Invited commentary

How to publish (more) effectively?

Cezary Watala

Department of Haemostatic Disorders, Medical University of Lodz, Poland

Submitted: 13 December 2005
Accepted: 13 December 2005

Arch Med Sci 2005; 1, 4: 201-204

Corresponding author:
Prof. Cezary Watała
Department of Haemostatic Disorders
Medical University of Lodz
Medical University Hospital No. 2
113 Zeromskiego
90-549 Lodz, Poland
E-mail: cwatala@csk.umed.lodz.pl

The presentation by Kenneth Dickstein titled "Hints and Tricks in Scientific Publications" is a good and brief guide for everybody interested in an effective scientific writing. It should be certainly forwarded to all younger and less experienced researchers, but also to the majority of experienced and successful science workers. It is really worth reading and thinking over.

The presentation is good, because it emphasizes all the crucial aspects of logical and reasonable scientific writing and – in turn – putatively effective publishing. The presentation touches the most important – but how seldom respected in a whole – commitments how to arrange, organize and finally polish our publication prepared for submission. It essentially consists of a set of useful hints on what to take care of and what to avoid. Importantly, it presents the reviewer's and editor's views on our submission, and states what is of crucial importance for their first impression on our work: it actually states how to increase the chance of our publishing success.

The presentation is brief but exhaustive: it is like an extract of the magnificent tutorials by Stephan Day and others on “How to write and publish a scientific paper”. It may not be considered a replacement for these books, but rather used as a synoptic manual summarizing the most important conceptual ideas of scientific writing, improving and revising. The last one, so frequently underestimated and neglected by young researchers, is often of a definitive importance for the final editorial decision on our manuscript. Our replies to the Reviewers' comments are sometimes regarded as a much better hallmark of our research experience and proficiency than the manuscript itself. Therefore, do not skimp your time when revising!

The basic difficulty in scientific writing for young researchers is to prepare an outline of a final manuscript. In fact, as they usually do not know how to cope with this problem, they simply neglect it, obviously unaware that good manuscripts are nearly always generated from good outlines. What an outline presents is a logical structure of the arguments supposed to appear in our paper, with advantages and disadvantages or gaps of our study. The outline should address the answers or comments on the following topics/problems:

a) What is the central message of our paper? It does not have to be long; one sentence stating of what are you trying to say in your manuscript is often enough.
b) What were the basic question(s) and problem(s) that “triggered” you to start the study? What was known before you started the study? Is your study a repetitive one or completely novel? What answers were needed to address the problem(s)? – preferably list the key point(s). What was your approach to answer the question(s)?
c) What is the population of elements which you worked with? No matter whether humans, animals or crystals of a solid chemical – they represent the group of objects that you studied in a hope that your conclusions would refer not only to these particular objects enrolled, but they would be wider,

more universal, more representative to all other elements of the same or similar type. You need to take care of how your studied population represents the general one. To be entitled to argue on that you need to employ the appropriate sampling procedure. Give the details on how you did it, what are the methods used to conduct the study, who was the object of your study; in one word: define materials and methods.

d) Define your principal findings and results. The most important ones are probably encapsulated within your central message, but list all of them that you feel ought to be included.
e) Make a brief list of your conclusions and the implications of the study. What is new and why do you think it matters? State the limitations. How far – you think – your outcomes might be of an impact for possible changes in the future in approaches, techniques, clinical practice etc.?
f) Create a logical tree of ideas and conclusions, arranged chronologically or by order of importance. The resulting structure should be iterative or at least recurrent – a given key point or idea should be derived of the earlier discussed issue. Identify the conclusions that contribute to each key point and define whether literature supporting your principal findings and conclusions exists.

There are plenty of detailed important issues listed above that are worth a more profound discussion, however, the correct presentation of the most original part of our paper – our results – is certainly in front of others. The correct, regular description of our results in a manuscript is not an easy task at all. One of the major obstacles is how to reasonably arrange the text of our manuscript describing our findings in such a way to guarantee the clarity and legibility of this section and to disclose the impact of our outcomes to the potential reader. Below is a brief summary of what we should remember to state when describing our experimental approach and findings in various types of clinical studies.

1. How to report the employed research design and how to justify its use?

Introduction

a) state the purpose of the study; identify the relationships that were studied and the reasons for studying them
b) if the study was designed to test one or more a priori hypotheses, state these hypotheses
c) state (not "describe") how the original data were obtained

Materials and methods – what was the study focused on?

a) specify whether the study was observational (retrospective, prospective) or experimental (randomized controlled trials); define units of interest
b) describe the population studied and the population to which the results are to be generalized
c) provide definitions for all explanatory variables (independent, contributory, dummy, concomitant, risk factors, predictive, prognostic etc) and all response variables (dependent, endpoints, outcomes)
d) specify the minimum change or difference in the response variable(s) that is considered to be ***clinically*** (not "statistically") important
e) indicate whether the study was approved (with respect to human or animal subjects)

Materials and methods – how was the study planned?

a) describe the study design
b) describe fully the treatment under study and the protocol under which it was administered
c) if the groups are to be paired, report criteria and rationale for such a pairing
d) describe any potential confounding variables and report how you controlled them
e) specify how the sample size was determined; give accepted power and significance
f) specify the sampling technique(s)
g) give the inclusion and exclusion criteria
h) describe the circumstances of how the informed consent was obtained
i) specify how the participants were assigned to experimental groups (control or treatment)
j) if applicable, describe the technique of masking (blinding)
k) describe placebo medications, 'sham' procedures or any alternative or concomitant treatments received by control group
l) describe how you collected the data
m) describe the planned nature and duration of the study
n) describe – if any – quality-control methods used to ensure completeness and accuracy of data collection
o) describe the administrative structure of multicenter trials

