quarter of 2001–02.[2] The estimated annual use fell from about 66 000 to 51 000 treatments; if this decline in ECT use continued at the same rate, then the straight line extrapolated from the last data point would reach zero by the year 2011–12. There has never been another national survey. A partial survey of English ECT clinics in the first quarter of 2006 suggested a further fall, to only about 27 000, which was in line with the extrapolation.[3] The re-appraisal prompted us to review the rate of ECT usage in our clinic since 2006.

The rates of ECT usage in 2006 and 2011 were almost identical, that is, 0.82 and 0.83 individual treated patients per 10 000 population in the City of Edinburgh. Likewise, the rates in the intervening years were also almost identical. We therefore conclude, at least for Edinburgh, that the rate of ECT use has been stable for the past 6 years.

The electronic data collection system in our ECT clinic was updated at the end of 2004, and included a record of the primary psychiatric diagnosis of referred patients. The number of referred patients diagnosed with a severe depressive episode (both with and without psychotic features) varied little in these 6 years, from 23 to 28 patients. This gave a crude referral rate of 25 patients with severe depression per year per total population of 500 000.

If we are treating just as many patients with severe depression as 5 or 6 years ago, then this must continue to be resourced. It is not just ECT practitioners that have heard the suggestion about the demise of ECT. Senior managers locally have expressed surprise to hear that there is still a need for the ECT clinic at the Royal Edinburgh Hospital. This concerned us because when the availability of ECT was reduced in Glasgow, ECT use fell.[4] The search for the Holy Grail is laudable, but patients with severe depression still need access to ECT.

1   Allan C, Kalu U-G, Sexton CE, Ebmeier KE. Transcranial stimulation in depression. *Br J Psychiatry* 2012; **200**: 10–11.

2   Department of Health. *Electro Convulsive Therapy: Survey Covering the Period from January 2002 to March 2002, England. Statistical Bulletin 2003/08.* Department of Health, 2003.

3   Bickerton D, Worrall A, Chaplain R. Trends in the administration of electroconvulsive therapy in England. *Psychiatr Bull* 2009; **33**: 61–3.

4   McAllister M, Crabb J, Brodie D, Krishandas R. Effect of closing facilities on electroconvulsive therapy use in Glasgow. *J ECT* 2011; **27**: 131–3.

Allan Scott, Royal Edinburgh Hospital, Morningside Terrace, Edinburgh EH10 5HF, UK. Email: Fiona.J.Morrison@nhslothian.scot.nhs.uk; **Simone Burns**, Royal Edinburgh Hospital, Edinburgh, UK

Methodological discrepancies in the update of a meta-analysis

Leichsenring & Rabung[1] reported that long-term psychodynamic psychotherapy (LTPP) is superior to less intensive forms of psychotherapy in complex mental disorders. Based on 10 trials, they found an overall effect size (ES) of 0.55. We found several methodological discrepancies in their study.

First, it seems surprising that the Q-test indicated no significant unexplained variance, as the between-group effect size of one of the primary studies[2] (ES = 1.76) is quite outstanding in Fig. 2. To shed light on this issue, we recalculated the overall effect size using a random effects meta-analysis based on the values from Fig. 2. Our meta-analysis replicated Leichsenring & Rabung[1] in the main. In contrast to Leichsenring & Rabung however, we found a significant unexplained variance ($Q = 25.33$, d.f. = 9, $P = 0.003$) and a larger overall confidence interval of 0.29–0.82

(in contrast to 0.41–0.67 as reported by Leichsenring & Rabung). Additionally, computing an outlier analysis, a significant outlying study effect size was found ($P < 0.001$). Including the moderator considering the impact of this study yields an effect size of 0.44 (95% CI 0.27–0.61, $P < 0.001$). The moderator effect, interpreted as the difference between the effect of the outlying study and the grand mean, was 1.32 (95% CI 0.57–2.07, $P < 0.001$). After removing the outlying study, there was no significant unexplained variance ($Q = 11.56$, d.f. = 8, $P = 0.172$).

Second, we calculated the fail-safe N according to Rosenthal;[3] 16 non-published studies with an effect size of 0 had to be included in the analysis to change the results of the meta-analysis (ES = 0.44) from significant to non-significant (ES < 0.16). As 16 is below 55 ($5K + 10$), the effect cannot be regarded as robust.

