Letter to the Editor

A comment on Lynch et al. (2009)

Meta-analysis (MA) is an essential tool for summarizing evidence for a specific intervention, but is prone to bias and not objective per se. Because many MAs have failed to report procedures in a transparent way that enables readers to assess strengths and weaknesses, a group of researchers developed the QUORUM guidelines (Moher et al. 1999; update: Moher et al. 2009). These list 19 major criteria which are deemed essential for transparent reporting of the method and results in a systematic review (overall there are 27 guidelines, referring to title, abstract, introduction, methods, results, discussion and funding). Lynch et al. (2009) only comply with five of these. For example, they do not present the full electronic search strategy including search terms or describe the process of study selection (e.g. screening, determining eligibility) or the process of data extraction (e.g. were different raters involved in the data extraction and how did they agree?), they do not list and define all variables for which data was sought and, although they emphasize the risk of over-interpreting results from methodologically weak studies, they do not describe methods for assessing risk of bias in the included studies, such as quality of randomization and blinding or drop-out rates. Moreover, they do not transparently describe the synthesis of results. The results section contains no flow diagram of the study selection or numbers of studies screened and there is no description of the included studies with regard to relevant study characteristics.

This lack of reporting makes it extremely difficult to understand their selection of studies. For example, one study that used an active control design (Levine et al. 1998) and was included in other meta-analyses (Lincoln et al. 2008; Wykes et al. 2008) were not even listed in the list of excluded studies (see supplementary online Appendix in Lynch et al. 2009). Whereas a study by Hogarty et al. (1997) that used an intervention that was not considered as CBT by the author of that study or the authors of other meta-analyses, was included. Some studies were excluded because of using additional elements in the intervention, such as motivational interviewing or family inclusion whereas others were included although they also used motivational interviewing (Haddock et al. 2009) or involved family members (Drury et al. 1996). Other exclusion criteria are listed more explicitly but lack a strong rationale. For example, why was the label ‘pilot study’ an exclusion criteria, given all other criteria were fulfilled? This resulted in the exclusion of two relevant studies. Further, why was relapse restricted to defined symptom changes whereas studies focusing on rehospitalization or follow-up symptom scores – for which beneficial effects of CBT have been demonstrated (Lincoln et al. 2008) – were excluded? Despite other disadvantages, rehospitalization rates or days would have been the least prone to observer bias, which is what the authors were aiming at. Alone, the exclusion of studies that focused on rehospitalization reduced the pool of relevant studies by another five. Finally, a number of not previously defined exclusion criteria were added in the results section or appeared in the list of excluded studies, such as co-morbid substance abuse, the use of cognitive remediation as a control intervention, exceeding a certain percentage of affective psychoses or the use of 5-year follow-up periods. These criteria reduced the number of included studies by a further five. As the authors do not, in fact, restrict their analyses to blind or
active-controlled studies, it is difficult to ascertain what the 13 studies that survived this selection process have in common.

**What do we learn from this meta-analysis?**

Several recent MAs (Zimmermann et al. 2005; Lincoln et al. 2008; Wykes et al. 2008) that identified small to medium effects for CBT have also investigated the effect of study quality on effect size. In particular, the MA by Wykes et al. (2008), which included 34 RCTs, tested the moderating effect of different aspects of study quality alone and in combination. They also found overall effect sizes to be smaller, albeit still significant, in studies using blind symptom ratings. Rating bias in observer-rated scales is a problem that is not restricted to psychological interventions (Margraf et al. 1991) and might be solved by focusing more on self-rating scales, which have been shown to assess positive symptoms with adequate reliability (Lincoln et al. in press). The MA by Lynch et al. is also not the first to take the study design into account. A separate integration of effect size according to whether studies included an active control intervention in addition to TAU or merely TAU has been conducted in other MAs which also demonstrate effect sizes to be smaller in the active control group designs, but not absent (Jones et al. 2004; Zimmermann et al. 2005; Lincoln et al. 2008). Due to the larger data basis and more transparent methodology the results from these MAs are more conclusive than those resulting from the selective methodology employed by Lynch et al.

