HSCC*, the title is limited to a maximum of 30 words. It should clearly indicate the subject matter, but be brief and interesting in order to attract attention (Lilleyman 1998). It should include terms that researchers may be searching for electronically but should not include any abbreviations or acronyms. If the study design is not included in the title (this is recommended for randomised controlled trials), it must be described in the Abstract. If the findings of a study are likely to be specific to the location of the study, then the location should be included in the title.

Abstract

The Abstract for *HSCC* should be non-structured, not exceeding 300 words, and followed by up to six key words. Many researchers, particularly those searching for papers relevant to their own topic, may only read the Abstract, so it is important that it summarises the important details and conveys the key messages of the study.

For quantitative research manuscripts, the Abstract should include study objectives; the type of study design; the location, setting and dates of data collection; the selection and number of participants with a participation rate if appropriate; details of any interventions, instruments and key outcome measures; a summary of

the main findings with key numerical details if possible; and interpretation, conclusions and implications for research, practice or policy. Dates of data collection are particularly important in health and social care research, as they may indicate whether the data and implications are up-to-date. The Abstract should not contain references. Although the Abstract submitted to *HSCC* should not itself be structured, one way of helping to make sure that it contains all relevant information is to draft it with structured headings, and remove the headings before submission.

The journal's publishers, Wiley–Blackwell, recommend that authors optimise their Abstracts for search engines to increase the chances of their papers being selected and read, and eventually cited in another paper. Their suggestions, including the choice of a clear descriptive title and the repetition of key phrases in the Abstract are available online (http://authorservices.wiley.com/bauthor/seo.asp).

The Abstract should be accompanied by up to 6 keywords, up to 3 bullet points on ''What is already known about this topic'', and up to 3 bullet points on ''What this paper adds'', with a total of no more than 110 words across all bullet points.

Main text

In *HSCC*, the main text of quantitative research manuscripts should follow the widely used Introduction–Methods–Results-and-Discussion (IMRAD) format, which reflects the quantitative research process (ICMJE 2008).

Introduction

The Introduction should give the background to the study, explaining what is known about the subject, what is not known and why the study was performed. It may include a brief literature review, but this should focus on material relevant to the research question being addressed. The Introduction should *not* include any findings or conclusions.

The general aims of the study may be stated at the end of the Introduction. Reporting guidelines in the literature do not agree on the placement of specific aims or objectives. The Universal Requirements (ICMJE 2008) and, for observational studies, STROBE (von Elm *et al.* 2007, Vandenbroucke *et al.* 2007) require them to be in the Introduction; for randomised and non-randomised intervention studies, CONSORT (Schulz *et al.* 2010, Moher *et al.* 2010) and TREND (Des Jarlais *et al.* 2004) require them to be in the Methods section.

Methods

To comply with the Universal Requirements (ICMJE 2008), the Methods section for *HSCC* should only cover information known before the study starts; information known after data collection, such as the number of participants and their characteristics, should be reported in the Results section.

The Methods should clearly state the type of study design adopted, together with the setting, location and whether informed consent and ethical approval were obtained. Instruments used to collect data should be described, together with details of their validation and how they were administered in the study (e.g. by interview, by post, via diaries, etc.). Eligibility criteria for participants, and details of their source and access should be given, and the method of sampling from underlying populations (e.g. consecutive, random, stratified, etc.) should be stated. For randomised studies, to conform to CONSORT requirements, methods used to generate randomisation sequences and steps taken to conceal allocation should be described, as should details of who generated the allocation sequence, who enrolled the participants and who assigned them to groups. For surveys, the sampling frame and details of how probability samples were obtained should be described. Details of any interventions and how they were applied should be given. Any outcome measures, important explanatory variables, potential confounders and subgroups should be clearly identified.

Except for studies limited by resources or by time, the sample size should be justified by authors to demonstrate that the study was planned to recruit sufficient participants to answer the research question. Such reporting may both convince and educate the reader. Randomised studies should be powered according to the primary analysis (e.g. Campbell *et al.* 1995, 2004b, Whitley & Ball 2002, Schulz & Grimes 2005) while sample sizes for other study designs are often based on the number needed for the estimation of a key parameter such as prevalence (Hicks 2004) or conservative rules-of-thumb for relevant multivariate analyses (e.g. Green 1991, Peduzzi *et al.* 1996, MacCallum *et al.* 1999, Tabachnick & Fidell 2001). Machin *et al.* (2009) cover a variety of designs, give copious advice and include a software package to perform sample size calculations. There are also free software packages that may be downloaded from the Internet, such as the comprehensive G*Power (http://www.psycho.uni-duesseldorf.de/abteilungen/aap/gpower3/). A number of papers separately address the sample size needed for pilot studies (Lancaster *et al.* 2004, Julious 2005, Hertzog 2008).

Dates of recruitment or data collection must be given. For designs such as cohort studies, case–control studies and surveys, they should be given in the Methods (von Elm *et al.* 2007, Vandenbroucke *et al.* 2007). For randomised and non-randomised intervention studies, they

should be given in the Results (Des Jarlais *et al.* 2004, Moher *et al.* 2010, Schulz *et al.* 2010).

The Methods must include a description of the statistical methods used, with references for those that may be unfamiliar to readers of the journal. It is conventional to include a paragraph heading such as *Data analysis* or *Statistical analysis* near the end of the Methods section, followed by one or more paragraphs where the methods are concisely described. If relevant, authors should indicate whether any of the analyses were decided upon after inspecting the data (Vandenbroucke *et al.* 2007). The significance level for hypothesis testing should be formally stated, even if it is the conventional $\alpha = 0.05$.

Where appropriate, authors should describe how explanatory and outcome variables were included in analyses. Authors should assess any underlying assumptions of the statistical methods where possible. It is not necessary to report these results in detail, only that the relevant assumptions were examined, with a statement of the general findings.

The name and version of any software used for data analysis must be given. For *HSCC*, there is no need to provide references for commonly used statistical software such as SAS, SPSS, Stata or StatsDirect. Authors should be aware, for instance, that SPSS no longer stands for 'Statistical Package for the Social Sciences': it is a trademarked product name in its own right (https://www.spss.com/corpinfo/faqs.htm). At Release 18, after an acquisition by IBM, the software name was changed to IBM SPSS 18.

The Methods should *not* include any findings. However, details justifying the use of certain instruments or statistical methods, such as sample estimates of the internal consistency of scales and subscales, may be reported in the Methods to aid readability.

Results

The Results section should concentrate on a factual account of the findings with a minimum of interpretation. The Results should report the number of participants and the participation or response rate where applicable at different stages in the study. CONSORT and TREND strongly recommend and STROBE suggests the use of a flow diagram showing the numbers of participants at different stages. This can simplify the description and reduce the word count. This information should then be followed by a description of characteristics of the participants, broken down by group if necessary, with comparative information on non-participants where possible so that the reader may assess the extent and possible effect of non-participation.

The Results section should then contain, where applicable, descriptions of baseline values of key variables and any preliminary analyses. These should be followed by detailed results of primary analyses as determined by the aims and design of the study. These may range from estimates of key percentages in prevalence studies, estimates of outcome measures by group and between-group comparisons in randomised trials, to measures and tests of association with outcome measures that are both unadjusted and adjusted for values of other variables in observational designs. Confidence intervals should be given for main findings where possible; effect sizes may also be presented. These would then be followed by results of any secondary analyses such as analyses of secondary outcomes, subgroup analyses, or sensitivity analyses used to investigate the robustness of findings to the choice of analysis, the presence of missing data or potential bias (Vandenbroucke *et al.* 2007). Planned analyses and exploratory analyses should be clearly identified. The level of detail to be provided in all but the main analyses is left to the judgement of the authors, depending on the number of analyses, their importance and their impact on the central topic of the manuscript. If, for example, sensitivity analyses come to the same general conclusions as the main analysis, then it will be sufficient to say that the analyses were performed, without going into detail; otherwise appropriate details of the analyses should be given (Vandenbroucke *et al.* 2007) and the inconsistencies should be considered in the Discussion section.

The Results section should not repeat detailed results that are summarised separately in tables. The section may often be made more readable by reporting most of the dense numerical information in tables, leaving the text as a narrative overview to guide the reader through the important findings. *P*-values may be interpreted as a measure of the evidence against a null hypothesis (Kirkwood & Sterne 2003). Authors are reminded that when they are reporting the strict statistical significance of a result, they should be comparing the observed *p*-value against the chosen significance level α and not against other *p*-values. Non-significant test results should be reported in full for primary analyses, with observed *p*-values.

Discussion

The Discussion should contain a detailed interpretation of the findings and comparisons with those from other studies relevant to the research question. The Discussion must not include findings not previously reported in the Results section. While pointing out the strengths of the study, the Discussion should also include the authors' critical reflection on potential limitations of the study that may affect the validity or generalisabilty of the findings, such as problems due to sampling, recruitment, dropout, data collection or data analysis. The section should end with conclusions and implications for practice, policy or further research.

