

EDITORIAL

Publication bias and the integrity of psychiatry research¹

The cornerstone of evidence-based medicine is the belief that good quality research should form the basis of clinical practice and decision-making (Muir Gray, 1997). Psychiatry has kept abreast of this movement (Geddes *et al.* 1997) and claims have been made that randomized-controlled trials (the highest quality primary evaluative research) can be used to justify 65% of routine clinical decisions (Geddes *et al.* 1996). However, it is largely published research that forms the ‘knowledge base’ of the evidence movement. A fundamental difficulty arises when published research results are a biased sample of all research results – published and unpublished. Publication bias presents one such threat and has been much discussed in wider healthcare (Easterbrook *et al.* 1991; Dickersin & Min, 1993; Dickersin, 1997), but has been little discussed or researched in psychiatry, despite the fact that psychiatry is likely to be at least as prone to publication bias as other specialities.

WHAT IS PUBLICATION BIAS?

Publication bias is that which arises when research fails to be published on the basis of the direction or significance of its results (Dickersin, 1990). It has long been recognized that research with ‘negative’ or ‘uninteresting’ results is less likely to be published (Sterling, 1959), due to the failure of researchers and sponsors to submit negative research and the failure of editors to publish if it is submitted (Dickersin *et al.* 1992). A consequence is the danger that readers of journals are more likely to see studies showing results in a certain direction.

The recent move towards the production of systematic research reviews has generated renewed interest in ‘publication bias’ (Song & Gilbody, 1998), since their results are likely to be biased if publications included are not representative of all studies carried out. This move was, in part, prompted by the recognition that authors of non-systematic reviews often selectively report already published research, producing a related phenomenon – citation bias (Ravnskov, 1992). One important landmark in the recognition of the potential of publication and related bias to produce misleading (spuriously exaggerated) results comes from the meta-analysis of trials of magnesium following myocardial infarction (Egger & Davey-Smith, 1995). In this case, an incorrect meta-analytical result from biased under-powered studies was only revealed by the publication of a mega-trial.

IS PSYCHIATRY SUBJECT TO A PUBLICATION BIAS?

Empirical investigations of publication bias in other spheres of medicine have shown that publication bias does not occur uniformly across all areas of healthcare research. Some forms of research are more prone to publication bias than others. For example, research that is of small sample size, poorer methodological quality or is commercially sponsored is less likely to be published if the results are uninteresting or not favourable to the sponsoring company. Conversely, those studies that are well powered, well designed and the result of multi-centre collaboration and funded from the public purse are likely to be published, irrespective of their results (Dickersin, 1990; Dickersin *et al.* 1992).

Evaluative research in psychiatry demonstrates risk factors for publication bias. For example, a recent survey of 2000 published randomized trials in the field of schizophrenia (Thornley & Adams,

¹ Address for correspondence: Dr Simon M. Gilbody, NHS Centre for Reviews and Dissemination, University of York, York YO10 5DD.

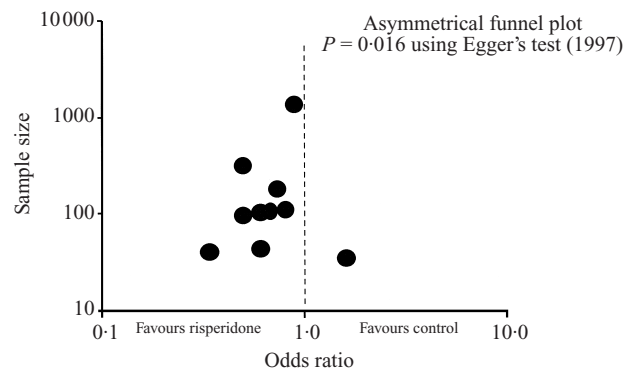


FIG. 1. Asymmetrical funnel plot. Short-term clinical improvement in randomized trials comparing risperidone with conventional neuroleptics. (Data from Kennedy *et al.* 1998, reproduced with kind permission from Update Software: Oxford.)

