

Editorials

Bias in the evaluation of antidepressants

MATTHEW HOTOPF and CORRADO BARBUI

Tricyclic antidepressants (TCAs) are no longer widely prescribed as first choice treatments for depression. In their place have come the selective serotonin reuptake inhibitors (SSRIs). Sales of SSRIs and other new antidepressants have grown massively partly at the expense of tricyclics and partly because of burgeoning antidepressant prescribing in general (Lawrenson *et al.*, 2000; Middleton *et al.*, 2001; Rosholm *et al.*, 2001; Isacsson *et al.*, 1999; Ciuna *et al.*, 2004). Are the SSRIs really so much better?

SSRIs have been marketed as being as effective as tricyclics, but better tolerated and safer in overdose. There are probably numerous other perceived advantages – they may be less sedating so cause fewer psychomotor symptoms, and as “cleaner” drugs are less likely to have pharmacodynamic interactions with other drugs. However if one examines evidence purely from the accepted gold standard to compare two treatments, any advantages are insignificant. It is true that for every 30 or so patients started on an antidepressant, under trial conditions, one more will remain on the antidepressant at six weeks if an SSRI was prescribed, indicating a small advantage in terms of tolerability (Hotopf *et al.*, 1997a; Barbui *et al.*, 2000). In terms of efficacy, TCAs seem to have the edge in comparison with SSRIs. Perry, who reviewed the Hamilton scores in trials comparing TCAs with SSRIs, found TCAs more effective than SSRIs in patients suffering from severe depression with melancholic features

(Perry, 1996). Similar conclusions were reached by Joffe, who found a non significant advantage of TCAs against SSRIs in terms of effect size (Joffe *et al.*, 1996), while Anderson, who systematically reviewed more than 100 randomised clinical trials (RCTs), showed that TCAs appeared more effective in inpatients but not in outpatients with depression (Anderson, 2000). We found that the reference tricyclic, amitriptyline, was slightly more effective than SSRIs as a class (Barbui & Hotopf, 2001a), and that this was explained by a more favourable profile in inpatients but not in outpatients (Barbui *et al.*, 2004a). However, other meta-analyses which compare the two classes find no difference (Anderson, 2001).

Is the evidence from the RCTs reliable? Most trials comparing the two classes are remarkably homogenous in design, generally comparing moderately depressed outpatients at six weeks, and mainly using the Hamilton Depression Rating Scale (HRSD) as an endpoint. We have commented elsewhere on the low quality of these studies, with small sample sizes, short follow up, and highly selected participants (Hotopf *et al.*, 1997b; Barbui & Hotopf, 2001b). The homogeneity of the trial designs means that within the limits of the primary research the use of meta-analysis is probably reasonable. A more serious concern is that of publication bias.

Funnel plots should be able to help us here. They work on the assumption that researchers are less likely to leave unpublished the results of large trials, than they are with small trials. A graph plotting the number of participants in the study (on the Y axis) against the effect size (on the X axis) will show an asymmetrical pattern if there is significant publication bias, with small “negative” studies (i.e. ones which favour the older comparison agent) being conspicuous by their absence. For the meta-analyses of tricyclics and SSRIs the funnel plots have generally been symmetrical, suggesting publication bias is absent. But

Address for correspondence: Professor M. Hotopf, Department of Psychological Medicine, Institute of Psychiatry, Weston Education Centre, Cutcombe Rd, London SE5 9RJ (United Kingdom).

E-mail: m.hotopf@iop.kcl.ac.uk

Declaration of Interest: none.

should that reassure us? The funnel plot may give a very clear answer when there have been “mega-trials” comparing two treatments (ie studies with thousands of participants), as mega-trials are unlikely to go unpublished. For the antidepressant trials there is not much spread in sample size between the largest and smallest trial.