References

References for *HSCC* must be in Harvard style. In the text, the authors' surnames should be cited followed by the date of publication, with an ampersand between names when there are two authors, e.g. Peacock & Kerry (2007). Where there are three or more authors, the first author's name is followed by *et al.* in the text, e.g. von Elm *et al.* (2007). When multiple references are cited within the manuscript in parentheses, they are listed in chronological order and separated by commas, e.g. (Lancaster *et al.* 2004, Grissom & Kim 2005, Lang 2006).

The reference list should be in alphabetical order, each reference including (where relevant): authors' surnames and initials; year of publication in parentheses; title of article with name of journal and volume number, or title of book with edition, editors, publisher and place of publication. Up to seven authors may be included for a reference in the reference list; for references with eight or more authors, the first three authors should be named followed by *et al.* in the reference list. For journal articles, page numbers should be included routinely; page numbers from books need only be included when referencing direct quotes or paraphrases. Unpublished work should be cited in the text only. Only references to articles or books genuinely in press should be included in the reference list.

Software such as EndNote or Reference Manager may be used to manage and format references. Both support the reference style for *HSCC* – see http://www.endnote.com/support/enstyles.asp (search for the journal name) and http://www.refman.com/support/rmstyles.asp (filter by 'Public Health' to access 'Health & Social Care').

Tables and figures

Tables and figures should be numbered independently in order of reference from the text (Table 1, Table 2, etc. and Figure 1, Figure 2, etc.). For *HSCC*, they should not be embedded in the main text but included one-per-page in numerical order at the end of the manuscript following the References. Each table or figure must be cited in the text [e.g. 'see Table 1' or '(Table 1)'] with instructions on its approximate placement (e.g. 'Please insert Table 1 about here') inserted on a new line after the relevant paragraph. All tables and figures must be essential and must not be repeated in detail in the text.

Submitting a manuscript to *HSCC*

Manuscripts must be submitted online to *HSCC* via the website at http://mc.manuscriptcentral.com/hscc in Word 2003 format (.doc) or Rich Text Format (.rtf). Files submitted in Word 2007 format (.docx) are currently automatically rejected. As noted above, the main manuscript should contain a title, an Abstract, keywords and bullet points, the main text, References, a list of titles for tables and figures, and the tables and figures themselves. Special graphical format files are not usually required for statistical figures – images from statistical packages pasted into Word are generally acceptable.

As recommended by the Uniform Requirements (ICMJE 2008), manuscripts should be double-spaced with generous margins to facilitate reviewing and editing. The journal does not support footnotes in the text, although footnotes are allowed for tables and figures. For *HSCC*, the word count for the main text should not exceed 5000 words; currently, there is no limit on the number of tables or figures, but only essential ones should be included.

Before the main manuscript is uploaded, authors are asked to enter a title of up to 30 words and an Abstract of up to 300 words in text boxes. Text may be copied and pasted into the text boxes from a Word document. During the submission process, authors will subsequently be asked to upload at least two files to the online editorial system: a Title Page (containing the title of the manuscript; names, qualifications and affiliations of the authors; and the full postal address, email address and telephone number of the corresponding author) and a Main Document (containing the main manuscript including the Abstract, but without any author details to allow blinded peer review). Where necessary, other files may be uploaded.

Inappropriate methods of analysis

Vandenbroucke *et al.* (2007) summarise a central tenet of data analysis very neatly: 'In general, there is no one correct statistical analysis but, rather, several possibilities that may address the same question, but make different assumptions.' Altman *et al.* (2000) give general advice on statistical analysis and presenting findings, while Kirkwood & Sterne (2003) include a useful chapter on strategies for analysis. Most quantitative manuscripts submitted to *HSCC* describe the results of appropriately chosen statistical analyses. However, there are a number of inappropriate methods that occur again and again.

Comparing groups at baseline following randomisation

Relatively few manuscripts submitted to *HSCC* involve randomised trials, but most of those that do mistakenly include results of statistical tests to compare characteristics of the randomised groups at baseline. This is despite

the practice being indicated as inappropriate on the CONSORT website (CONSORT item 15, http://www.consort-statement.org/consort-statement/) and the explanation and elaboration document (Moher *et al.* 2010) that authors should be following. Statistical tests test hypotheses about underlying populations, and the null hypotheses concerning no differences between groups should be automatically true following randomisation. Randomisation should have no impact on the sampling process, so that randomised groups can be considered to be samples from the same population. Where appropriate, analysis of outcome measures post-intervention should take account of baseline values to improve statistical power (Vickers & Altman 2001).

Confounding in observational studies

Most quantitative manuscripts submitted to *HSCC* report the results of observational study designs such as cross-sectional surveys and longitudinal cohort studies. Such designs inevitably involve confounding, where the observed association between an explanatory variable and the outcome variable is partly due to other factors associated with both variables that are not on the causal path (Vandenbroucke *et al.* 2007). Kirkwood & Sterne (2003) describe how confounding variables should be carefully chosen; in many situations, important confounders cannot be measured, so it may be necessary to interpret the results with caution. Authors should be aware that analyses limited to pairs of variables, such as correlations and *t*-tests, that do not take other variables into account may only provide limited explanations. Regression analyses, such as multiple linear regression, logistic regression or Cox regression, may be used to adjust for confounding variables, but authors should be careful not to word their interpretation to conclude causality purely from the statistical results (Vandenbroucke *et al.* 2007).

Assessing underlying assumptions

Authors submitting manuscripts to *HSCC* often fail to comment on whether underlying assumptions were satisfied. Introductory statistical textbooks such as Bland (2000), Kirkwood & Sterne (2003) and Field (2009) explain the assumptions underlying simple parametric and nonparametric tests. Tabachnick & Fidell (2001) comprehensively describe the assumptions underlying linear regression, logistic regression, survival analysis and factor analysis, noting that multicollinearity can affect the stability of estimates in all of those methods, not just in linear regression. Authors should therefore check for evidence of multicollinearity when their main analysis involves linear regression, logistic regression, a survival analysis regression or factor analysis. Normal distributions are assumed for certain statistical methods, although in many situations, the methods work well under minor departures from the assumption. Pett (1997) and Tabachnick & Fidell (2001) discuss how to assess the assumption of a Normal distribution in some detail. Authors sometimes mistakenly assume that individual variables should have Normal distributions for paired *t*-tests and linear regression. For paired *t*-tests, it is the difference variable that should be (approximately) Normal; for linear regression, the residuals should be (approximately) Normal. We do not require full details of the assessment of assumptions to be reported in the manuscripts; a statement that they have been investigated and a general comment on the findings is usually sufficient.

Building regression models

The aim in regression modelling is to find an interpretable model explaining the associations between several explanatory or independent variables and an outcome or dependent variable. Perhaps because they are readily available in the software and widely reported in the literature, some authors fitting regression models uncritically employ automated variable selection techniques based on purely statistical criteria (e.g. stepwise regression) to find a solution involving a small number of explanatory variables. Such methods have been widely criticised as producing optimistically biased *p*-values, confidence intervals for coefficients and goodness-of-fit statistics (e.g. R^2); estimates of coefficients that are too large in absolute terms; and models that can be difficult to interpret (Thompson 1995, Harrell 2001, Miles & Shevlin 2001, Cohen *et al.* 2003, Babyak 2004). Kirkwood & Sterne (2003) point out that in observational studies, it is essential to control for sensibly chosen confounding variables, and that these cannot be determined by purely statistical criteria. A simulation study indicated that automated variable selection tended to be poor at selecting authentic predictors and good at selecting irrelevant 'noise' variables (Derksen & Keselman 1992). While we would consider articles where automatic variable selection has been applied in an exploratory manner, regression models that have been developed from theoretical models reflecting causal priority, confounding and relevance to the research are more likely to produce meaningful results (Cohen *et al.* 2003). Manuscripts where models have been developed more thoughtfully will be of more interest to the reader and are more likely to be accepted. Such approaches include: hierarchical modelling, where the variables are introduced in blocks in a logical or causal order (Miles

& Shevlin 2001, Tabachnick & Fidell 2001); moderation, where a moderator variable affects the causal relationship between an exploratory variable and an outcome variable (http://davidakenny.net/cm/moderation.htm); and mediation, where an exploratory variable has an effect on a mediator variable, which in turn has an effect on an outcome variable (http://davidakenny.net/cm/mediate.htm) (Baron & Kenny 1986, Miles & Shevlin 2001). Miles & Shevlin (2001) and Kirkwood & Sterne (2003) give good advice on the handling of interactions in modelling. Authors are reminded that statistical models simplify reality, and that they have to describe and interpret their findings clearly and meaningfully to the reader. Kirkwood & Sterne (2003) suggest leaving the analysis of interactions to the final stage of modelling.