1998), showed the randomized knowledge base to be dominated by under-powered pharmacological studies, which are likely to be commercially sponsored, although the precise extent of commercial influence is difficult to establish. Similarly, we know that much psychiatric research remains unpublished, since pharmaceutical companies frequently cite 'data on file' (and therefore outside the public domain) in their promotional material. Additionally, a commonly used 'educational' marketing tool is the presentation of off-prints of non-systematic research summaries or individual trials that are unlikely to be representative of the totality of research. Despite this, publication and related bias is rarely discussed in the psychiatric literature, although some systematic reviews investigate the issue in specific areas of enquiry.

One method by which publication bias can be examined is through the use of funnel plots (Light & Pillemer, 1984), which rely on the fact that larger studies are published, irrespective of their results, whereas smaller studies are only selectively published. Funnel plots chart effect size against sample size. In the absence of publication bias, when all studies are plotted as individual data points on this graph, then a symmetrical funnel should be seen. Results from small studies are more prone to random variation and will consequently be dispersed symmetrically (i.e. randomly) about some central overall effect size (i.e. that which would be obtained in a well-powered study). As sample size increases then the degree of random variation about this central axis decreases, producing a symmetrical inverted funnel. If some smaller studies are excluded non-randomly because of their direction of effect, then the funnel will look asymmetrical about its base. Funnel plot asymmetry is highly suggestive of publication and related bias (Egger *et al.* 1997) and indicates that the research in the area under review should, at the least, be treated with suspicion.

Unfortunately, few systematic reviews in any area of healthcare (including psychiatry) routinely test for publication related bias in this way (Song *et al.* 2000), although rigorous systematic reviews, such as those published on the Cochrane Database of Systematic Reviews generally do so. For example, a recent review of studies comparing the effects and costs of risperidone with conventional neuroleptics (Kennedy *et al.* 1998) demonstrates funnel plot asymmetry (Fig. 1), indicating the published literature on this topic presents an over-optimistic view of the effectiveness of this newer drug.

A related phenomenon is the selective or multiple reporting of positive studies and the disaggregation of large scale multi-centre trials by reporting individual centres as if they were different trials ('salami slicing'). This problem is compounded by the failure of authors to acknowledge the existence of duplicate publications and by 'shifting first authorship', making it difficult to detect. Examples of this practice in psychiatry have been highlighted in the case of risperidone by Huston & Moher (1996) and more recently by Duggan *et al.* (1999) in a systematic Cochrane review of olanzapine, where one trial had been published (in one form or another) in 83 separate publications.

In short, if evidence of publication and related bias is sought in some key areas of psychiatric policy and practice, then it is found. Furthermore, in many areas publication might be seen as a sophisticated marketing tool by the sponsors of research, making it seem as if there is more positive evidence than there really is. So if the psychiatric literature is potentially subject to a publication related bias, then what is to be done?

ADDRESSING PUBLICATION BIAS IN PSYCHIATRY

A number of approaches have been proposed to reduce the occurrence and aid the detection and eradication of publication bias (Song *et al.* 2000). Although some of these approaches are potentially useful, several difficulties arise in their application in psychiatry, which are outlined below.

Searching for unpublished studies

Any review of the literature should make efforts to search for unpublished studies (NHS Centre for Reviews and Dissemination, 1996) by for example, contacting the key experts in the field and, in the case of drug trials, the pharmaceutical industry. However, even when these steps are taken, we cannot be sure that all relevant unpublished research has been uncovered, either because of commercial sensitivity or the difficulties in identifying original authors. Similarly, when unpublished research is obtained, there is little way of knowing whether this is the totality of unpublished research or just more selective release into the public domain.

A clear example of the limitations of this approach is seen in the systematic review of risperidone outlined above, which follows best practice in searching for unpublished research but is still subject to a potential bias. The search for unpublished research should therefore be seen as a necessary but imperfect guard against publication bias.

Checking for publication bias

Given the inadequacy of methods to protect against publication bias, advances in the detection of the presence of publication and related biases are welcome. Of the few studies which do check for potential publication bias, the funnel plot represents the most widely used method (Song *et al.* 2000), and examples of the use of this method do exist in psychiatry (e.g. Hotopf *et al.* 1997*a*; Kennedy *et al.* 1998). Formal statistical tests have been developed to augment the use of the funnel plot (by removing subjective bias), of which one by Egger *et al.* (1997) shows the greatest power to detect the presence of publication and related biases (see Song *et al.* 2000, for methodological review).