Last, to gain better insight into the interpretation of the overall effect size as small, medium or large, we calculated a Bayesian meta-analysis following Higgins *et al*'s methodology.[4] The Bayesian analysis essentially replicated the findings of our random effects meta-analysis. In addition, we found the probability of the overall effect size to be small (ES < 0.5) at 72.5%. Thus, in contrast to Leichsenring & Rabung,[1] we found that the overall effect size was small rather than medium or large.

Therefore, we would greatly appreciate caution against a conclusion that the overall effectiveness of LTPP for treating complex mental disorders should now be considered as definitely proven.

1   Leichsenring F, Rabung S. Long-term psychodynamic psychotherapy in complex mental disorders: update of a meta-analysis. *Br J Psychiatry* 2011; **199**: 15–22.

2   Bateman A, Fonagy P. The effectiveness of partial hospitalization in the treatment of borderline personality disorder, a randomized controlled trial. *Am J Psychiatry* 1999; **156**: 1563–9.

3   Rosenthal R. The file drawer problem and tolerance for null results. *Psychol Bull* 1979; **86**: 638–41.

4   Higgins JPT, Thompson SG, Spiegelhalter DJ. A re-evaluation of random-effects meta-analysis. *J R Statist Soc A* 2009; **172**: 137–59.

Soeren Kliem, PhD student, Department of Psychology, Technical University of Braunschweig, Germany. Email: soerenkliem@gmx.net; **Johannes Beller**, **Christoph Kroeger**, Department of Psychology, Technical University of Braunschweig, Germany

Author's reply: When trying to replicate some results of our meta-analysis,[1] Kliem and colleagues reported some methodological discrepancies.[2] These discrepancies, however, are due to modifications in their statistical approach as compared with the one we originally reported.

First, in contrast to our results,[1] Kliem *et al* reported significant heterogeneity between studies for overall outcome as indicated by the Q statistic. As stated in our meta-analysis, we had aggregated the effect size estimates across studies, adopting a random effects model, which is more appropriate than a fixed effects model if the aim is to make inferences beyond the observed sample of studies.[1,2] Applying a random effects model, the aggregated effect size for overall outcome was 0.54, and heterogeneity was not significant ($Q = 11.72$, $P = 0.23$, $I^2 = 23$). Thus, there was no need for an additional outlier analyses or for the exclusion of any study. As Rosenthal's fail-safe N was 66, which is above 60 ($5K + 10$), the effect can be regarded as robust. Kliem *et al*, however, apparently applied the fixed effects model to test for heterogeneity. The use of a fixed effects model, however, addresses another research question and consequently yields different results.

Furthermore, Kliem *et al* reported a larger confidence interval for the overall effect size. The confidence interval that they calculated corresponds well to the respective interval that we reported in Table 1, which was 0.26–0.83.[1] A narrower confidence interval, however, was erroneously reported by us in the forest plot owing to a transcription error (Fig. 2, 95% CI 0.41–0.67)[1] – see correction.

Second, after excluding the study by Bateman & Fonagy,[3] which they regarded as an outlier, Kliem *et al* reported a fail-safe *N* of 16 unpublished studies with an effect size of 0, which would need to be added in the meta-analysis to change the result from significant to non-significant. However, we could not replicate these findings. After excluding the study by Bateman & Fonagy, we found a fail-safe *N* of 69, which is above 55 ($5K + 10$), again indicating that the effect is robust. Apparently, Kliem *et al* erroneously did not calculate a fail-safe *N* according to Rosenthal but according to Orwin's method.[4] Consequently, they did not assess how many studies with $ES = 0$ had to be included to change the result from significant to non-significant but to 'not significantly different from 0.16' – an irrelevant result.

Third, results of Bayesian meta-analyses depend largely on the specification of prior assumptions on the treatment effect and between-trial variance. Since Kliem *et al* did not provide any information about the assumptions of their analyses, it is impossible to interpret the presented result reasonably.

In summary, we could not confirm the discrepancies reported by Kliem *et al*. We did not find substantial heterogeneity or any cogent indication of publication bias. The effect in favour of long-term psychodynamic psychotherapy was confirmed as robust. Instead, we could show that most of those 'discrepancies' seem to be based on differing methodological approaches.