Analysing designs with clustering

Occasionally, a manuscript will report on a study design that includes clustering of participants, either naturally through cluster sampling or experimentally within a cluster-randomised trial. When participants in the same cluster are likely to be more similar than participants from different clusters, the analysis should take the clustering into account. Ignoring the clustering, as authors often do, results in confidence intervals that are narrower and *p*-values that are smaller than they should be, potentially leading to incorrect conclusions of significance (Kerry & Bland 1998). Kirkwood & Sterne (2003) describe appropriate methods of analysis, including estimating a summary measure for each cluster and analysing at the cluster level using standard methods, using robust standard errors that are corrected for clustering, using random effects or multilevel models to model similarity within clusters, and using generalised estimating equations to adjust standard errors and estimated parameters for clustering.

Assessing effects of missing data

A more common fault in submitted manuscripts is a lack of appreciation of the impact of missing data on findings, whether due to non-participation during initial recruitment, attrition during a prospective study or the exclusion of incomplete data in an analysis involving several variables (Streiner 2002). In their Discussion, where authors consider potential limitations to their study, they should reflect on the potential impact of missing data on the sample used for their main analyses and possibly also on the results of those analyses. The number of missing values across key variables should be reported (Vandenbroucke *et al.* 2007). Methods for handling missing data exist (Wood *et al.* 2004, McKnight *et al.* 2007), but none give a perfect solution as they depend on assumptions that it may not be possible to assess. In manuscripts submitted to *HSCC*, the most common approach is a 'complete case analysis', where only participants with complete data across the variables considered are included in the main analyses. However, this may introduce a bias if the subset of participants included differs in important respects from the subset not included, and it is inefficient in not using all the available data (Vandenbroucke *et al.* 2007). Authors using complete case analysis should compare characteristics of participants used for analysis with characteristics of those excluded, or against an appropriate reference sample or population. This may identify potential bias but not necessarily the reasons why data are missing (Streiner 2002). There are advanced methods for dealing with missing data such as linear mixed models (Mallinckrodt *et al.* 2008) and multiple imputation (Sterne *et al.* 2009), but currently these are not widely used by authors submitting to *HSCC*. Authors using such methods must explain any assumptions being made about why data may be missing, and compare the results against those of a complete case analysis in a sensitivity analysis, with any important differences being discussed (Vandenbroucke *et al.* 2007). Detailed advice on the handling of missing values can be found at the Missing Data website (http://www.missingdata.org.uk/).

Categorising continuous data

Many authors readily convert values of continuous variables into two or more categories, either for clinical reasons or for convenience in analysis. While this may be acceptable to simplify a descriptive presentation, it involves a major loss of precision and information, with very different values at either end of a continuous interval being treated as equivalent and similar values either side of a cut-off point being treated as very different.

Altman (2000), MacCallum *et al.* (2002) and Babyak (2004) describe the problems that may arise when continuous variables are grouped into categories for statistical analysis, such as a loss in statistical power, reduced effect sizes and biased estimates in regression. Dichotomisation or categorisation may be acceptable if the intention of the analysis is to reflect recognised cut-points, such as those defining 'normal' and 'abnormal' ranges for psychological measures, or if the continuous variable does not have a simple, e.g. linear, relationship with an outcome. Dichotomisation based on sample means or medians should be avoided as this makes it difficult to compare results across studies and categorisation into several ordered categories is usually better (Altman & Royston 2006). Authors who intend to categorise continuous

variables are strongly encouraged to assess its impact on their analyses.

Authors submitting to *HSCC* routinely test the association between an ordinal variable and a dichotomous variable using Pearson's chi-squared test. Pointing out that this test does not take the order of categories into account, Bland (2000) describes a suitable alternative, the chi-squared test for trend. His advice is to use Pearson's chi-squared test when order does not matter and the chi-squared test for trend when it does.

Interpreting odds ratios

Authors submitting to *HSCC* commonly misinterpret odds ratios, such as those resulting from logistic regression, as relative risks or risk ratios, for example, mistakenly stating that an odds ratio of 2.0 means 'twice as likely'. Kirkwood & Sterne (2003) describe the difference between odds ratios and relative risks, noting that numerically an odds ratio is always further away from 1.0 than the corresponding relative risk, and showing that an odds ratio of 2.0 can correspond to a relative risk anywhere between 1.0 and 2.0 depending on the prevalence of the outcome. An odds ratio of 2.0 should be interpreted as meaning that the *odds* of an outcome are twice as high compared with the odds in a reference category. As an example, Klinkenberg *et al.* (2005) investigated care received and place of death among the 342 members of a representative cohort of older people in the Netherlands who died between 1995 and 1999. The authors fitted several logistic regression models to estimate associations between various factors and whether a person died at home or in a hospital, including a simple model with two covariates, care arrangement (no care, formal care, informal care, and both formal and informal care) and region (West, South and North-East Holland). The adjusted odds ratios for region were reported in table 6 of their paper as West: 2.00 (95% C.I. 1.00 to 4.06), South: 3.27 (95% C.I. 1.42 to 7.52), North-East: 1.00 (the reference category). Adjusted for care arrangements, the authors concluded that 'For sample members in the Western part of the country, the odds of dying at home were twice those for persons in the North-East. For people in the South, the odds were over three times as high'.

Interpreting statistical significance and non-significance

Authors should remember that non-significant *p*-values do not show that there is 'no difference' or 'no association' present – absence of evidence is not evidence of absence (Altman & Bland 1995). This applies whether they occur in the results of simple hypothesis tests, like *t*-tests and chi-squared tests, or more complex analyses, such as regression models. As Bland (2000) points out, it is not possible to prove statistically that no effect exists. It is more correct to conclude that there is no evidence of a difference or association. Many authors mistakenly believe that non-significant test results should not be reported. If the test was considered to be important, the test results should be reported in full, including the value of the observed *p*-value.

Authors often forget to include a practical interpretation of findings especially when the number used for analysis is relatively large or small: for the reader's benefit, interpretation needs to go beyond the reporting of *p*-values. Small studies may be underpowered to detect important differences or associations as being statistically significant; large studies may detect statistical significance when observed differences or associations are small and unimportant (Sheldon 2000). Authors should reflect and comment whether observed confidence intervals or effect sizes have a practical importance.

General advice on statistical reporting

The advice on statistical reporting given here is mainly specific to *HSCC*, but the basic principles apply to other journals.

Statistical keywords, acronyms and symbols

Several commonly used words have a particular statistical meaning. Authors should avoid using words such as the following except in their statistical sense: *association*, *correlation*, *dependent*, *effect*, *incidence*, *independent*, *normal*, *outlier*, *parameter*, *population*, *power*, *prevalence*, *random*, *range*, *reliability*, *sample*, *sensitivity*, *significance/significant*, *specificity*, *validity* and *variance*. There are also many widely accepted statistical abbreviations and symbols (Lang & Secic 1997), which can be used without additional explanation (examples are given in Tables 2 and 3).

Reporting numbers

Peat *et al.* (2002) give 'golden rules for reporting numbers'. The convention is to report whole numbers <10 in words unless they are part of a list or if they have units (e.g. 5 cm). Whole numbers of 10 or more are conventionally reported in digits unless they start a sentence (e.g. Twenty-five per cent of participants …). Decimal numbers <1.0 such as correlations and *p*-values should have a zero before the decimal point (e.g. $r = -0.50$, $p = 0.012$). There should be no space between a number and a percentage sign (e.g. 25%, 25.0%) but a space between a number and a unit (e.g. 10 mm, 50.0 kg), with

Table 2 Commonly used statistical abbreviations (adapted from Lang & Secic 1997)

Acronym	Meaning
ANCOVA	Analysis of covariance
ANOVA	Analysis of variance
C.I.	Confidence interval
df	Degrees of freedom
ICC	Intra-class correlation
IQR	Interquartile range
MANOVA	Multivariate analysis of variance
NNT	Number needed to treat
OR	Odds ratio
RR	Risk ratio or relative risk
SD	Standard deviation
SE	Standard error

These abbreviations may be used in the text, tables or figures without needing to be explained to the reader.

Table 3 Commonly used statistical symbols (adapted from Lang & Secic 1997)

Symbol	Meaning
B	Unstandardised regression coefficient
F	*F*-distribution, -test or -statistic, e.g. ANOVA *F*=, partial *F*=
p, *P*	Probability of test statistic being more extreme than observed value under null hypothesis, *p*-value, *P*-value, e.g. *p*= or *P*=
r, *R*	Pearson's product–moment correlation, e.g. Pearson's *r*=
r^2, R^2	Coefficient of determination
t	*t* distribution, test or statistic, e.g. *t*=, paired *t*=
U	Mann–Whitney *U*-test statistic, e.g. *U*=
Z	Statistic with standard Normal distribution, e.g. Mann–Whitney *Z*=
α	Cronbach's alpha coefficient of internal consistency, e.g. Cronbach's α=
α	Probability of Type I error (significance level)
β	Probability of Type II error (1 – statistical power)
β	Standardised regression coefficient
χ^2	Chi-squared distribution, test or statistic, e.g. χ^2= or chi-square=
κ or kappa	Cohen's kappa statistic, e.g. Cohen's κ= or Cohen's kappa=
ρ or rho	Spearman's rank order correlation, e.g. Spearman's ρ= or Spearman's rho=
τ or tau	Kendall's concordance correlation, e.g. Kendall's τ= or Kendall's tau=

measurements in SI units where appropriate (see, e.g. Baron & McKenzie Clarke 2008 and http://physics.nist.gov/cuu/Units/units.html).

