

Publication More than Once: Duplicate Publication and Reuse of Text

Elise Langdon-Neuner, MSc*

Editor-in-chief of The Write Stuff (The Journal of the European Medical Writers Association), Vienna, Austria.

Publication of articles in biomedical journals disseminates the results of scientific research, which ultimately advances knowledge and improves public health. Publication also benefits the authors and the pharmaceutical companies who sponsor research. The benefit to authors is in career promotion and securing funding. The benefit to pharmaceutical companies is influencing doctors to prescribe their products. It follows that the greater the number of publications the greater will be the benefit to authors and sponsors. This scenario drives temptation to publish data from the same research study more than once. Consequently republication of entire articles or information contained in articles is not uncommon.^{1,2} The problem is such publication can be unethical and against the public interest, as it was in the second study described below. It also wastes the time of editors and reviewers in processing the papers as well as readers' time in reading what they have read before.

Two studies in particular highlight the beneficial (data dissemination and enhanced reception of publications) and the bad (inflation of results that may endanger patients) effects of multiple publications. In the first study,³ the original research articles published by applicants for research fellowships were evaluated. The aim was to test the extent to which the number of publications from a single research project affected the reception of these publications in the scientific community as signalled by citations. The study found that citation counts for articles increased significantly with each article published. Furthermore, the longer the articles, the more the citation count increased with the number of papers published. Short articles provide other scientists with less content to quote. The investigators concluded that researchers benefit from publishing more than one paper per study. This study did not distinguish between articles containing data that overlapped with the other publications and articles containing different data from the same study.

The second study¹ examined the influence of covert

duplicate publication on meta-analysis. It looked at data from 9 clinical trials conducted on the drug ondansetron published in full reports of randomized controlled trials which tested the drug's efficacy. Twenty-three reports were found that contained duplicate data, and 21 of these had no cross-reference to any of the other reports. This resulted in data from the same patients being analysed more than once in the meta-analysis. 17% of the reports and 28% of the patient efficacy data were duplicated. Inclusion of the duplicated data in meta-analysis led to a 23% overestimation of the drug's efficacy, which is clearly contrary to public interest.

Two methods have been used by authors to increase the number of publications from a single study. One is republication of papers that are identical to or similar to the original paper reporting the same body of research. The other is separate publications reporting parts of a single piece of research. In this second method, authors break down their work into what has been called the least (minimum) publishable units. The distinction between the two methods can be blurred but the hallmark of dishonest authors in both cases is the omission of cross-citations to the other publication(s), i.e. covert duplication.

There is not always a clear distinction between these two forms and the literature is complicated by the variety of terms that are used to describe republication. The first form of republication mentioned above is usually called duplicate publication and the second, when the division is inappropriate, is called salami publication. This article will use the term divided publication for the second type to denote both ethical and unethical instances. Other terms that refer to one or both types include: redundant, repetitive, overlapping, multiple, dual and prior. One study described 6 different types: republication of a published article with 1) identical sample and outcome, 2) a different sample and outcome, 3) different outcomes from the same sample, 4) the same outcome but reporting only part of a larger trial, 5) new

*Corresponding Author: Elise Langdon-Neuner, Editor-in-chief of The Write Stuff (The Journal of the European Medical Writers Association), Hauptstrasse 23, 3422 Altenberg, Austria. Tel: +43 1 20100 2067. Fax: +43 1 20100 525. Email: a.a.neuner@gmail.com.



data added to a preliminary study and 6) an assembly of two or more articles to form another article.⁴

Many, but not all, journals take their editorial policies from the guidelines laid down by the International Committee of Medical Journal Editors (ICMJE, www.icmje.org). Section III.D. gives guidance to journals on policies relating to overlapping publications under the heading 'Redundant Publication'. Duplicate publication according to the ICMJE is when a paper is published which overlaps substantially with another one that has already been published in print or electronic media. The degree of overlap that constitutes 'substantial' is not defined.

The New England Journal of Medicine (NEJM) attempts to define 'substantial' by contrasting huge clinical trials where it is legitimate to describe important outcomes separately, e.g. the Framingham Heart Study, which could not have been reported in a single paper, with reports of studies that only have few (they say several dozen) patients.⁵ It states that reports of studies with few patients should not be split into different manuscripts. A specific example given by the NEJM relates to a controlled intervention at a birthing centre where the editors believed that results on the mothers and those on the infants should not have been sent to different journals. There is also comment on instances when authors perceived an overlap to be far less substantial than did editors. They advise that when deciding whether reports are redundant, authors should ask themselves whether a single paper would be more cohesive and more informative than two.

The BMJ under the heading 'Redundant publication' states that they want to make up their own minds on the degree of overlap. For them to do so, authors are required to send them previous publications that overlap by more than 10% with the article being submitted to the BMJ (<http://resources.bmj.com/bmj/authors/article-submission/publication>).