Unfortunately, difficulties arise in the uniform application of this method in psychiatry since the use of funnel plot analysis relies upon two criteria being satisfied. First, studies must be sufficiently similar in terms of participants and interventions to justify a formal statistical pooling in the form of a meta-analysis. Secondly, the published literature must include a sufficient number of studies with a wide range of sample sizes, providing a mix of smaller studies and one or more larger studies with which to construct a funnel plot. The risperidone example outlined above provides a rare example where there have been both a number of smaller studies and a larger study with over 1000 patients. Unfortunately, psychiatric research is dominated by smaller under-powered studies (Hotopf *et al.* 1997*b*; Thornley & Adams, 1998) and the application of this method will therefore be limited when there are either no large trials or only a small number of trials in total. For example studies comparing clozapine with other atypical anti-psychotics share the same risk factors for publication bias as those in the risperidone example – small sample size and commercial influence (Tuunainen & Gilbody, 1998). However, funnel plots cannot be constructed in this case since there are relatively few individual studies ($N = 5$) and these are uniformly under-powered, with no large trials available to compare effect size. Similarly, in many instances, psychiatric systematic reviews do not (and should not) employ formal meta-analytical pooling, since studies are often too heterogeneous to combine in a meaningful way to produce a summary statistic. For example, a recent systematic

review of interventions for deliberate self-harm (Hawton *et al.* 1998) does not pool widely different psychosocial approaches, but instead chooses to discuss the strengths and weaknesses of each individual study. The authors comment that the research is dominated by small size positive studies and therefore subject to potential publication bias, but are unable to test this assertion.

Despite these difficulties, methods such as funnel plots should always be applied in psychiatry where this is both appropriate and possible, but the extent of publication bias in psychiatry will ultimately remain unknown. The limitations of these methods make it all the more important to consider what steps can be taken to prevent publication bias arising in psychiatric research.

An amnesty for unpublished research

One recent initiative has been a 'trials amnesty', whereby researchers have been encouraged to register unpublished research, so that the results are available to interested parties and are in the public domain (Smith & Roberts, 1997). The amnesty was widely publicized through editorials in major biomedical journals in 1997, although no psychiatric journal was part of this initiative. The details of these studies are included on the electronic Cochrane Library and at the time of writing 150 studies have been identified through this amnesty, but no unpublished psychiatric research has emerged.

It is clear, therefore, that a voluntary amnesty is unlikely to address the problem of publication bias in both psychiatry and wider healthcare. A more formal and mandatory mechanism is therefore required.

A prospective trials register

For over 20 years there have been calls for a mandatory prospective register of clinical trials, whereby trials are registered at inception, irrespective of eventual publication, and the granting of ethical approval and acceptance of their results by bodies such as the FDA are contingent upon this registration (Simes, 1986; Easterbrook, 1987, 1992; Dickersin, 1988; Stern & Simes, 1997). Publication bias would become irrelevant, since it would not matter if such research were unpublished if its existence was registered in the public domain. Some areas of subspecialty such as oncology have developed their own voluntary registers for research, and there have been calls for such a register in psychiatry (e.g. Adams & Gelder, 1994) which are as yet unheeded.

Increasingly, the continued existence of publication bias is now difficult to defend and there have been a number of recent initiatives about which psychiatry should take note and be prepared to participate in. In the commercial sector two major pharmaceutical companies (Schering and Glaxo-Wellcome) have pledged to make available all clinical research on a product once it has ceased to be commercially sensitive and has gained product approval (Sykes, 1998). Moreover, commitments have been made to ensure that all individual studies are registered at inception and are given a unique code number to aid the detection of duplicate publication. Details of these studies are soon to be included on an Internet based register of completed or ongoing unpublished trials (www.controlled-trials.com), which will be searchable in the same way as other electronic databases such as Medline. At the time of writing none of the major manufacturers of psychotropics has given a public undertaking to follow this lead. In the current climate, this position is likely to become increasingly untenable (Rennie, 1999).