The base total for a percentage should always be clear to a reader, and numbers should usually accompany percentages unless readability is compromised. Peat *et al.* (2002) suggest reporting percentages to one decimal place if the sample size is larger than 100 and to the nearest whole number if it lies between 20 and 100, with percentages not used if the sample size is <20. For readability, the same number of decimal places should be used throughout the text, although a compromise may be needed when percentages from groups of very different sizes are compared. The same number of decimal places should be used for percentages throughout the tables, although this need not be the same as the number used in the text. It is often acceptable to round percentages to the nearest whole number in the text for readability and report percentages to one decimal place in tables unless the sample size is small.

Estimates of the centre of a distribution (e.g. means and medians) should be accompanied by corresponding estimates of the variation about the centre (e.g. standard deviations and ranges or interquartile ranges). Summary statistics in the units of the measured variable (such as means, standard deviations and confidence intervals) should be given to one more decimal place than the number used for reporting values of that variable (Altman *et al.* 2000). Coefficients such as Cronbach's alpha, correlations, *R*-square and Cohen's kappa that are <1.0 in magnitude and most test statistics, relative risks and odds ratios may be given to two decimal places.

The use of the '±' symbol should be avoided in the text or tables to avoid potential confusion over what follows the symbol (Altman *et al.* 2000). Intervals such as ranges, confidence intervals and interquartile ranges should be reported using the word 'to' to separate the lower and upper limits; this avoids confusion that occurs with the use of a hyphen or dash as the separator when either limit is negative. Means and standard deviations in the text may be reported in a form such as 'mean 12.3 (SD 4.5)'. 'M', 'xbar' or 'x' should not be used as an abbreviation for 'mean'.

Reporting test results

Unless readability is affected, important test results should be reported in full, in either the text (usually within parentheses), the body of a table or the footnotes of a table. Reporting should include (where applicable) a suitably labelled test statistic with the value usually reported to two decimal places, degrees of freedom and the actual *p*-value to three decimal places or two significant figures consistently throughout the manuscript, with the *p* in lower case or upper case. *P*-value ranges such as '$p < 0.05$' and 'NS' should not be reported for individual

tests: the observed *p*-value should be given (Altman *et al.* 2000). By convention, however, *p*-values less than 0.001 should be reported as $p < 0.001$ or $P < 0.001$. A reviewer or reader may be able to assess whether an appropriate test has been used from the label of the test statistic, whether appropriate data have been analysed from the degrees of freedom, and the weight of evidence against the null hypothesis from the observed *p*-value.

In large tables, it may not be possible to report complete test results without over-complicating the table. One alternative would be to report the actual *p*-values providing it was clear which tests were being used. If this is not possible, then the level of significance may be shown conventionally using superscripted asterisks, * for $p \leq 0.05$, ** for $p \leq 0.01$ and *** for $p \leq 0.001$, described in a footnote.

Test results should not be reported in isolation: appropriate descriptive statistics should also be given to allow the reviewer or reader to make sense of the test results. Where possible and particularly for comparisons of key outcomes between groups, test results must be accompanied by relevant confidence intervals. For instance, for a comparison of the population means of two groups using a *t*-test, the following information could be reported in the text or tables: sample size, mean and standard deviation for each group; *t*-test statistic, degrees of freedom and actual *p*-value; and 95% confidence interval for the difference in means. While authors may report effect sizes where applicable, it is not currently a requirement for *HSCC*. Advice on effect sizes is given by various authors (Cohen 1988, 1992, Grissom & Kim 2005, McGrath & Meyer 2005, Field 2009) and online (http://www.uccs.edu/~faculty/lbecker/es.htm).

Tables and figures

Although simple findings may be presented in the text, tables should be used to summarise detailed information in an easy-to-follow format. Each table should have a separate function, and the contents of tables should not be repeated in the text. Tables may be useful for keeping detailed statistical findings out of the text to improve readability, so that the Results section can, as far as possible, concentrate on a narrative overview of the findings.

In a national survey of mental health social workers in the UK, Evans *et al.* (2005) investigated the impact of having approved social worker (ASW) status on work patterns and workload stresses. As examples of good practice, they presented many of their numerical findings for ASW versus non-ASW – descriptive statistics (means, standard deviations, medians and ranges) and results of *t*-tests and Mann–Whitney *U*-tests – compactly in tables.

Tables should follow a consistent layout, with, for example, groups always defining the rows or always defining the columns, so that the reader can easily follow a series of tables. The title should be at the top of a table (t for table and t for top) and titles should be informative without duplicating column headings. As emphasised by Nichol & Pexman (1999), tables should be self-contained, avoiding the need for a reader to cross-reference the text to understand a table. Any abbreviations or acronyms should be explained in full, either in the title, the body of the table or in a footnote to the table.

Figures may be used to emphasise key findings or present complicated information that would be difficult to describe easily in words or numbers. A simple diagram showing the behaviour of a single variable is unlikely to be accepted if it can be summarised numerically in less space. Two examples of diagrams that do present detailed information in an easy-to-follow graphical format are the flow diagram showing the involvement of participants at different stages in a study and Kaplan–Meier survival plots in time-to-event analyses. Advice on the design of participant flow diagrams is given in the elaboration and explanation paper for the CONSORT statement (Moher *et al.* 2010), while Pocock *et al.* (2002) give practical advice for survival plots. The title should be at the foot of a figure (f for figure and f for foot). Nichol & Pexman (2003) point out that, like tables, figures should be able to stand alone and that any abbreviations or acronyms should be explained in full, either in the title or in footnotes to the figure.

Suggestions for reporting results of advanced analyses

In addition to approaches such as *t*-tests, chi-squared tests and correlations, a range of more advanced statistical methods are widely used to analyse complex data in health and social care research. Descriptions of the methods described here are accompanied by references to papers from *HSCC* that show examples of good practice. The advice given for reporting results of these analyses is specific to *HSCC*, but it may also apply more generally. The tabular formats for reporting the results of analyses introduced here are suggestions and authors would not be expected to follow them exactly.

Analysis of variance and related methods

Analysis of variance (ANOVA) is a statistical method for comparing population means of a continuous variable in groups defined by one or more variables. Howell (2007) and Field (2009) cover the different versions of ANOVA in great detail. The most widely used version, one-way ANOVA, compares population means in three or more independent groups. Repeated measures ANOVA compares population means in dependent groups (usually

the same variable measured at three or more time points). Mixed-model ANOVA is used when there are both dependent and independent groups. Analysis of covariance (ANCOVA) extends ANOVA to adjust the comparison for other explanatory variables (covariates). ANCOVA is commonly used in the analysis of randomised trials to adjust the post-intervention comparison of a continuous variable for measurement of the variable at baseline to control for baseline imbalance and maximise the power of detecting a difference between groups (Vickers & Altman 2001).

Hacking *et al.* (2008) reported on a survey of people with mental health needs taking part in participatory art projects in England. They used mixed-model ANOVA to assess separately the changes in nine outcome measures from baseline to follow-up by subgroup. As an example of good practice, they presented descriptive statistics by group and relevant *F*-test results in a well-presented table.

Schneider *et al.* (2002) explored associations between different forms of service organisation for community mental health-care and the quality of life for service users in four English health districts. They used ANCOVA to analyse a Life Satisfaction Scale score at 6-month follow-up by service configuration, with Life Satisfaction Scale score and an Affect Balance Scale score at baseline as covariates. They reported the results of this key analysis in the form of a standard ANOVA table similar to Table 4.

The findings of ANOVA and ANCOVA are often reported compactly in the body of the text in the form of the most relevant *F*-test results, but in specific situations, results may also be reported in the form of the classical ANOVA table found in textbooks, as recommended by Lang & Secic (1997). This has columns showing sources of variation, sums of squares, mean squares, *F*-statistics, degrees of freedom and *p*-values (Table 4).

Multiple regression

Multiple linear regression is a statistical method for assessing the association between a set of explanatory variables and a continuous outcome variable. Excellent practical guidance is given by Tabachnick & Fidell (2001), Miles & Shevlin (2001) and Cohen *et al.* (2003). Nichol & Pexman (1999) suggest ways of presenting results when all explanatory variables are included simultaneously or hierarchically in blocks defined by the researcher; their hierarchical presentation may also be used for stepwise regression, where statistical criteria are used to select variables. Their tables feature both unstandardised regression coefficients (*B*) and standardised regression coefficients (β), although Greenland *et al.* (1986) warn against interpreting standardised coefficients as effect sizes. Lang & Secic (1997) recommend reporting unstandardised coefficients, standard errors, 95% confidence intervals, and test statistics and *p*-values for individual variables. More simply, Peacock & Kerry (2007) suggest reporting unstandardised coefficients, 95% confidence intervals and corresponding *p*-values.