According to the ICMJE guidelines, duplicate publication can be acceptable in specific circumstances, e.g. if there is a need to reach the widest possible audience (e.g. guidelines) or the papers are aimed at different audiences, e.g. if they are in different languages. The guidelines give detailed stipulations to be followed in these circumstances: the editors of both journals must have agreed to the republication and a statement must be published with the paper making it clear that it is a secondary publication. For translations, permission should be sought from the journal that published the original paper and from the copyright holder. It is worth noting though that the National Library of Medicine does not index translations when the original article was published in a journal that is indexed in Medline.

Divided publication is considered ethically unacceptable when a number of articles impart the same data or results. It has also been suggested to be unacceptable if it leads to an inflated evaluation of the writer's output and deprives the reader of understanding the study as a whole.⁶ However, publishing different aspects of the same study in separate papers is not necessarily wrong. The Framingham study is

one example of where separate papers reporting on subsets of data were published without violating publication ethics. Word limits for articles stipulated by journals might also leave authors with little choice other than to split data. If a general medicine journal such as NEJM or BMJ rejects a paper, publication may then only be possible by sending one paper to an obstetrics journal and the other to a paediatric journal. Likewise, different publication routes might be appropriate if one paper concentrates on results and another on methodology. Even with reports of small clinical trials one paper could report outcomes of a study at 2 years and another one 3 years later could report outcomes at 5 years. In all cases, the papers should be cross-referenced or at least authors should declare any data overlap to editors when submitting their articles.⁷

As well as reusing the same data, authors might recycle ideas or reuse text that they have used before in another article. This duplication is often referred to as self-plagiarism. An article by Patrick Scanlon described an incident when a committee was asked to establish whether some authors who published two articles in two separate journals were guilty of self-plagiarism.⁶ Substantial parts of the introductions, literature reviews and methods were identical. The statistical analysis and findings and discussion of each were different. The committee decided that reusing material that established the context for the reader was acceptable. What was important was that the findings of each article made an original contribution to scientific knowledge. Against the authors was that they had failed to cite the other paper. The committee, however, did not consider this to be a case of unethical duplication because different subsets of data had been used, and there had been only two publications rather than a clear intention to deceive by systematic multiple publications.

Plagiarism is to steal somebody else's idea and present it as your own (e.g. see Concise Oxford Dictionary), in other words failing to cite the source of the idea. Additionally, if phrases from a published paper are repeated verbatim they should be placed inside quotation marks, but curiously direct quotes are fairly rare in medical journals. Phrases that are paraphrased or reported indirectly should still be clearly attributed to the original author. A recent case of plagiarism that has caused a cloud to hang over Croatia's research community⁸ was discovered by Iain Chalmers when he was searching for studies to include in a systematic review.⁹ He found that a substantial part of the text and some data in a paper from the Croatian professor, Asim Kurjak, published in *Acta Medica Iugoslavica* in 1974 were identical to those published in an earlier paper by Noble et al. in the *Journal of Obstetrics and Gynaecology of the British Commonwealth*. Asim Kurjak had not cited the other authors' paper. The increasing use of software (e.g. eTBLAST) that identifies similar text and the placing of the results in a database (Déjà vu, see <http://spore.swmed.edu/dejavu>) freely accessible to the public for scrutiny should act as a deterrent against plagiarism and duplicate publication in the future. Ongoing



research by Errami et al.¹⁰ of citations in the Medline database is identifying duplication by authors of their own work and plagiarisms of other authors' work. Furthermore, publishers are beginning to take action on this information.¹¹

There is, however, a question as to how self-plagiarism can exist when you cannot steal from yourself. Scanlon asked this question and concluded that whether an author cites his own previous work should depend on the amount of text that he is recycling and whether readers will profit from seeing the original text.⁶ In my view, the content of the text is critical. For example, if I were to use the first two sentences of this article in another article again (which I have no intention of doing) I would probably not cite this article as the original source of my own idea. As Simon Chapman comments, text reporting original data needs to be differentiated from analysis and commentary: for significant contributors to a debate to only have one opportunity to state their views or be required to constantly rephrase those ideas would be ridiculous.¹² However, maybe not everybody agrees with this view. Glen Griffin in an editorial titled 'Don't plagiarize-even yourself!' states, "Even short quotes from a previously published article should be set in quotation marks and referenced back to the original".¹³ Another consideration is that authors should want to increase citations to their work because citations are beneficial to their careers so failure to do so, when it would be appropriate, could raise suspicion of an intention to deceive.

In an exchange on an editor's internet forum, Stewart Handysides argued against changing the wording in a methods section of a new paper from the original if the original was clear, the words were the author's in the first place, and a citation to the original was given.¹⁴ He makes the point that not to reproduce a good text would disadvantage a new reader, especially one without access to archives. Miguel Roig, who otherwise urges placing identical text in quotes or substantial paraphrasing, also agrees that a more liberal approach is needed for highly technical text.¹⁵ Nevertheless, copying your own work cannot always be done with impunity. Even if you give a citation, you might be in breach of copyright law if you have assigned copyright to a publisher of your article. On the other hand, an author could probably successfully defend such a claim if the amount of text reused is small relative to the entire original article, and republication of this amount of text does not harm the market value of the original article for the copyright holder. Court actions based on copyright of scientific text are not common. Therefore, it is difficult to know how many words or what percentage of the original article would be considered a small amount by a court. It has been suggested that an author challenged with breach of copyright could defend reuse of up to 30% of his or her own text.¹⁶ A court would also take account of the relative importance of the text, e.g. the text containing the core message of the article is relatively more important than the other text in the article.

