BJPsych | The British Journal of Psychiatry (2012) 200, 446–453. doi: 10.1192/bjp.bp.111.098103

Review article

Methodology and reporting of systematic reviews and meta-analyses of observational studies in psychiatric epidemiology: systematic review

Traolach S. Brugha, Ruth Matthews, Zoe Morgan, Trevor Hill, Jordi Alonso and David R. Jones

Background
Relatively little is known of the use of systematic review and synthesis methods of non-randomised psychiatric epidemiological studies, which play a vital role in aetiological research, planning and policy-making.

Aims
To evaluate reviews of psychiatric epidemiological studies of functional mental disorders that employed synthesis methods such as systematic review or meta-analysis, or other forms of quantitative review.

Method
We searched the literature to identify appropriate reviews published during the period 1996 to April 2009. Selected reviews were evaluated using published review guidelines.

Results
We found 106 reviews in total, of which 38 (36%) did not mention method of data abstraction from primary studies at all. Many failed to mention study quality, publication bias, bias and confounding. In 73 studies that performed a meta-analysis, 58 (79%) tested for heterogeneity and of these, 47 found significant heterogeneity. Studies that detected heterogeneity made some allowance for this. A major obstacle facing reviewers is the wide variation between primary studies in the use of instruments to measure outcomes and in sampling methods used.

Conclusions
Many deficiencies found in systematic reviews are potentially remediable, although synthesis of primary study findings in a field characterised by so many sources of heterogeneity will remain challenging.

Declaration of interest
None.

Systematic reviews and (where appropriate) meta-analysis have potentially great value in combining evidence from primary studies to inform policy- and decision-making with more accurate evidence syntheses than those available from narrative reviews. Such methods have been extensively used to summarise treatment evidence in the fields of clinical psychology and psychiatric therapeutics. Relatively little is known of the use of such methods in non-randomised psychiatric epidemiology studies, which have a vital role in aetiological research, planning and policy-making. There are examples of reviews of the use of systematic reviews and meta-analysis in various fields parallel to that performed here, as diverse as acupuncture and animal experiments, the latter including reporting guidelines.[1,2] More general guidelines for reporting of meta-analyses of observational studies are also available.[3]

Systematic review methods have been developed for use in medical research, including observational epidemiological studies.[4] However, the field of mental and behavioural disorders may pose particular challenges to systematic reviewers because of the relatively fragile understanding of mental health outcomes and determinants compared with physical health,[5,6] and because biologically based gold standard measures are not available. Outcomes include, for example, anxiety (and 'stress symptoms'), depression, functional psychosis (hallucinations, delusional beliefs) and physically unexplained somatic symptoms (unexplained pain), which make use of largely theoretically based definitions and measures, which are more difficult to assess validly in comparison with substance use and organically induced disorders. Observational studies also assess risk factors or determinants that are based on similar, theoretically based definitions and measures (adversity, personality, functioning), which could give rise to heterogeneity across different studies. The potential challenge of heterogeneity to study comparability may be due to the range of different measures used to assess such constructs.[7] Differences in study design and, in particular, sample design may also limit the comparability of different studies.

We anticipated that the synthesis of results could usefully be applied to two kinds of epidemiological study: methods for synthesising associations with disorder (i.e. risk factors) and methods for synthesising prevalence estimates. Reliable inferences from syntheses of prevalence estimates may be the more difficult of these because of their potential sensitivity to differences between study contexts and methods for measuring whether diagnostic criteria for a given disorder are met, which may result in heterogeneity and caveats related to combinability.

In this paper we report findings of a systematic review of systematic reviews of studies of non-organic mental disorders that make use of representative epidemiological samples, to estimate disease frequency and/or association with potential risk factors. We aimed to review the uses – good and bad – of synthesis methods in published reviews, giving reasons with examples for the recommended use of such methods. Our objective was to review all such methods, and not all the literature in which such methods are used. Having examined systematic reviews published up to 2005, we decided that it would be more useful for a scientific article to compare that initial period with more recent reviews and to consider whether the quality of more recent systematic reviews differed from earlier ones, although this was not an original aim of our study.