In a study of the impact of nutrition and eating disabilities on the quality of life of stroke survivors in South London, Perry & McLaren (2004) measured quality of life using a validated measure (the Quality of Life Index/head injury version) as a continuous outcome variable. They estimated correlations between this and a number of other continuous validated measures for activities of daily living, eating disabilities, nutritional status and risk, dietary intake, cognitive function, social support, anxiety and depression. They then fitted a

Table 4 Suggested layout for table reporting key numerical results of analysis of variance or analysis of covariance (*N* = xxx)

Source of variation	Degrees of freedom	Sum of squares	Mean square	*F*-statistic	*P*-value
First main effect	x	xxx.xx	xx.xx	xx.xx	0.xxx
Second main effect	x	xxx.xx	xx.xx	xx.xx	0.xxx
…	…	…	…	…	…
Last main effect	x	xxx.xx	xx.xx	xx.xx	0.xxx
First interaction term	x	xxx.xx	xx.xx	xx.xx	0.xxx
Second interaction term	x	xxx.xx	xx.xx	xx.xx	0.xxx
…	…	…	…	…	…
Last interaction term	x	xxx.xx	xx.xx	xx.xx	0.xxx
First covariate	x	xxx.xx	xx.xx	xx.xx	0.xxx
Second covariate	x	xxx.xx	xx.xx	xx.xx	0.xxx
…	…	…	…	…	…
Last covariate	x	xxx.xx	xx.xx	xx.xx	0.xxx
Residual	xxx	xxx.xx	xx.xx		
Total*	xxx	xxx.xx	xx.xx		

*The Total row may be suppressed if desired.

stepwise multiple regression model including age, an index of local deprivation and these validated measures. As an example of good practice, they reported an informative table containing details of the stepwise selection, including changes in the coefficient of determination (R^2) with *F*-tests showing the significance of the changes. They also presented a second table giving details of the final fitted model, including estimated unstandardised regression coefficients, standard errors and 95% confidence intervals.

Table 5 suggests how important information could be presented for a standard multiple regression where all independent variables are entered into the model at the same time.

In hierarchical multiple regression, it may be useful to show the regression results at each stage, and present statistics on the change between one block and the next. Table 6 suggests a way of presenting results for this approach.

Logistic regression

Logistic regression is a statistical method for assessing the association between explanatory variables and a categorical (usually dichotomous) outcome variable. The classic practical reference for logistic regression is Hosmer & Lemeshow (2000), while Harrell (2001) and Kirkwood & Sterne (2003) also cover it very well. Nichol & Pexman (1999) suggest ways of presenting results of logistic regression, concentrating on log-odds, their standard errors and odds ratios, adjusted for other variables in the model. Lang & Secic (1997) also recommend reporting test statistics for individual variables and associated *p*-values. Log-odds are not as easy to interpret as odds ratios, and Peacock & Kerry (2007) present a simpler alternative reporting adjusted odds ratios, their 95% confidence intervals and corresponding *p*-values. They also suggest presenting unadjusted odds ratios to show how odds ratios change when adjusted for other variables, commenting that the confidence intervals for unadjusted odds ratios need not be reported if space is limited.

Peters *et al.* (2004) surveyed the use of community-based NHS services for urinary incontinence in the Bristol area of England. They used a dichotomous outcome variable taking the values yes or no, depending on whether a respondent aged 65 or over had or had not used the services within the last 6 months. They first estimated the association between individual socio-demographic, health-related, clinical and personal variables and the outcome variable using simple logistic regression models. The authors presented these results very clearly in tables reporting numbers and percentages or other appropriate descriptive statistics, odds ratios, their 95% confidence intervals and associated *p*-values for each variable in turn. They subsequently fitted logistic regression models within the four groups of variables, and then across the four groups, reporting their findings in a compact table presenting odds ratios, 95% confidence intervals and *p*-values.

Table 7 suggests how important results for logistic regression could be presented. For this table, all explanatory variables are included at the same time in the model.

In hierarchical logistic regression, as in hierarchical multiple regression, it may be useful to show the regression results at each stage, and present statistics on the change between one block and the next. Table 8 suggests a way of presenting results for hierarchical logistic regression.

Time-to-event (survival) analysis

Time-to-event analysis or survival analysis covers statistical methods used to estimate the probability of an event occurring in time, where the event may occur for some individuals and not for others. This may also include estimating the association between explanatory variables

Table 5 Suggested layout for table reporting key numerical results of standard multiple regression

Independent variable	Unstandardised coefficient (*B*)	95% C.I.	*P*-value
First continuous or dichotomous variable	x.xx	x.xx to x.xx	0.xxx
First categorical variable with more than two categories			0.xxx*
First non-reference category	x.xx	x.xx to x.xx	0.xxx
Second non-reference category	x.xx	x.xx to x.xx	0.xxx
…	…	…	…
Last non-reference category	x.xx	x.xx to x.xx	0.xxx
…	…	…	…

Reported in footnotes to table: measure of overall significance (e.g. result of regression ANOVA *F*-test) and variance explained (e.g. R^2, adjusted R^2).

*The availability of overall *p*-values for categorical variables may depend on the software and how it is used.

Table 6 Suggested layout for table reporting key numerical results of hierarchical multiple regression

	Partial model			Full model		
Independent variable	Unstandardised coefficient (*B*)	95% C.I.	*P*-value	Unstandardised coefficient (B)	95% C.I.	*P*-value
First continuous or dichotomous variable (in both models)	x.xx	x.xx to x.xx	0.xxx	x.xx	x.xx to x.xx	0.xxx
First categorical variable with more than two categories (in both models)			0.xxx*			0.xxx*
First non-reference category	x.xx	x.xx to x.xx	0.xxx	x.xx	x.xx to x.xx	0.xxx
Second non-reference category	x.xx	x.xx to x.xx	0.xxx	x.xx	x.xx to x.xx	0.xxx
…	…	…	…	…	…	…
Last non-reference category	x.xx	x.xx to x.xx	0.xxx	x.xx	x.xx to x.xx	0.xxx
…	…	…	…	…	…	…
First continuous or dichotomous variable (in full model only)				x.xx	x.xx to x.xx	0.xxx
First categorical variable with more than two categories (in full model only)						0.xxx*
First non-reference category				x.xx	x.xx to x.xx	0.xxx
Second non-reference category				x.xx	x.xx to x.xx	0.xxx
…				…	…	…
Last non-reference category				x.xx	x.xx to x.xx	0.xxx
…	…	…	…	…	…	…

Reported in footnotes to table: partial model – measure of overall significance (e.g. result of regression ANOVA *F*-test) and variance explained (e.g. R^2, adjusted R^2); full model – measure of overall significance (e.g. result of regression ANOVA *F*-test) and variance explained (e.g. R^2, adjusted R^2); differences between partial model and full model – significance of change (e.g. result of partial *F*-test) and change in variance explained (e.g. change in R^2).
*The availability of overall *p*-values for categorical variables may depend on the software and how it is used.

Table 7 Suggested layout for table reporting key numerical results of logistic regression (*N* = xxx)

	1st outcome category	2nd outcome category				
Independent variable	*N* (row%)	*N* (row%)	Unadjusted odds ratio	Adjusted odds ratio	95% C.I.	*P*-value
First continuous variable	xx	xx	x.xx	x.xx	x.xx to x.xx	0.xxx
Types of descriptive statistics*	xx.x [xx.x]	xx.x [xx.x]				
First dichotomous variable						
Reference category	xx (xx.x%)	xx (xx.x%)	1.00	1.00		
Non-reference category	xx (xx.x%)	xx (xx.x%)	x.xx	x.xx	x.xx to x.xx	0.xxx
First categorical variable with more than two categories						0.xxx‡
Reference category	xx (xx.x%)	xx (xx.x%)	1.00	1.00		
First non-reference category	xx (xx.x%)	xx (xx.x%)	x.xx	x.xx	x.xx to x.xx	0.xxx
Second non-reference category	xx (xx.x%)	xx (xx.x%)	x.xx	x.xx	x.xx to x.xx	0.xxx
…	…	…	…	…	…	…
Last non-reference category	xx (xx.x%)	xx (xx.x%)	x.xx	x.xx	x.xx to x.xx	0.xxx
…	…	…	…	…	…	…

Reported in footnotes to table: measure of overall significance (e.g. result of likelihood ratio chi-squared test) and variation or pseudo-variance explained (e.g. Nagelkerke R^2).
*For example, mean [SD] or median [range], as appropriate.
‡The availability of overall *p*-values for categorical variables may depend on the software and how it is used.

and the event rate relative to that in a reference group (as a hazard ratio, which can be interpreted as the relative risk of the event occurring). Good references for survival analysis include Tabachnick & Fidell (2001), Harrell (2001) and Kirkwood & Sterne (2003). Lang & Secic (1997) recommend presenting a Kaplan–Meier survival