1 Leichsenring F, Rabung S. Long-term psychodynamic psychotherapy in complex mental disorders: update of a meta-analysis. *Br J Psychiatry* 2011; **199**: 15–22.

2 Hedges LV, Vevea JL. Fixed- and random-effects models in meta-analysis. *Psychol Meth* 1998; **3**: 486–504.

3 Bateman A, Fonagy P. The effectiveness of partial hospitalization in the treatment of borderline personality disorder, a randomized controlled trial. *Am J Psychiatry* 1999; **156**: 1563–9.

4 Orwin RG. A fail-safe N for effect size in meta-analysis. *J Educ Behav Stat* 1983; **8**: 157–9.

Falk Leichsenring, Department of Psychosomatics and Psychotherapy, University of Giessen, Ludwigstrasse 76, 35392 Giessen, Germany. Email: Falk.Leichsenring@psycho.med.uni-giessen.de

doi: 10.1192/bjp.200.5.429a

Identifying responders in randomised controlled trials for depression

The usefulness of antidepressants in patients with mild or moderate depression has been questioned and therefore Thase *et al*[1] re-analysed randomised controlled trial (RCT) data for escitalopram. They calculated that a subgroup of patients (14%) with mild or moderate depression did respond to treatment and not to placebo, but that the subgroup of patients with severe depression who responded to treatment was larger (23%).

The findings from Thase *et al*[1] illustrate what is known from the philosophy of science about RCTs: if T causes O in a population, this only implies that T causes O in at least some members of that population.[2] Reporting average results can be misleading if the population is not causally homogeneous.

Ideally, one should try to identify the specific subgroup of responders. The authors could consider analysing their data further by looking at demographic factors such as age, because younger people are less likely to respond,[3] and the presence of particular symptoms such as anxiety, which might adversely influence outcome as well,[4] if these data are available. However, it is unlikely that one can identify the subgroups of responders with 100% accuracy with these additional data.

More generally, Thase *et al*'s results show that by reporting average results, important findings from RCTs might be missed. Reporting binary outcomes such as number or percentage of patients improved or in remission should be encouraged in psychiatric research, even if the primary outcome variable is considered to be a continuous interval variable. This applies to all RCTs with non-homogeneous populations. Unfortunately, CONSORT requirements do not make this compulsory at the moment.[5]

1 Thase ME, Larsen KG, Kennedy SH. Assessing the 'true' effect of active antidepressant therapy *v.* placebo in major depressive disorder: use of a mixture model. *Br J Psychiatry* 2011; **199**: 501–7.

2 Cartwright N. Are RCTs the gold standard? *BioSocieties* 2007; **2**: 11–20.

3 Tsapakis EM, Soldani F, Tondo L, Baldessarini RJ. Efficacy of antidepressants in juvenile depression: meta-analysis. *Br J Psychiatry* 2008; **193**: 10–7.

4 Coryell W, Fiedorowicz JG, Solomon D, Leon AC, Rice JP, Keller MB. Effects of anxiety on the long-term course of depressive disorders. *Br J Psychiatry* 2012; **200**: 210–5.

5 Schulz KF, Altman DG, Moher D. CONSORT 2010 statement: updated guidelines for reporting parallel group randomized trials. *Ann Int Med* 2010; **152**: 1–8.

Dieneke Hubbeling, Consultant Psychiatrist, Wandsworth Crisis and Home Treatment Team, South West London and St George's Mental Health NHS Trust, 16 Glenburnie Road, London SW17 7DJ, UK. Email: dieneke@doctors.org.uk

doi: 10.1192/bjp.200.5.430

Corrections

Long-term psychodynamic psychotherapy in complex mental disorders: update of a meta-analysis. *BJP*, **199**, 15–22. Figure 2 (p. 19): total effect size CI should read 0.26 to 0.83.

Predicting the development of schizophrenia: authors' reply (letter). *BJP*, **200**, 255. Jefter Chuma is Clinical Studies Officer with NIHR Mental Health Research Network (North West Hub), and Prem Mahadun is Consultant Psychiatrist with the Trafford Crisis Resolution and Home Treatment Team.

doi: 10.1192/bjp.200.5.430a