Recent disclosures that GlaxoSmithKline deliberately “buried” two unfavourable randomised controlled trials giving key data on children and adolescents with major depression suggests to us that publication bias may remain a very serious limitation to the entire literature comparing SSRIs and tricyclics (Parker *et al.*, 2003). If industry is prepared to conceal crucial safety and efficacy information on the treatment of children with depression with their compounds for purely commercial reasons, it is plausible that the same companies will have been as unscrupulous with data in adults. Furthermore, the publication bias on the paroxetine trials in children was not limited to small poorly conducted trials and as the recent systematic review for the UK National Institute of Clinical Excellence shows, the two unpublished trials were each bigger than the smaller (but commercially favourable) published trial (Whittington *et al.*, 2004). The funnel plot (and other formal statistical tests which work on the same principle) would not be able to detect publication bias under these circumstances.

This also raises a concern about the purpose of many of the trials we have reviewed. The sample sizes of most such trials is so small that highly clinically significant differences between agents simply could not have been detected. Thus even if the newer antidepressants performed no better than placebo, many of these trials would have indicated no statistically significant difference when compared with the reference compound. Unfortunately this is often not reflected in the conclusion of the trials, which tend to be that the newer agent is as effective as the old one. These trials are effectively a form of marketing, and are often used to highlight a particular advantage of the new drug over the old, which may have been a type I error.

There is evidence of other forms of bias in randomised trials of antidepressants, which we have reviewed. For example, if one looks at all trials comparing fluoxetine and other antidepressants, and then categorises them according to whether fluoxetine was the new agent or the comparator, fluoxetine is slightly more effective when it is the new agent. In other words there is evidence of the so-called “wish bias” in which possibly due to observer bias, or publication bias, the drug performs better when it is new than when it is old (Barbui *et al.*, 2004b). In addition fluoxetine tended to be prescribed in higher doses when it was the new agent, although this was not respon-

sible for the difference in efficacy (Barbui *et al.*, 2002).

The main disadvantage of the SSRIs when they were first introduced was the very significant difference in price between them and the tricyclics, an issue of less importance now as they come off patent, but one which remains important as another generation of antidepressants becomes available. We have reviewed the economic analyses comparing any groups of antidepressants, and found only three published RCTs to include economic analyses (Barbui *et al.*, 2003). These show few differences between antidepressants in terms of total health care costs.

An alternative study design – the database analysis – has been much more widely published, almost always with direct support from individual pharmaceutical companies. Database analyses use large administrative healthcare databases to compare costs between individuals newly prescribed different antidepressants. Database analyses have many inherent methodological weaknesses (Hotopf *et al.*, 1996). Firstly, nothing is known about actual outcome, other than future health care use. Secondly, the design is not randomised, so it is possible that confounders – known or unknown – are responsible for any differences observed. For example it may be that individuals prescribed tricyclics have more severe depression, or previous episodes for which tricyclics had been prescribed. Such patients would be expected to have a worse outcome, making tricyclics appear more expensive than they would otherwise be. Thirdly, as a retrospective nature of these analyses mean that it is often very hard to follow the analytic strategy used. This means that investigators perform multiple analyses using different endpoints, and subtly different methods, to calculate costs. Although our review found few consistent patterns in these studies, what was striking was the finding that the supporting company never seemed to publish evidence against their own antidepressant (Barbui *et al.*, 2003). In four database analyses comparing sertraline with fluoxetine, three (all funded by Eli Lilly, the manufacturer of fluoxetine) found fluoxetine was associated with lower overall healthcare costs. The fourth study, funded by Pfizer (the manufacturer of sertraline) found sertraline was cheaper.

In conclusion, the SSRIs have become treatments of choice for depression, not because of any obvious advantages over their predecessors which could be demonstrated using randomised controlled trials. We suggest that the only significant advantage was their greater safety in overdose. However with the current controversy over the potential for SSRIs to induce suicidal behaviour continuing, even this advantage may be of marginal benefit.