Table 8 Suggested layout for table reporting key numerical results of hierarchical logistic regression

	Partial model			Full model		
Independent variable	Adjusted odds ratio	95% C.I.	*P*-value	Adjusted odds ratio	95% C.I.	*P*-value
First continuous or dichotomous variable (in both models)	x.xx	x.xx to x.xx	0.xxx	x.xx	x.xx to x.xx	0.xxx
First categorical variable with more than two categories (in both models)			0.xxx*			0.xxx*
First non-reference category	x.xx	x.xx to x.xx	0.xxx	x.xx	x.xx to x.xx	0.xxx
Second non-reference category	x.xx	x.xx to x.xx	0.xxx	x.xx	x.xx to x.xx	0.xxx
…	…	…	…	…	…	…
Last non-reference category	x.xx	x.xx to x.xx	0.xxx	x.xx	x.xx to x.xx	0.xxx
…	…	…	…	…	…	…
First continuous or dichotomous variable (in full model only)				x.xx	x.xx to x.xx	0.xxx
First categorical variable with more than two categories (in full model only)						0.xxx*
First non-reference category				x.xx	x.xx to x.xx	0.xxx
Second non-reference category				x.xx	x.xx to x.xx	0.xxx
…				…	…	…
Last non-reference category				x.xx	x.xx to x.xx	0.xxx
…	…	…	…	…	…	…

Reported in footnotes to table: partial model – measure of overall significance (e.g. result of likelihood ratio chi-squared test) and variation or pseudo-variance explained (e.g. Nagelkerke R^2); full model – measure of overall significance (e.g. result of likelihood ratio chi-squared test) and variation or pseudo-variance explained (e.g. Nagelkerke R^2); difference between partial model and full model – significance of change (e.g. result of likelihood ratio chi-squared test).
*The availability of overall *p*-values for categorical variables may depend on the software and how it is used.

plot showing the estimated probability over time of survival (the event not happening) for each group. Pocock *et al.* (2002) give practical advice for the presentation of such plots. Lang & Secic (1997) also recommend reporting median survival times and survival rates for each group (both with 95% confidence intervals), and results of either the log-rank test or Breslow's test to compare survival curves between groups. When Cox proportional hazards regression is used to estimate the association between explanatory variables and event rates, Lang & Secic (1997) recommend reporting regression coefficients, standard errors, test statistics and *p*-values, hazard ratios and their 95% confidence intervals for each explanatory variable. Peacock & Kerry (2007) suggest reporting Kaplan–Meier survival plots, survival rates and results of the log-rank test, but simplify the reporting of Cox regression results to unadjusted and adjusted hazard ratios with 95% confidence intervals.

Trappes-Lomax *et al.* (2006) reported on a non-randomised controlled trial evaluating the effectiveness of a short-term residential rehabilitation facility for patients aged 55 or over discharged from community hospitals in two areas of Devon, England. The researchers were interested in two time-to-event outcomes within a 12-month follow-up period: the primary outcome was days until a participant died or went into residential or nursing care, and the secondary outcome was days until a participant was readmitted to hospital. The authors reported their time-to-event findings in a clear format. For each outcome, they presented a Kaplan–Meier survival curve and survival rates for each group, an unadjusted hazard ratio and a hazard ratio adjusted for key baseline covariates (age, sex, whether the participant lived alone, days in hospital at baseline and level of dependency).

Median survival times by group could be presented as shown in Table 9 (Peat & Barton 2005), while survival rates by group at a key follow-up time point could be presented as shown in Table 10 (Lang & Secic 1997). The results of Cox proportional hazards regression could be presented as suggested in Table 11, following Peacock & Kerry (2007).

Factor analysis

Factor analysis covers a variety of statistical techniques for identifying small numbers of hidden factors that may underlie multivariate data. It can be a difficult analysis to conduct, but Pett *et al.* (2003) give excellent practical guidance, as do Fabrigar *et al.* (1999), Tabachnick &

Table 9 Suggested layout for table reporting survival times* by group (N = xxx)

Group‡	Number at risk	Number of events	Number censored‡	Median survival time*	95% C.I.
Total sample	xxx	xxx	xxx	xx.x‡	xx.x§ to xx.x§
First group	xxx	xxx	xxx	xx.x	xx.x to xx.x
Second group	xxx	xxx	xxx	xx.x	xx.x to xx.x
…	…	…	…	…	…
Last group	xxx	xxx	xxx	xx.x	xx.x to xx.x

*Death may not always be the event of interest, so the wording may need to be changed appropriately.
‡The table will be easier to a reader to interpret if shortened meanings are given for groups rather than 'Group 1', etc.
‡For example, lost to follow-up.
§Values of survival times in days may be rounded to the nearest whole number.

Table 10 Suggested layout for table reporting survival rates* by group at given time point‡ (N = xxx)

Group‡	Number at risk	Number of events	Number censored§	Survival rate*	95% C.I.
Total sample	xxx	xxx	xxx	xx.x%	xx.x% to xx.x%
First group	xxx	xxx	xxx	xx.x%	xx.x% to xx.x%
Second group	xxx	xxx	xxx	xx.x%	xx.x% to xx.x%
…	…	…	…	…	…
Last group	xxx	xxx	xxx	xx.x%	xx.x% to xx.x%

*Death may not always be the event of interest, so the wording may need to be changed appropriately.
‡Two or more time points may be important for a study.
‡The table will be easier to a reader to interpret if shortened meanings are given for groups rather than 'Group 1', etc.
§For example, lost to follow-up.

Table 11 Suggested layout for table reporting results of Cox proportional hazards regression (N = xxx)

Independent variable	Unadjusted hazard ratio	Adjusted hazard ratio	95% C.I.	*P*-value
First continuous variable	x.xx	x.xx	x.xx to x.xx	0.xxx
First dichotomous variable				
Reference category	1.00	1.00		
Non-reference category	x.xx	x.xx	x.xx to x.xx	0.xxx
First categorical variable with more than two categories				0.xxx*
Reference category	1.00	1.00		
First non-reference category	x.xx	x.xx	x.xx to x.xx	0.xxx
Second non-reference category	x.xx	x.xx	x.xx to x.xx	0.xxx
…	…	…	…	…
Last non-reference category	x.xx	x.xx	x.xx to x.xx	0.xxx
…	…	…	…	…

*The availability of overall *p*-values for categorical variables may depend on the software and how it is used.

Fidell (2001) and Costello & Osborne (2005). Pett *et al.* (2003) also describe the important information that should be reported for factor analysis. This may include: how variables were initially chosen and the types of variables involved; a preliminary assessment of the suitability of factor analysis; estimated correlations and measures of sampling adequacy; methods used for extracting and rotating factors; how the number of factors was determined and which factor solutions were explored; estimated communalities showing the variance explained for each variable; the variance explained for each factor; rotated factor loadings; Cronbach's alpha coefficient of internal consistency for factor-based subscales; and a practical interpretation of the rotated factors. Pett *et al.* (2003) give various suggestions for ways of displaying important information in tables.

McCrone *et al.* (2006) applied factor analysis in the development of an index of mental health needs of a Local Authority in England. They considered different combinations of extraction and rotation methods and different factor solutions. As an example of good practice, they reported loadings for a four-factor solution for a principal components analysis with Varimax rotation, with loadings <0.4 suppressed for clarity, together with the percentage of variation explained by each factor. They also reported the Local Authorities with the three highest and lowest scores and gave a reasoned interpretation of the factors.

Tables 12 and 13 suggest possible layouts for reporting the important numerical results from a factor analysis. In Table 12, all factor loadings are presented, with those for items considered to be primarily loading on each factor shown in bold. In Table 13, only the largest factor loadings are presented; loadings for items not

Table 12 Suggested layout for table reporting key numerical results of factor analysis with more important loadings in bold (N = xxx)

Item*	Final communalities	Rotated factor loadings			
		Factor 1	Factor 2	…	Factor *k*
First item primarily loading on factor 1	0.xx	**0.xx**	0.xx	…	0.xx
Second item primarily loading on factor 1	0.xx	**0.xx**	0.xx	…	0.xx
…	…	…	…	…	…
Last item primarily loading on factor 1	0.xx	**0.xx**	0.xx	…	0.xx
First item primarily loading on factor 2	0.xx	0.xx	**0.xx**	…	0.xx
Second item primarily loading on factor 2	0.xx	0.xx	**0.xx**	…	0.xx
…	…	…	…	…	…
Last item primarily loading on factor 2	0.xx	0.xx	**0.xx**	…	0.xx
…	…	…	…	…	…
First item primarily loading on factor *k*	0.xx	0.xx	0.xx	…	**0.xx**
Second item primarily loading on factor *k*	0.xx	0.xx	0.xx	…	**0.xx**
…	…	…	…	…	…
Last item primarily loading on factor *k*	0.xx	0.xx	0.xx	…	**0.xx**
Percentage of variance explained		xx.x%	xx.x%	…	xx.x%
Cronbach's alpha for subscale score		0.xx	0.xx	…	0.xx

*Items sorted by highest loadings on factor 1, then on factor 2, and so on. The table will be easier to a reader to interpret if shortened meanings are given for items rather than 'Item 1', etc.