Finally, even if the integration of all effect sizes from blind and actively controlled studies failed to find an effect for CBT, the conclusion that CBT is ineffective might be overly hasty. Although Lynch et al. set the premise that the control interventions must be unspecific, a closer look at the included control interventions reveals some of them to involve rather specific elements that are not always clearly distinguishable from CBT. For example, Durham et al. (2003) used a psycho-dynamic approach, developed to enable patients with psychosis to come to terms with past psychotic episodes and understand them in the context of their life history and feelings. The studies by Bechdolf et al. (2004) and Valmaggia et al. (2005) which also failed to find CBT superior to the control intervention used psycho-education in the control intervention, which is certainly specific and has even been demonstrated to be effective under certain conditions (Lincoln et al. 2007). Exclusion of these three studies would have increased the effect size in the MA by Lynch et al. by over 50%. Although, authors of a MA cannot be held responsible for inconsistencies in the primary outcome studies, they are responsible for selecting and integrating studies in a way that allows them to draw valid conclusions with regard to their primary hypotheses. In order to analyse the impact of CBT over and above non-specific effects resulting from therapist contact and supportive listening, an analysis of individual well-designed studies might have been more convincing.

In sum, it can be noted that Lynch et al. point the finger at some known weaknesses in the evaluation research of CBT for psychosis. However they do not add much to the existing knowledge, apart from underlining once again that effect sizes are smaller in blind studies and when there are strong control interventions. In light of the evidence from other MAs and the methodological constraints in the MA by Lynch et al. the absence of significant effects should not be over-interpreted.

**Declaration of Interest**

None.

**References**


T. M. Lincoln
Section for Clinical Psychology and Psychotherapy, Faculty of Psychology, Philipps-Universität Marburg, Gutenbergstr. 18, 35032 Marburg, Germany
(Email: lincoln@staff.uni-marburg.de)

The authors reply

Meta-analysis is, as Lincoln points out, a tool. As such, it is doubtful whether anyone can be prescriptive about how and when it should be used. Recent examples of meta-analyses which were carried out on a subset of all the available data, and which did not go into exhaustive detail in their methods, but which nevertheless had clinically useful findings include Geddes et al. (2009) on lamotrigine for bipolar depression, Cuijpers et al. (2009) on psychotherapy for depression and Leucht et al. (2009) on atypical neuroleptics for schizophrenia. None of these studies featured flow charts.

Can tinkering with the studies we included and excluded in our meta-analyses make the pooled effectiveness of CBT for schizophrenia significant? The study of Levine et al. (1998), which Lincoln highlights, had six patients in the CBT arm and six in the control (supportive therapy) arm, and so was excluded on the (stated) grounds of being too small. Adding this study (ES = −2.23) and three other small/pilot studies [Haddock et al. 1999 (n = 8, 10, ES +0.57); Turkington & Kingdon, 2000 (n = 10, 5, ES −1.14); Cather et al. 2005 (n = 15, 13, ES +0.04)] to the meta-analysis of CBT against symptoms makes little difference to the pooled effect size (−0.09, 95% CI −0.25 to 0.06, p = 0.22).

Arguing for the exclusion of the study of Hogarty et al. (1997) from the meta-analysis of CBT against relapse in schizophrenia faces two problems. First, their definition of personal therapy emphasized identification and management of psychosis-related affect dysregulation through a process of internal coping, and so conforms to definitions of CBT. Second, this study was included in the meta-analyses of Pilling et al. (2002), the Cochrane review (Jones et al. 2004), and the original and revised NICE guidelines (NICE, 2003, 2009). It should also be noted that the lack of a significant pooled effect for CBT in our meta-analysis does not depend on the inclusion of this study (pooled OR for the remaining seven studies: 1.13, 95% CI 0.84–1.52, p = 0.42).

We feel it was uncontroversial to exclude two studies that Lincoln alludes to, which compared CBT to befriending (Jackson et al. 2008) and to social skills training (Lecomte et al. 2008), because they contained significant numbers of patients with affective psychosis. On the other hand it required truly solomonic judgement to decide whether or not to include a study which compared CBT to cognitive remediational therapy (Penadés et al. 2006), an intervention which, while potentially therapeutic, would not be expected to have any effect on psychotic symptoms. As it happens, however, none of these studies found a significant advantage for CBT.

The study using motivational interviewing plus CBT, which Lincoln considers we should have included or at least justified excluding (Haddock et al. 2003), was carried out on dual-diagnosis patients, not on patients just with schizophrenia.
Lincoln’s final point is that control interventions like befriending, supportive counselling and psychoeducation might not be completely therapeutically inert. This begs the question: if CBT can not be shown to be better than these, does it really deserve such passionate advocacy?

Declaration of Interest
None.

References


PETER MCKENNA1, DAMIAN LYNCH2 AND KEITH LAWS3

1 Benito Menini CASM, Sant Boi, Barcelona, and CIBERSAM, Spain
2 Stobhill Hospital, Glasgow, UK
3 School of Psychology, University of Hertfordshire, UK

(Email: mckennapeter1@googlemail.com)