Mega-trials are needed

Psychiatry is bedevilled by the problem of small size underpowered studies, for which meta-analyses and systematic reviews are a necessary but imperfect solution. What are really needed are mega-trials registered in the public domain with sufficient power to establish the effectiveness or otherwise of common interventions and policy initiatives. The paradox remains that until there are mega-trials, then the existence of publication bias among smaller studies is difficult to detect and that the detection of publication bias is most important in those areas where small size trials predominate and there are no mega-trials. It should be remembered that the existence of the few large scale trials

in pharmacotherapy has not prevented their duplicate publication and disaggregation as a sophisticated form of commercial research dissemination (Huston & Moher, 1996; Duggan *et al.* 1999; Rennie 1999).

CONCLUDING COMMENTS

Despite psychological researchers being the first to recognize the importance of publication bias (Sterling, 1959), this issue has been all but ignored in the sphere of mental health. Publication bias and the factors that cause it need to be dealt with if psychiatry is going to become more 'evidence based'. On a practical level, publication bias threatens the validity of our research knowledge base. More importantly, the continued existence of publication bias represents an abuse of the trust and time that patients give through their participation in clinical research, and is essentially a form of scientific misconduct (Chalmers, 1990).

The recognition of the potential consequences of publication bias has led to important advances in its minimization and detection. However, these methods are either rarely employed when they should be in psychiatry, or are likely to be of limited use. This is unfortunate when, compared to other specialities, psychiatry is likely to be especially prone to publication bias. Ultimately, the only answer to publication bias in psychiatry is a mandatory clinical trials register. The arguments for this are both practical and moral.

SIMON M. GILBODY AND FUJIAN SONG

We are grateful to Professor Trevor Sheldon of York Health Policy Group for his comments. S.G. is supported by the UK Medical Research Council Fellowship scheme and the work discussed in this editorial is funded by the UK Health Technology Assessment Programme (project 95/12/01).