Table 13 Suggested layout for table reporting key numerical results of factor analysis with less important loadings suppressed (N = xxx)

Item*	Final communalities	Rotated factor loadings			
		Factor 1	Factor 2	…	Factor *k*
First item primarily loading on factor 1	0.xx	0.xx			
Second item primarily loading on factor 1	0.xx	0.xx			
…	…	…			
Last item primarily loading on factor 1	0.xx	0.xx			
First item primarily loading on factor 2	0.xx		0.xx		
Second item primarily loading on factor 2	0.xx		0.xx		
…	…		…		
Last item primarily loading on factor 2	0.xx		0.xx		
…	…			…	
First item primarily loading on factor *k*	0.xx				0.xx
Second item primarily loading on factor *k*	0.xx				0.xx
…	…				…
Last item primarily loading on factor *k*	0.xx				0.xx
Percentage of variance explained		xx.x%	xx.x%	…	xx.x%
Cronbach's alpha for subscale score		0.xx	0.xx	…	0.xx

*Items sorted by highest loadings on factor 1, then on factor 2, and so on. The table will be easier to a reader to interpret if shortened meanings are given for items rather than 'Item 1', etc.

considered to be primarily associated with factors are suppressed (the loadings are not zero or missing but omitted to improve clarity).

Summary

Despite the considerable time and effort that can be involved in the collection and analysis of data and the drafting and re-drafting of manuscripts, the publication of research findings is a necessary but rewarding process. Researchers have an ethical obligation to their funding body and those who took part in the study to publish their results in a timely manner. While publications can enhance career prospects, reputations and the likelihood of further funding, they also extend knowledge in their chosen area by disseminating answers to research questions, even though this may only seem like adding a tiny brick to a huge wall. Writing a research paper need not be a daunting task, given the many sources of advice available to authors, from the Uniform Requirements and accepted standards such as CONSORT, to guidelines to authors and examples of papers in journals. One other important source of advice deserves a mention: the sterling work performed during the peer-review process by reviewers, who can spend many hours or days constructively commenting on a manuscript and suggesting ways the research or the reporting can be strengthened. Many authors submitting to *HSCC* have commented on the usefulness of feedback from reviewers in helping them improve their manuscripts. We all have the same aims: a better understanding of and an improvement in health and social care in the community.

Authors intending to submit manuscripts to *HSCC* reporting the results of quantitative research are recommended to study the author guidelines and this article. Manuscripts conforming to the expected structure where the statistical reporting follows the advice presented are likely to have a smoother and consequently faster path through the peer-review process. While specific to *HSCC*, the general principles may also be applied to statistical reporting in dissertations, theses and reports, and to manuscripts submitted to other journals. Journals other than *HSCC* may require different structures for manuscripts or adopt different styles for reporting statistics, and authors are strongly advised to consult relevant author guidelines and copies of recently published papers in those journals.

Acknowledgements

I would like to thank the anonymous reviewers for their helpful and constructive comments.

References

Altman D.G. (2000) Statistics in medical journals: some recent trends. *Statistics in Medicine* **19**, 3275–3289.

Altman D.G. & Bland J.M. (1995) Statistics notes: absence of evidence is not evidence of absence. *BMJ* **311**, 485.

Altman D.G. & Royston P. (2006) Statistics notes: the cost of dichotomising continuous variables. *BMJ* **332**, 1080.

Altman D.G., Gore S.M., Gardner M.J. & Pocock S.J. (2000) Statistical guidelines for contributors to medical journals. In: D.G. Altman, D. Machin, T.N. Bryant and M.J. Gardner (Eds) *Statistics with Confidence*, 2nd edn. BMJ Books, London.

American Psychological Association (2001) *Publication Manual of the American Psychological Association*. American Psychological Association, Washington, DC.

Babyak M.A. (2004) What you see may not be what you get: a brief, nontechnical introduction to overfitting in regression-type models. *Psychosomatic Medicine* **66**, 411–421.

Bailar J.C. III & Mosteller F. (1988) Guidelines for statistical reporting in articles for medical journals: amplifications and explanations. *Annals of Internal Medicine* **108**, 266–273 [WWW document]. URL http://www.people.vcu.edu/~albest/Guidance/guidelines_for_statistical_reporting.htm

Baron R.M. & Kenny D.A. (1986) The moderator-mediator variable distinction in social psychological research: conceptual, strategic and statistical considerations. *Journal of Personality and Social Psychology* **51**, 1173–1182.

Baron D.N. & McKenzie Clarke H. (2008) *Units, Symbols, and Abbreviations: A Guide for Authors and Editors in Medicine and Related Sciences*, 6th edn. Royal Society of Medicine Press, London.

Bland M. (2000) *An Introduction to Medical Statistics*, 3rd edn. Oxford University Press, Oxford.

Bossuyt P.M., Reitsma J.B., Bruns D.E. *et al.* (2003a) Towards complete and accurate reporting of studies of diagnostic accuracy: the STARD initiative. *Clinical Chemistry* **49** (1), 1–6 [WWW document]. URL http://www.stard-statement.org/

Bossuyt P.M., Reitsma J.B., Bruns D.E. *et al.* (2003b) The STARD statement for reporting studies of diagnostic accuracy: explanation and elaboration. *Clinical Chemistry* **49** (1), 7–18 [WWW document]. URL http://www.stard-statement.org/

Boutron I., Moher D., Altman D.G., Schulz K. & Ravaud P. (2008a) Methods and processes of the CONSORT group: example of an extension for trials assessing nonpharmacologic treatments. *Annals of Internal Medicine* **148** (4), W60–W67.

Boutron I., Moher D., Altman D.G., Schulz K. & Ravaud P. (2008b) Extending the CONSORT statement to randomized trials of nonpharmacologic treatment: explanation and elaboration. *Annals of Internal Medicine* **148** (4), 295–309.

Campbell M.J., Julious S.A. & Altman D.G. (1995) Estimating sample sizes for binary, ordered categorical, and continuous outcomes in two group comparisons. *BMJ* **311**, 1145–1148.

Campbell M.K., Elbourne D.R. & Altman D.G. (2004a) CONSORT statement: extension to cluster randomised trials. *BMJ* **328** (7441), 702–708.

Campbell M.K., Thomson S., Ramsay C.R., MacLennan G.S. & Grimshaw J.M. (2004b) Sample size calculator for cluster randomized trials. *Computers in Biology and Medicine* **34**, 113–125.

Cohen J. (1988) *Statistical Power Analysis for the Behavioral Sciences*, 2nd edn. Lawrence Erlbaum Associates, Hillsdale, NJ.

Cohen J. (1992) A power primer. *Psychological Bulletin* **112** (1), 155–159.

Cohen J., Cohen P., West S.G. & Aiken L.S. (2003) *Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences*, 3rd edn. Lawrence Erlbaum Associates, Mahwah, NJ.

Costello A.B. & Osborne J.W. (2005) Best practices in exploratory factor analysis: four recommendations for getting the most from your analysis. *Practical Assessment, Research & Evaluation* **10** (7) [WWW document]. URL http://pareonline.net/pdf/v10n7.pdf

Derksen S. & Keselman H.J. (1992) Backward, forward and stepwise automated subset selection algorithms: frequency of obtaining authentic and noise variables. *British Journal of Mathematical and Statistical Psychology* **45**, 265–282.

Des Jarlais D.C., Lyles C., Crepaz N. & The TREND Group (2004) Improving the reporting quality of nonrandomized evaluations of behavioral and public health interventions: the TREND statement. *American Journal of Public Health* **94** (3), 361–366.

von Elm E., Altman D.G., Egger M., Pocock S.J., Gotzsche P.C. & Vandenbroucke J.P. (2007) Strengthening the reporting of observational studies in epidemiology (STROBE) statement: guidelines for reporting observational studies. *Preventative Medicine* **45**, 247–251.

Evans S., Huxley P., Webber M. *et al*. (2005) The impact of 'statutory duties' on mental health social workers in the UK. *Health & Social Care in the Community* **13** (2), 145–154.

Fabrigar L.R., Wegener D.T., MacCallum R.C. & Strahan E.J. (1999) Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods* **4** (3), 272–299.

Field A. (2009) *Discovering Statistics Using SPSS*, 3rd edn. SAGE Publications, London.

Gagnier J.J., Boon H., Rochon P., Moher D., Barnes J. & Bombardier C. (2006) Reporting randomized, controlled trials of herbal interventions: an elaborated CONSORT statement. *Annals of Internal Medicine* **144** (5), 364–367.