Method

We searched EMBASE, MEDLINE and PsycLIT to identify reviews of psychiatric epidemiological studies (including two or more primary population studies) that employed synthesis methods such as systematic review or meta-analysis or other forms of quantitative review. Initially, we searched the period from 1996

https://doi.org/10.1192/bjp.bp.111.098103 Published online by Cambridge University Press

to July 2005 (a summary of the findings contributed to a European Public Health Action Report);[8] the second period considered reviews up to April 2009. Search terms are shown in Appendix 1, and were designed to be sensitive (potentially over-inclusive) to avoid missing any relevant articles. The search strategy was developed by one of the authors (a subject specialist, T.S.B.) and an information officer (Mary Edmunds Otter, Department of Health Sciences, University of Leicester, UK). The filter used to identify systematic reviews was adapted from two strategies recommended by the Centre for Reviews and Dissemination.[9]

Abstracts were obtained for all papers identified in the electronic searches, and reviewed independently by three authors (R.M., T.S.B., Z.M.) for inclusion, according to criteria drawn up by T.S.B. The review included studies in which the health outcomes included the functional psychoses (ICD-10, Chapter V, code F2), mood disorders (ICD-10 code F3) and neurotic disorders (ICD-10 code F4). Studies of 'hard outcomes' for which there are sufficiently clear and established approaches to and examples of synthesis review methods (survival, suicide, organic brain disorder such as dementia and brain damage, drug or alcohol misuse) were not included in this review, although research on these more clearly definable outcomes may also be needed. Further details of inclusion and exclusion criteria are available on request. Any disagreements were discussed by the authors undertaking data extraction, and a consensus reached on whether the paper should be included. Full-text articles were retrieved for all studies identified as potentially relevant from the abstracts, as well as for those where their relevance was unclear. Each full-text paper was reviewed by one author (R.M., Z.M. or T.H.) to establish whether it met inclusion criteria. Where it was unclear whether an article should be included, a separate reviewer (T.S.B.) also read the paper and a consensus was reached on its inclusion. The selected reviews were evaluated using guidelines drawn up by the authors, based on the literature.[3,10,11] Reviews were classified as prevalence or association studies and according to whether meta-analyses had been performed; papers using weighted averages were included with those described as 'meta-analyses' even where the term was not employed.

Classification of reviews

Classification of articles as studies of either prevalence or association was occasionally problematic. Two articles by Saha *et al* and one from Singer appeared at first glance to be concerned with the association of a mental disorder with an outside factor, and would be expected to be coded accordingly.[12–14] However, on further scrutiny it became clear that these reviews were meta-analysing prevalence data, and then performing further non-meta-analytical methods to produce summary rates. For the purposes of this review these articles have been treated and classified in the tables as prevalence studies.

Coding of reviews

During the process of extracting information from the articles (online Appendix DS1) and its inclusion in our tables there were many instances where reviews failed to mention important issues such as whether publication bias was assessed or how (and whether) study quality was assessed. In these cases the article was coded as not having mentioned that particular method. In addition, several reviews by Saha *et al* that re-analysed data from two previous reviews did not include this information[12,13,15] and directed the reader to the original reports.[16,17] Therefore these reviews have been coded in the same way as original reviews that did not mention such criteria.

Results

We identified 1153 articles from the search strategies, of which we classed 245 as potentially relevant after reading the abstracts (Fig. 1). After these 245 papers had been read in full, 103 were selected as relevant. A further 4 papers were identified through searching the reference lists of the reviewed papers. Of the 107 articles, 32 focused on prevalence only, 17 looked at prevalence and association, and 58 reported only on associations with disorder. One article by McGrath *et al*,[18] focusing on prevalence, was a review of previous reviews,[16,17] and is discussed here but not included in the tables of results. The total number of papers therefore represented is 106. A summary of these papers is given as an online supplement.