The US Office for Research Integrity (ORI) does not rank

covert duplicate publication as a research fraud, although some believe that it should.⁸ In contrast, it has declared plagiarism to be a research misconduct. At the same time, it indicates that text can be reused and states that it "generally does not pursue the limited use of identical or nearly-identical phrases which describe a commonly-used methodology or previous research because ORI does not consider such use as substantially misleading to the reader or of great significance." (<http://ori.dhhs.gov/policies/plagiarism.shtml>). Furthermore, the ORI does not consider self-plagiarism to be a form of misconduct and will not take direct action against authors who self-plagiarise, although it states that it will notify the institution(s) from which the duplicate publications/grants originated. Remember plagiarism is where the original source is not attributed. The ORI's stance has disappointed Miguel Roig because he views a failure to disclose an earlier publication as analogous to data falsification if this leads to the same data being used more than once in a meta-analysis (<http://www.wame.org/wame-listserve-discussions/ori-and-self-plagiarism>).

Although the position on republication of identical articles is fairly clear, current guidelines are lacking when it comes to divided publications and reuse of text. Until guidance is given as to what is and what is not ethical, authors can only rely on the various opinions and 'case law' provided by the literature. The cases presented to the Committee on Publication Ethics (COPE), reported together with the committee's advice, can be added to the examples I have given as reference points (<http://www.publicationethics.org.uk/cases/>).

In summary, authors should be advised that:

Republication of an article is only acceptable if the journal that published the original consents and publication is accompanied by a statement that the article is a republication.

Republication of parts of an article is acceptable provided the articles report on different data or use different analysis of the same data and provided

The articles cite each other and the source of the data is clear and

On submission to a journal the editor is informed of the existence of related submissions or publications, even if they are in a different language.

Authors should also be aware of the debates surrounding the grey area of self-plagiarism and would be well advised to make an effort to reformulate the text that they have published before.

Acknowledgements

I would like to thank Miguel Roig and Madeline Frame for their critical review of the manuscript.



References

1. Tramèr M, Reynolds DJM, Moore RA, McQuay HJ. Impact of covert duplicate publication on meta-analysis: a case study. *BMJ*, 1997;315:635-640.
2. Schein M, Paladugu R. Redundant surgical publications: tip of the iceberg? *Surgery* 2001;129:655-661.
3. Bornmann L, Daniel HD. Multiple publication on a single research study: Does it pay? The influence of number of research articles on total citation counts in biomedicine. *Journal of the American Society for Information Science and technology* 2007;58:1100-1107.
4. Von Elm E, Poggia G, Walder B, Tramèr MR. Different patterns of duplicate publication. An analysis of articles used in systematic reviews. *JAMA* 2004;291:974-980.
5. Kassirer JP, Angell M. Redundant publication: a reminder. *NEJM* 1995;333:449-450.
6. Scanlon PM. Song for myself: an anatomy of self-plagiarism. *Plagiary* 2007. http://www.plagiary.org/papers_and_perspectives2007.htm (26 November 2007).
7. Committee on Publication Ethics (COPE) Guidelines on Good Publication Practice. <http://www.publicationethics.org.uk/guidelines/reports/2003/2003pdf15.pdf> (26 November 2007).
8. Godlee F. Editor's choice. Plagiarism and punishment. *BMJ* 2007;335. <http://www.bmj.com/cgi/content/full/335/7627/0> (26 November 2007).
9. Chalmers I. Role of systemic reviews in detecting plagiarism: case of Asim Kurjak. *BMJ* 2006;333. <http://www.bmj.com/cgi/content/full/333/7568/594> (26 November 2007).
10. Errami M, Hicks JM, Fisher W, Trusty D, Wren JD, Lang TC, Garner HR. Déjà vu: a study of duplicate citations in Medline. *Bioinformatics* 2008 24:243-249.
11. Cooney E. Journal Investigating duplicate publication against Boston researcher. http://www.boston.com/news/health/blog/2008/01/a_harvard_rheum.html (27 January 2008).
12. Chapman S. What's so precious about originality? *BMJ* 2007;334:1251.
13. Griffin GC. Don't plagiarize-even yourself! *Postgraduate Medicine*, 1991;89:15-16.
14. Langdon-Neuner E. EASE-forum digest: June-September 2007. *European Science Editing* 2007;33:109.
15. Roig M. Some reflections on plagiarism: the problem of paraphrasing in the sciences. *European Science Editing* 2007;33:38-41.
16. Samuelson P. Self-plagiarism or fair use? *Communications of the ACM* 1994;37:21-25.