

In Coleman's synthesis, the PAR measure has been applied inappropriately and, we believe, reported misleadingly. For example, the reported PAR for completed suicide is particularly high at 35%. For several reasons, readers should not interpret this figure as meaning that over a third of all suicides among women of reproductive age could be prevented if none of them underwent abortion. An inherent assumption in the PAR is that all other things would remain equal after the removal of a risk factor, which is clearly not true for abortion in this instance. Further, the aetiology of suicide is extremely complex, and in most cases cannot be attributed to a single adverse life event that is the one measured in a particular study. Women who die by suicide at some time following an abortion are likely to carry multiple distal and proximal risk factors as they proceed along their life course, as is true for most people of any age or gender who die by suicide, and it is fallacious to suggest that abortion can be isolated from other causal factors in these limited data-sets.

Second, in the first paragraph of the Discussion (p.183), Coleman states with apparent certainty that '...nearly 10% of the incidence of mental health problems was shown to be directly attributable to abortion.' This is about as unambiguous a statement of causality as could possibly be made, in the face of clear guidance on the potential pitfalls of drawing such conclusions when applying the PAR.² Having stated the causality of the association with such certainty, the author then appears to backtrack in her concluding remarks (pp. 185–186) by making the following ambiguous statement, clearly contradicting the view expressed at the start of her Discussion:

'Although an answer to the causal question is not readily discerned based on the data available, as more prospective studies with numerous controls are being published, indirect evidence for a causal connection is beginning to emerge.'

Following publication of just such a 'prospective study with numerous controls' in the *New England Journal of Medicine* in 2011,³ it might be appropriate for Priscilla Coleman (and colleagues supportive of her views) to reconsider their conclusions. This recent study³ provides the best data available from the largest unbiased sample on the association (or lack thereof) between excess risk of mental illness and abortion because that study is based on a large population sample, with measurement of mental illness both before and after the abortion event study.³ That study 'does not support the hypothesis that there is an increased risk of mental disorders after a first-trimester induced abortion' (quoted from abstract).

- 1 Coleman PK. Abortion and mental health: quantitative synthesis and analysis of research published 1995–2009. *Br J Psychiatry* 2011; **199**: 180–6.
- 2 Rockhill B, Newman B, Weinberg C. Use and misuse of population attributable fractions. *Am J Public Health* 1998; **88**: 15–9.
- 3 Munk-Olsen T, Laursen TM, Pedersen CB, Lidegaard Ø, Mortensen PB. Induced first-trimester abortion and risk of mental disorder. *N Engl J Med* 2011; **364**: 332–9.

Kathryn M. Abel, Centre for Women's Mental Health, University of Manchester, UK. Email: kathryn.abel@manchester.ac.uk; **Ezra S. Susser**, Department of Epidemiology, Mailman School of Public Health, Columbia University, New York, New York, USA; **Peter Brocklehurst**, Institute for Women's Health, University College London, and Policy Research Unit – Maternal Health and Care, National Perinatal Epidemiology Unit, University of Oxford, Oxford, UK; **Roger T. Webb**, Centre for Women's Mental Health, University of Manchester, UK

doi: 10.1192/bjp.200.1.74a

In her review of research on the mental health effects of abortion, Coleman¹ stated:

'In this highly politicised area of research it is imperative for researchers to apply scientifically based evaluation standards in a systematic, unbiased manner when synthesising and critiquing research findings. If not, authors open themselves up to accusations of shifting standards based on conclusions aligned with a particular

political viewpoint. Moreover, the results may be dangerously misleading and result in misinformation guiding the practice of abortion.'

However, Coleman failed to follow well-accepted scientific standards for the conduct and reporting of systematic reviews and meta-analyses. Further, Coleman's failure to state her obvious conflicts of interest in this review raises serious questions about biases in her analysis. Hence, the review is open to serious questions about the author's scientific standards, methods, political viewpoints, and potentially misleading conclusions.

Widely accepted standards for systematic reviews and meta-analyses are contained in the published AMSTAR, MARS, MOOSE, and PRISMA statements.^{2–5} None of these standards were cited or followed by Coleman. AMSTAR is the only validated instrument for assessing the methodological quality of systematic reviews and meta-analysis. We assessed Coleman's review according to the AMSTAR statement, and found that it failed to meet any of the eleven basic requirements for systematic reviews and meta-analysis included in AMSTAR.