Authors generally gave comprehensive details of search strategies employed, including details of electronic databases searched, exact search terms, dates covered by search and other methods used. Only four studies gave no details about the search strategy. One hundred and one (95%) reviews searched at least one electronic database, with eighty-six (81%) searching two or more databases (Table 1). The most common database searched was MEDLINE, although PsycLIT, PsycINFO, EMBASE, Ovid, PubMed and CINAHL were also searched frequently.

The majority of studies gave details of inclusion and exclusion criteria used to select individual studies for detailed review (96 studies, 91%). Only 35 gave details of the method used to apply these criteria; of these, 3 used a single reviewer, 28 used several reviewers, reaching a consensus where disagreements arose, and 4 took a sample of studies for consensus review. Just under half of the reviews (50 studies) gave the proportion of initially identified studies that met selection criteria.

Extraction of data and assessment of study quality

Sixty-eight (64%) studies gave details of guidelines used to abstract data, although only forty-five (42%) described the method of abstraction (Table 2). Of the 106 studies, 63 (59%) did not mention the method of abstraction: 38 studies (36%) made no mention of data abstraction at all; 25 (24%) studies mentioned data abstraction but not the method of abstraction; 45 (43%) described the method of abstraction (studies may have mentioned more than one method and therefore the total exceeds 106). Seventy (66%) studies made some mention of study quality, with nineteen formally assessing the quality of primary studies, and four carried out a sensitivity analysis excluding studies of poor quality.

Methods of synthesis

In the 48 papers concerned with meta-analysing prevalence, the most common method (13 papers) was the use of random or fixed effects meta-analysis models. Calculation of means, weighted for study size, was used in 10 studies, whereas 3 studies used a Bayesian approach to synthesis.[19–21] Of 22 studies that stated they did not perform a meta-analysis, 13 gave a summary measure such as the median prevalence, or a range of prevalences, with 7 studies giving the prevalence for individual studies in addition. Nine studies gave individual results only.

For studies of risk factors associated with psychiatric disorder or outcome, the most commonly employed methods of analysis were fixed or random effects meta-analysis models (24 studies), with 14 further papers using approaches described as weighted averages and two others the Mantel–Haenszel pooled odds

https://doi.org/10.1192/bjp.bp.111.098103 Published online by Cambridge University Press