REFERENCES

- Anderson I.M. (2000). Selective serotonin reuptake inhibitors versus tricyclic antidepressants: a meta-analysis of efficacy and tolerability. *Journal of Affective Disorders* 58, 19-36.
- Anderson I.M. (2001). Meta-analytical studies on new antidepressants. *British Medical Bulletin* 57, 161-178.
- Barbui C. & Hotopf M. (2001a). Amitriptyline versus the rest: still the leading antidepressant after 40 years of randomised controlled trials. *British Journal of Psychiatry* 178, 129-144.
- Barbui C. & Hotopf M. (2001b). 40 years of antidepressant drug trials. *Acta Psychiatrica Scandinavica* 104, 92-95.
- Barbui C., Hotopf M., Freemantle N., Boyton J., Churchill R., Eccles M.P., Geddes J.R., Hardy R., Lewis G. & Mason J.M. (2000). Treatment discontinuation with selective serotonin reuptake inhibitors (SSRIs) versus tricyclic antidepressants (TCAs). *Cochrane Database of Systematic Reviews*. Issue 4.
- Barbui C., Hotopf M. & Garattini S. (2002). Fluoxetine dose and outcome in antidepressant drug trials. *European Journal of Clinical Pharmacology* 58, 379-386.
- Barbui C., Percudani M. & Hotopf M. (2003). Economic evaluation of antidepressive agents: a systematic critique of experimental and observational studies. *Journal of Clinical Psychopharmacology* 23, 145-154.
- Barbui C., Guaiana G. & Hotopf M. (2004a). Amitriptyline for inpatients and SSRIs for outpatients with depression? Systematic review and meta-regression analysis. *Pharmacopsychiatry* 37, 93-97.
- Barbui C., Cipriani A., Brambilla P. & Hotopf M. (2004b). Wish bias in antidepressant drug trials? *Journal of Clinical Psychopharmacology* 24, 126-130.
- Ciuna A., Andretta M., Corbari L., Levi D., Mirandola M., Sorio A. & Barbui C. (2004). Are we going to increase the use of antidepressants up to that of benzodiazepines? *European Journal of Clinical Pharmacology* 60(9), 629-634.
- Hotopf M., Lewis G. & Normand C. (1996). Are SSRIs a cost-effective alternative to tricyclics? *British Journal of Psychiatry* 168, 404-409.
- 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 clinical trials in depression: a systematic review of methodology. *Journal of Epidemiology and Community Health* 51 354-358.
- Isacson G., Boethius G. & Henriksson S. (1999). Selective serotonin reuptake inhibitors have broadened the utilisation of antidepressant treatment in accordance with recommendations. Findings from a Swedish prescription database. *Journal of Affective Disorders* 53, 15-22.
- Joffe R., Sokolov S. & Streiner D. (1996). Antidepressant treatment of depression. A meta-analysis. *Canadian Journal of Psychiatry* 41, 613-616.
- Lawrenson R.A., Tyrer F. & Newson R.B. (2000). The treatment of depression in UK general practice: selective serotonin reuptake inhibitors and tricyclic antidepressants compared. *Journal of Affective Disorders* 59, 149-157.
- Middleton N., Gunnell D. & Whitley E. (2001). Secular trends in antidepressant prescribing in the UK, 1975-1998. *Journal of Public Health Medicine* 23, 262-267.
- Parker G., Anderson I.M. & Haddad P. (2003). Clinical trials of antidepressant medications are producing meaningless results. *British Journal of Psychiatry* 183, 102-104.
- Perry P.J. (1996). Pharmacotherapy for major depression with melancholic features: relative efficacy of tricyclic versus selective serotonin reuptake inhibitor antidepressants. *Journal of Affective Disorders* 39, 1-6.
- Rosholm J.U., Andersen M. & Gram L.F. (2001). Are there differences in the use of selective serotonin reuptake inhibitors and tricyclic antidepressants? A prescription database study. *European Journal of Clinical Pharmacology* 56, 923-929.
- Whittington C.J., Kendall T., Fonagy P., Cottrell D., Cotgrove A. & Boddington E. (2004). Selective serotonin reuptake inhibitors in childhood depression: systematic review of published versus unpublished data. *Lancet* 363, 1341-1345.