Green S.B. (1991) How many subjects does it take to do a regression analysis? *Multivariate Behavioral Research* **26** (3), 499–510.

Greenland S., Schlesselmann J.J. & Criqui M.H. (1986) The fallacy of employing standardized regression coefficients and correlations as measures of effect. *American Journal of Epidemiology* **123** (2), 203–208.

Grissom R.J. & Kim J.J. (2005) *Effect Sizes for Research – A Broad Practical Approach*. Lawrence Erlbaum Associates, Mahwah, NJ.

Hacking S., Secker J., Spandler H., Kent L. & Shenton J. (2008) Evaluating the impact of participatory arts projects for people with mental health needs. *Health & Social Care in the Community* **16** (6), 638–648.

Hall G.M. (Ed.) (1998) *How to Write a Paper*, 2nd edn. BMJ Books, London.

Harrell F.E. Jr (2001) *Regression Modeling Strategies: With Applications to Linear Models, Logistic Regression, and Survival Analysis*. Springer-Verlag, New York.

Hertzog M.A. (2008) Considerations in determining sample size for pilot studies. *Research in Nursing & Health* **31**, 180–191.

Hicks C.M. (2004) *Research Methods for Clinical Therapists*, 4th edn. Churchill Livingstone, Edinburgh.

Hosmer D.W. & Lemeshow S. (2000) *Applied Logistic Regression*, 2nd edn. John Wiley, New York.

International Committee of Medical Journal Editors (2008) *Uniform requirements for manuscripts submitted to biomedical journals: writing and editing for biomedical publication* [WWW document]. URL http://www.icmje.org/

Julious S.A. (2005) Sample size of 12 per group rule of thumb for a pilot study. *Pharmaceutical Statistics* **4**, 287–291.

Kerry S. & Bland J.M. (1998) Analysis of a trial randomised in clusters. *BMJ* **316**, 54.

Kirkwood B.R. & Sterne J.A.C. (2003) *Essential Medical Statistics*, 2nd edn. Blackwell Science, Oxford.

Klinkenberg M., Visser G., Broese van Groenou M.I., van der Wal G., Deeg D.J.H. & Willems D.L. (2005) The last 3 months of life: care, transitions and the place of death of older people. *Health & Social Care in the Community* **13** (5), 420–430.

Lancaster G.A., Dodd S. & Williamson P.R. (2004) Design and analysis of pilot studies: recommendations for good practice. *Journal of Evaluation in Clinical Practice* **10** (2), 307–312.

Lang T. (2006) Documenting research in scientific articles: guidelines for authors. *Chest* **130**, 1263–1268.

Lang T.A. & Secic L. (1997) *How to Report Statistics in Medicine*. American College of Physicians, Philadelphia, PA.

Lilleyman J.S. (1998) Titles, abstracts and authors. In: G.M. Hall (Ed.) *How to Write a Paper*, 2nd edn. BMJ Books, London.

MacCallum R.C., Zhang S., Preacher K. & Hong S. (1999) Sample size in factor analysis. *Psychological Methods* **4** (1), 84–99.

MacCallum R.C., Widaman K.F., Zhang S., Preacher K. & Rucker D.D. (2002) On the practice of dichotomization of quantitative variables. *Psychological Methods* **7** (1), 19–40.

Machin D., Campbell M.J., Tan S.B. & Tan S.H. (2009) *Sample Size Tables for Clinical Studies*, 3rd edn. Wiley-Blackwell, Chichester.

Mallinckrodt C.H., Lane P.W., Schnell D., Peng Y. & Mancuso J.P. (2008) Recommendations for the primary analysis of continuous endpoints in longitudinal clinical trials. *Drug Information Journal* **42**, 303–319.

McCrone P., Thornicroft G., Boyle S., Knapp M. & Aziz F. (2006) The development of a Local Index of Need (LIN) and its use to explain variations in social services expenditure on mental health care in England. *Health & Social Care in the Community* **14** (3), 254–263.

McGrath R.E. & Meyer G.J. (2005) When effect sizes disagree: the case of r and d. *Psychological Methods* **11** (4), 386–401.

McKnight P.E., McKnight K.M., Sidani S. & Figueredo A.J. (2007) *Missing Data: A Gentle Introduction*. The Guildford Press, New York.

Miles J. & Shevlin M. (2001) *Applying Regression and Correlation: A Guide for Students and Researchers*. SAGE Publications, London.

Moher D., Hopewell S., Schulz K.F. *et al*. (2010) CONSORT 2010 Explanation and Elaboration: updated guidelines for reporting parallel group randomised trials. *BMJ* **340**, c869.

Nichol A.A.M. & Pexman P.M. (1999) *Presenting Your Findings: A Practical Guide for Creating Tables*. American Psychological Association, Washington, DC.

Nichol A.A.M. & Pexman P.M. (2003) *Displaying Your Findings: A Practical Guide for Creating Figures, Posters, and Presentations*. American Psychological Association, Washington, DC.

Peacock J. & Kerry S. (2007) *Presenting Medical Statistics from Proposal to Publication: A Step-by-step Guide*. Oxford University Press, Oxford.

Peat J. & Barton B. (2005) *Medical Statistics: A Guide to Data Analysis and Critical Appraisal*. BMJ Books, London.

Peat J., Elliott E., Baur L. & Keena V. (2002) *Scientific Writing: Easy When You Know How*. BMJ Books, London.

Peduzzi P., Concato J., Kemper E., Holford T.R. & Feinstein A.R. (1996) A simulation study of the number of events per variable in logistic regression analysis. *Journal of Clinical Epidemiology* **49** (12), 1373–1379.

Perry L. & McLaren S. (2004) An exploration of nutrition and eating disabilities in relation to quality of life at 6 months post-stroke. *Health & Social Care in the Community* **12** (4), 288–297.

Peters T.J., Horrocks S., Stoddart H. & Somerset M. (2004) Factors associated with variations in older people's use of community-based continence services. *Health & Social Care in the Community* **12** (1), 53–62.

Pett M.A. (1997) *Nonparametric Statistics for Health Care Research*. SAGE Publications, Thousand Oaks, CA.

Pett M.A., Lackey N.R. & Sullivan J.J. (2003) *Making Sense of Factor Analysis: The Use of Factor Analysis for Instrument Development in Health Care Research*. SAGE Publications, Thousand Oaks, CA.

Piaggio G., Elbourne D.R., Altman D.G., Pocock S.J. & Evans S.J.W. (2006) Reporting of noninferiority and equivalence randomized trials: an extension of the CONSORT statement. *JAMA* **295**, 1152–1160.

Pocock S.J., Clayton T.C. & Altman D.G. (2002) Survival plots of time-to-event outcomes in clinical trials: good practice and pitfalls. *Lancet* **359**, 1686–1689.

Schneider J., Wooff D., Carpenter J., Brandon T. & McNiven F. (2002) Community mental healthcare in England: associations between service organisation and quality of life. *Health & Social Care in the Community* **10** (6), 423–434.

Schulz K.F. & Grimes D.A. (2005) Sample size calculations in randomised trials: mandatory and mystical. *Lancet* **365**, 1348–1353.

Schulz K.F., Altman D.G. & Moher D. for the CONSORT group (2010) CONSORT 2010 Statement: updated guidelines for reporting parallel group randomised trials. *BMJ* **340**, 698–702.

Sheldon T. (2000) Estimating treatment effects: real or the result of chance. *Evidence Based Nursing* **3**, 36–39.

Sterne J.A.C., White I.R., Carlin J.B. *et al.* (2009) Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. *BMJ* **338**, b2393.

Streiner D.L. (2002) The case of the missing data: methods of dealing with dropouts and other research vagaries. *Canadian Journal of Psychiatry* **47**, 68–75.

Tabachnick B.G. & Fidell L.S. (2001) *Using Multivariate Statistics*, 4th edn. Allyn & Bacon, Needham Heights, MA.

Thompson B. (1995) Stepwise regression and stepwise discriminant analysis need not apply here: a guidelines editorial. *Educational and Psychological Medicine* **55**, 525–534.

Trappes-Lomax T., Ellis A., Fox M., Taylor R., Power M., Stead J. & Bainbridge I. (2006) Buying Time I: a prospective, controlled trial of a joint health/social care residential rehabilitation unit for older people on discharge from hospital. *Health & Social Care in the Community* **14** (1), 49–62.

Vandenbroucke J.P., von Elm E., Altman D.G. *et al.* (2007) Strengthening the reporting of observational studies in epidemiology (STROBE): explanation and elaboration. *Epidemiology* **18** (6), 805–835.

Vickers A.J. & Altman D.G. (2001) Analysing controlled trials with baseline and follow up measurements. *BMJ* **323**, 1123–1124.

Whitley E. & Ball J. (2002) Statistics review 4: sample size calculations. *Critical Care* **6**, 335–341.

Wood A.M., White I.R. & Thompson S.G. (2004) Are missing outcome data adequately handled? A review of published randomized controlled trials in major medical journals. *Clinical Trials* **1**, 368–376.