Following AMSTAR, specific flaws of the Coleman review are as follows:

- 1 there was no public *a priori* design
- 2 there was no duplicate study selection or duplicate data extraction
- 3 the author did not describe the search strategy in sufficient detail
- 4 the review was limited to published studies, contrary to all published standards
- 5 a list of excluded studies was not provided
- 6 the author did not provide sufficient descriptive information on included studies, including demographic characteristics of participants
- 7 the scientific quality of included studies was not documented
- 8 scientific quality of included studies was not considered in formulating conclusions
- 9 appropriate methods were not used in combining the findings of studies (Coleman clearly violated the rule for avoiding dependencies in meta-analysis, when she synthesised 36 effects from 22 studies in Fig. 1)
- 10 the likelihood of publication bias was not assessed
- 11 conflicts of interest and sources of support were not acknowledged (no financial disclosures were made and no other potential conflicts were acknowledged).

An article in the *British Journal of Psychiatry*⁶ calls attention to the importance of non-financial conflicts of interest in the psychiatric literature. Coleman has at least two types of conflict of interest here. Among the most important of such conflicts is an agenda-driven bias, by which authors seek to influence legislation and social policy. David Reardon is a co-author with Coleman on seven articles included in the review and an author on an additional study in the review that does not involve Coleman as a co-author. Reardon is quite explicit about his agenda to instil fear of abortion as a way of facilitating passage of anti-abortion legislation.⁷

Coleman is the first author on 6 studies and co-author on 5 additional studies in her review; thus, she authored or co-authored fully half of the 22 studies included. According to the *Cochrane Handbook*,⁸ this is another potential conflict of interest, since it may 'unduly influence judgements made in a review (concerning, for example, the inclusion or exclusion of studies, assessments of the risk of bias in included studies or the interpretation of results) ... This should be disclosed in the review

and, where possible, there should be an independent assessment of eligibility and risk of bias by a second author with no conflict of interest.' Coleman did not obtain an independent assessment of the studies she authored or co-authored, nor did she acknowledge these conflicts in the review.

Coleman's conclusion that the results of the studies in her review are 'quite consistent' (p. 183) is belied by visual inspection of the Forest plots, which include non-overlapping confidence intervals. Coleman should have reported results of heterogeneity tests (chi-squared and I^2), which probably would have shown significant heterogeneity in results across studies (presumably that is why she chose the random effects model).

Some of the commentaries on Coleman's review appear to be uninformed by current scientific standards for reviews. Comments by Fergusson *et al* are particularly misleading. Faced with variations in the methodological quality of available studies, it is essential for reviewers to weed out weaker studies. Valid conclusions can only be based on valid studies.

It is unclear how this paper got through peer review at the *Journal*. It appears that peer reviewers and the Editor ignored published standards for systematic reviews and meta-analyses. Given the serious methodological flaws contained in Coleman's review and the author's failure to report obvious conflicts of interest, we believe the article should be retracted.

- 1 Coleman PK. Abortion and mental health: quantitative synthesis and analysis of research published 1995–2009. *Br J Psychiatry* 2011; **199**: 180–6.
- 2 Shea BJ, Grimshaw JM, Wells GA, Boers M, Andersson N, Hamel C, et al. Development of AMSTAR: a measurement tool to assess the methodological quality of systematic reviews. *BMC Med Res Methodol* 2007; **7**: 10.
- 3 APA Publications and Communications Board Working Group on Journal Article Reporting Standards. Reporting standards for research in psychology: Why do we need them? What might they be? *Am Psychol* 2008; **63**: 839–51.
- 4 Stroup DF, Berlin JA, Morton SC, Olkin I, Williamson GD, Rennie D, et al. Meta-analysis of observational studies in epidemiology: a proposal for reporting. *JAMA* 2000; **283**: 2008.
- 5 Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLOS Med* 2009; **6**: e1000097.
- 6 Maj M. Non-financial conflicts of interests in psychiatric research and practice. *Br J Psychiatry* 2008; **193**: 91–2.
- 7 Reardon DC. A defense of the neglected rhetorical strategy (NRS). *Ethics Med* 2002; **18**: 23–32.
- 8 Higgins JPT, Green S (eds). *Cochrane Handbook for Systematic Reviews of Interventions*. Version 5.1.0. The Cochrane Collaboration, 2011.