Description of statistical methods

a) describe the comparisons made and the statistical procedures used
b) state whether the analysis is on the basis of intention-to-treat or other method
c) describe any planned interim analyses and any stopping rules for the study
d) specify any procedures used to control for the multiple testing problem
e) report the levels of α (inference) and β (statistical power) errors

f) report whether the statistical tests were one- or two-tailed
g) identify the statistical program used to analyze the data

Results – describe the study as it was conducted

a) specify the beginning and ending dates of the data collection; give the reasons for selecting those dates
b) when applicable provide a summary (e.g. schematic) showing the number and disposition of participants at each stage
c) if applicable, describe the subjects who were eligible and available but were not approached to participate in the study
d) if applicable, describe the subjects who were evaluated for participation but who finally did not meet the inclusion criteria
e) describe the participants who were enrolled in the study but who did not complete it (drop-outs, withdrawals)
f) describe the participants who completed the treatment but who were lost to follow up
g) report the duration and nature of follow-up
h) describe the participants who completed the course of treatment and the follow-up examination(s)
i) indicate how representative the sample is for the population of interest
j) indicate the differences and similarities between control and treatment groups at baseline
k) indicate whether masking and allocation concealment were successful
l) for studies based on judgments from a few evaluators, provide a measure of consistency or agreement among the evaluators (interobserver variability)

Results – study outcomes

a) report absolute (and relative) changes or differences for all primary endpoints
b) report 95% confidence intervals for these changes or differences in primary endpoints
c) report *p* values for all primary analyses
d) present results of the study in figures or tables, whichever more appropriate
e) report statistical findings with enough detail to allow subsequent reanalysis or meta-analysis
f) report any potential confounding or interactive effects
g) indicate the degree to which study participants adhered to the protocol and explain any possible exceptions or deviations from the protocol
h) report potential treatment-related side effects and adverse events
i) report how outlying values were treated
j) explain any missing data

Discussion

a) first, describe the implications of the primary analyses
b) distinguish between statistical significance and clinical importance
c) discuss the results in the context of published literature
d) comment on generalizability of the results
e) state the limitations of the study; discuss any weaknesses in the research design or problems with data collection, analysis or interpretation
f) limit conclusions to those supported by the study; omit speculations

2. How to report data and descriptive statistics?

a) report all numbers with the appropriate degree of precision
b) when reporting percentages, always give the numerator and denominators of the calculations
c) when sample size >100, report percentages with the precision to no more than one decimal point, when sample size is less than 100, report percentages in whole numbers; when sample size is very low (<20), report actual numbers instead of percentages
d) when reporting changes in data as percent changes, use the formula: [final value – initial value)/initial value] x100%
e) when summarizing categorical data, specify the denominators of rates, ratios, proportions, and %
f) if continuous data have been separated by "cut-off-points" into ordinal categories, give the values of cut-off-points and the rationale for choosing them
g) when summarizing continuous data (with continuous distribution), provide appropriate measures of central tendency and dispersion
h) do not use mean and standard error to show within-population variability
i) use mean and SD only when describing approximately normally distributed data; identify the meaning of the used interval when reported for the first time
j) when comparing variabilities of two or more groups, use coefficients of variability rather than SD
k) report non-normally distributed (skewed) data with median and range or interquartile range (or other interpercentile range)
l) paired observations should be reported together
m) indicate whether and how markedly the non-normally distributed data were transformed into an approximately normal distribution; if data were transformed, convert units of measurement back to the original units for reporting
n) for small samples, present all data if appropriate, especially in cases when the descriptive statistics would be misleading; do not use percentages for small samples

3. How to report the outcomes of logistic regression analysis?

Logistic regression analysis is a model of multivariate statistics, very commonly used when we are interested for instance, to define whether some variables may be considered independent risk factors. In this approach we aim at prediction of one (binary) categorical response variable from one (simple logistic regression) or more (multiple logistic regression) explanatory variables.

a) describe the relationship of interest or the purpose of the analysis
b) identify the variables used in the comparison and summarize each with descriptive statistics
c) confirm that the assumptions of simple logistic regression analysis were met and state how each was checked
d) specify how the explanatory variables that appear in the final model were chosen
e) specify whether the potential explanatory variables were assessed for correlation or association
f) specify whether the independent variables were tested for interaction (tolerance)
g) summarize logistic regression equation in a table; include the number of observations in the analysis, the coefficient of the explanatory variable, and the associated standard error; the odds ratio with 95% confidence interval, and p value
h) specify whether the model was validated (does it make sense that a given explanatory variable predicts outcome, endpoint, etc.?)
i) report how any outlying data were treated in the analysis
j) give the name of statistical software used for calculations

4. How to report survival (time-to-event) analysis?

a) describe the studied relationship(s) and the reasons for studying it/them
b) describe clinical characteristics of the studied population
c) specify the starting and ending times that mark the beginning of the study and its termination
d) specify the nature of censored data
e) specify the statistical methods used to estimate the survival rate
f) confirm that requirements for survival analysis have been met
g) for each group, give the estimated survival rate at appropriate follow-up times, with confidence intervals, and the number of participants at risk of death at each time
h) specify the statistical methods used to compare two or more survival curves; report the actual *p* value of comparison
i) report the regression model used to assess the associations between survival rate and explanatory variables; report a measure of risk for each explanatory variable
j) describe the quality of life for survivals
k) if applicable, present full results in table(s)/graph