REFERENCES

- Adams, C. & Gelder, M. (1994). The case for establishing a register of randomised controlled trials of mental health care. *British Journal of Psychiatry* **164**, 433–436.
- Chalmers, I. (1990). Under-reporting research is scientific misconduct. *Journal of the American Medical Association* **263**, 1405–1408.
- Dickersin, K. (1988). Report from the panel on the Case for Registers of Clinical Trials at the Eighth Annual Meeting of the Society for Clinical Trials. *Controlled Clinical Trials* **9**, 76–81.
- Dickersin, K. (1990). The existence of publication bias and risk factors for its occurrence. *Journal of the American Medical Association* **263**, 1385–1389.
- Dickersin, K. (1997). How important is publication bias? A synthesis of available data. *Aids Education and Prevention* **9**, 15–21.
- Dickersin, K. & Min, Y. I. (1993). Publication bias: the problem that won't go away. *Annals of the New York Academy of Science* **703**, 135–46.
- Dickersin, K., Min, Y. I. & Meinert, C. L. (1992). Factors influencing publication of research results. Follow-up of applications submitted to two institutional review boards. *Journal of the American Medical Association* **267**, 374–378.
- Duggan, L., Fenton, M., Dardennes, R. M., El-Dosoky, A. & Indran, S. (1999). Olanzapine for schizophrenia (Cochrane Review). In *The Cochrane Library, Issue 3*. Update Software: Oxford.
- Easterbrook, P. (1987). Reducing publication bias. *British Medical Journal* **295**, 1347.
- Easterbrook, P. J. (1992). Directory of registries of clinical trials. *Statistics in Medicine* **11**, 345–423.
- Easterbrook, P. J., Berlin, J. A., Gopalan, R. & Matthews, D. R. (1991). Publication bias in clinical research. *Lancet* **337**, 867–872.
- Egger, M. & Davey-Smith, G. (1995). Misleading meta-analysis: lessons from 'an effective, safe, simple' intervention that wasn't. *British Medical Journal* **310**, 752–754.
- Egger, M., Davey-Smith, G., Schneider, M. & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *British Medical Journal* **315**, 629–634.
- Geddes, J. R., Game, D., Jenkins, N. E., Peterson, L. A., Pottinger, G. R. & Sackett, D. L. (1996). What proportion of psychiatric interventions are based on randomised evidence? *Quality in Health Care* **5**, 215–217.
- Geddes, J. R., Reynolds, S., Streiner, D. & Szatmari, P. (1997). Evidence based practice in mental health. *British Medical Journal* **315**, 1483–1484.
- Hawton, K., Arensman, E., Townsend, E., Bremner, S., Feldman, E., Goldney, R., Gunnell, D., Hazell, P., van-Heeringen, K., House, A., Owens, D., Sakinofsky, I. & Traskman-Bendz, L. (1998). Deliberate self harm: systematic review of efficacy of psychosocial and pharmacological treatments in preventing repetition. *British Medical Journal* **317**, 441–447.
- Hotopf, M., Hardy, R. & Lewis, G. (1997a). Discontinuation rates of SSRIs and tricyclic antidepressants: a meta-analysis and investigation of heterogeneity. *British Journal of Psychiatry* **170**, 120–127.
- Hotopf, M., Lewis, G. & Normand, C. (1997b). Putting trials on trial – the costs and consequences of small trials in depression: a systematic review of methodology. *Journal of Epidemiology and Community Health* **51**, 354–358.
- Huston, P. & Moher, D. (1996). Redundancy, disaggregation, and the integrity of medical research. *Lancet* **347**, 1024–1026.
- Kennedy, E., Song, F., Hunter, R. & Gilbody, S. (1998). Risperidone versus 'conventional' antipsychotic medication for schizophrenia (Cochrane Review). In *The Cochrane Library, Issue 3*. Update Software: Oxford. Updated quarterly.
- Light, R. J. & Pillemer, D. B. (1984). *Summing Up: The Science of Reviewing Research*. Harvard University Press: Cambridge, Mass.
- Muir Gray, J. A. (1997). *Evidence-based Healthcare: How to Make Health Policy and Management Decisions*. Churchill Livingstone: London.
- NHS Centre for Reviews and Dissemination (1996). *Undertaking Systematic Reviews of Research on Effectiveness: CRD Guidelines for Those Carrying Out or Commissioning Reviews CRD Report 4*. University of York: York.
- Rennie, D. (1999). Fair conduct and fair reporting of clinical trials. *Journal of the American Medical Association* **282**, 1766–1769.

- Ravnskov, U. (1992). Cholesterol lowering trials in coronary heart disease: frequency of citation and outcome. *British Medical Journal* **305**, 15–19.
- Simes, R. J. (1986). Publication bias: the case for an international registry of clinical trials. *Journal of Clinical Oncology* **4**, 1529–1541.
- Smith, R. & Roberts, I. (1997). An amnesty for unpublished trials: send us details on any unreported trials. *British Medical Journal* **315**, 622.
- Song, F. & Gilbody, S. (1998). Increase in studies of publication bias coincided with increasing use of meta-analysis. *British Medical Journal* **316**, 471.
- Song, F. S., Eastwood, A. J., Gilbody, S. M., Sutton, A. & Duley, L. (2000). Publication and related biases (Report commissioned by the NHS Health Technology Assessment Programme). *Technology Assessment* (in the press).
- Sterling, T. D. (1959). Publication decisions and their possible effects on inferences drawn tests of significance – or vice versa. *American Statistical Association Journal* **54**, 30–34.
- Stern, J. M. & Simes, R. J. (1997). Publication bias: evidence of delayed publication in a cohort study of clinical research projects. *British Medical Journal* **315**, 640–645.
- Sykes, R. (1998). Being a modern pharmaceutical company: involves making information available on clinical trial programmes. *British Medical Journal* **317**, 1172.
- Thornley, B. & Adams, C. E. (1998). Content and quality of 2000 controlled trials in schizophrenia over 50 years. *British Medical Journal* **317**, 1181–1184.
- Tuunainen, A. & Gilbody, S. (1998). Newer atypical antipsychotic medication versus clozapine for schizophrenia (Cochrane Review). *The Cochrane Library, Issue 3*. Update Software: Oxford.