Julia H. Littell, Graduate School of Social Work and Social Research, Bryn Mawr College, Bryn Mawr, Pennsylvania, USA. Email: jlittell@brynmawr.edu; **James C. Coyne**, Department of Psychiatry, Perelman School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania, USA

doi: 10.1192/bjp.200.1.75

Priscilla Coleman's recent meta-analysis¹ ignores guidelines for proper scientific conduct of meta-analyses of observational data. Her results violate at least three major principles of meta-analysis: she fails to assess the underlying validity of included studies; she fails to examine statistical heterogeneity; and she illogically combines estimates for distinct outcomes. Furthermore, she accuses previous reviews of lacking 'reasonable justification' for declining to quantitatively summarise effects, when declining to do so actually reflected sound epidemiological judgement.

Coleman contends that 'Through a process of systematically combining the quantitative results from numerous studies addressing the same basic question ... far more reliable results are produced than from particular studies that are limited in size and scope'. However, expert consensus suggests that 'the likelihood that the treatment effect reported in a systematic review

approximates the truth depends on the validity of the included studies...'² Coleman fails to assess the validity of included studies and erroneously asserts that 'as a methodology wherein studies are weighted based on objective scientific criteria, meta-analysis offers a logical, more objective alternative to qualitative reviews...'. In fact, studies in meta-analyses are typically weighted by sample size, which is not always related to study quality,³ and decisions on which studies to include and how to include them remain subjective. If poor-quality studies are included, as occurred in Coleman's review, a poor-quality quantitative estimate will be generated. Coleman combines statistically heterogeneous results, and illogically combines effect estimates for outcomes that vary substantially (i.e. marijuana use and suicide), thus generating a summary estimate void of meaning or utility.

Meta-analysis of observational data can be useful when carefully conducted. However, it is essential that a summary estimate be accompanied by a qualitative description of risk of bias in included studies (which Coleman's review lacked) since 'potential biases in the original studies, relative to biases in RCTs, make the calculation of a single summary estimate of effect of exposure potentially misleading'.⁴

Coleman ignores other essential requirements of a high-quality statistical meta-analysis.² She makes no attempt to present a replicable search strategy or article selection diagram. She attempts to justify excluding articles prior to 1995 by noting that study methodology has improved, but fails to adequately justify selected cut-off dates. Ultimately, she includes multiple methodologically weak studies, and excludes at least two older but methodologically stronger studies. She authored her review alone, despite Cochrane and PRISMA recommendations to involve multiple reviewers to reduce the possibility of investigator bias or error.^{2,5}

Coleman makes disingenuous accusations about previous reviews. For example, she claims that our 2008 systematic review⁶ 'overlooked' ten articles which met inclusion criteria, and 'lacked sufficient methodologically based selection criteria'. This unfounded attack is puzzling, particularly since in 2008, we directly emailed to Coleman the reasons (consistent with our methodologically based selection criteria detailed on p.437) for excluding seven of these ten articles. The remaining three (not previously enquired about) also fail to meet inclusion criteria: two had a follow-up period of less than 90 days and the other compared medical *v.* surgical termination.

Coleman continues to ignore the scientific importance of accounting for pregnancy intention in this body of literature. If women who abort (many of which are unintended pregnancies) are compared against women who deliver (many of which are intended pregnancies), effects of unintended pregnancy are difficult to disentangle from effects of abortion. Circumstances surrounding an intentional *v.* an unintentional conception or pregnancy may be related to mental health outcomes. Most aborted pregnancies in the USA were unintended.⁷ Coleman wrongly assumes that since nearly half of pregnancies in the USA are unintended, most births are too, failing to acknowledge that almost half of unintended pregnancies end in abortion.⁸ Thus, her assertion that 'the majority of women in the control groups in studies comparing abortion with term pregnancy actually delivered unintended pregnancies even if the variable was not directly assessed' has no empirical grounding. Similarly, her assertion that a 'no pregnancy' group may be a 'cleaner' comparison group ignores the fact that the 'no pregnancy' group would not have experienced unintended pregnancy.

The scientific validity and rigour of Priscilla Coleman's work has been questioned before.⁹ However, we are surprised and disappointed that the multiple egregious scientific errors in her