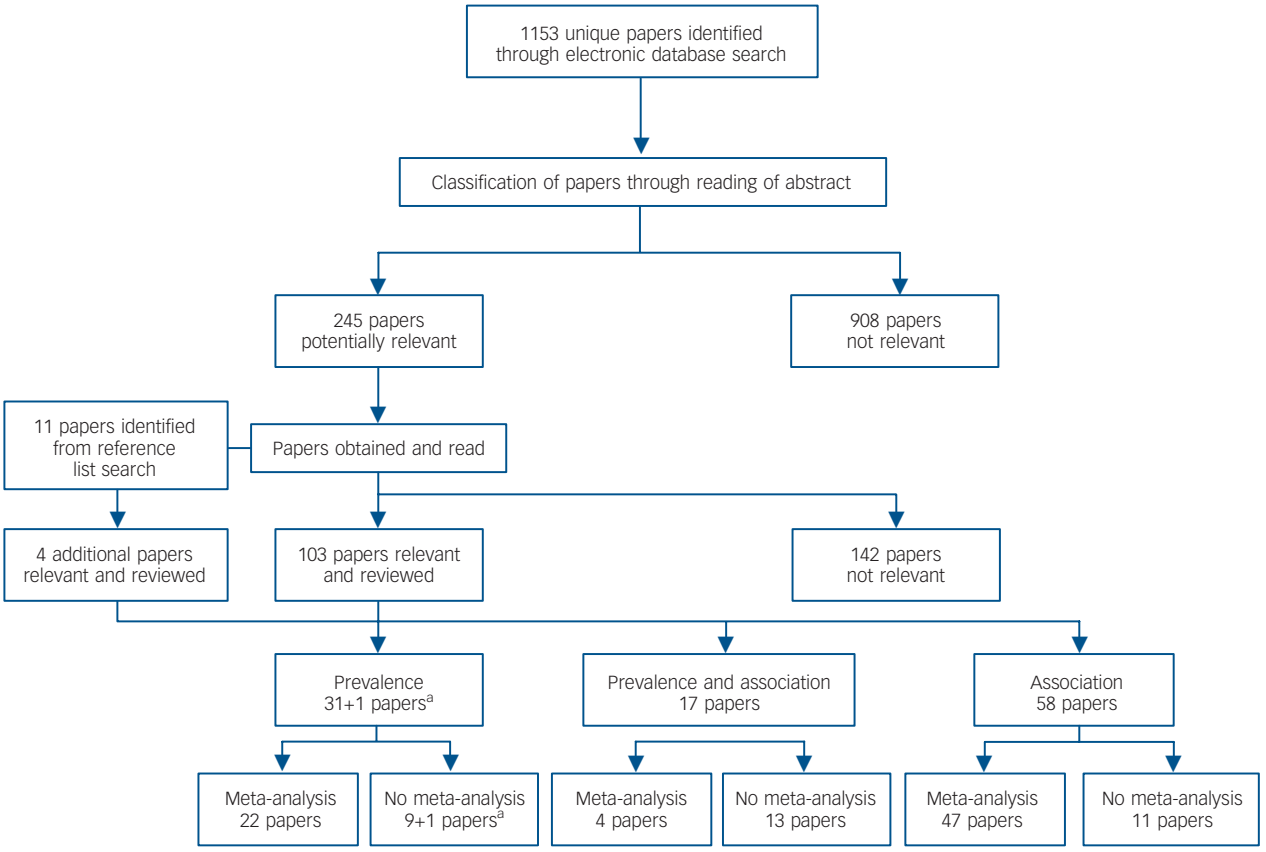

Fig 1 Flow chart of identified studies.

a. Data from the paper by McGrath *et al* is not included in the tables or main results.

Table 1 Search strategies employed

	Number of studies
Number of databases searched	
None	1
One	15
Two	30
Three	16
Four	18
Five or more	22
Not stated	4
Were search terms given?	
Exact search terms	69
Details given, but not exact terms	18
No information given	19
Other methods used for identifying primary papers[a]	
Reference lists of primary papers searched for additional papers	79
Contacted authors of primary papers identified in initial search	20
Contacted lead researchers in subject area	23
Hand-searched journals	17
Searched unpublished data or websites	11

a. Total exceeds 106 as studies used more than one method. Table does not include study by McGrath *et al*.[18]

Table 2 Data abstraction and assessment of primary study quality

	Number of studies
Data abstraction	
No mention	38
Method stated for abstraction	68
Method of abstraction[a]	
No mention	63
Two or more independent reviewers abstracted data	39
Random sample abstracted by two independent reviewers	3
Abstracted data checked by another reviewer	3
Study quality[a]	
No mention	36
Inclusion criteria included factors related to quality	40
Rated quality of primary studies	19
Discussed quality of primary studies	18
Sensitivity analysis excluding studies of poor quality	4
Total	106

a. Studies may have used more than one method of data abstraction or of assessing quality, so the total exceeds 106. Table does not include data from McGrath *et al*.[18]

ratio.[22,23] Other methods used were mixed effects Poisson regression,[24] unconditional logistic regression,[25] weighted least squares regression,[26,27] and a Bayesian hierarchical random effects model.[28] Two studies did not specify the exact method of analysis.[29,30] Of the 24 studies that did not employ a meta-analysis, 5 gave individual study results and 2 gave a narrative summary of the results. Eleven studies gave both a narrative summary and individual study results.

Testing and exploring heterogeneity

Between-study heterogeneity is a common feature of synthesis and has important implications for inferences drawn from it.

https://doi.org/10.1192/bjp.bp.111.098103 Published online by Cambridge University Press

Table 3 Heterogeneity in the 73 studies that used meta-analysis

	Number of studies
Did not mention heterogeneity	5
Discussed heterogeneity (but no formal test performed)	10
Tested for heterogeneity	58
Q statistic	34
Chi-squared test	16
Breslow–Day test	2
Other tests	6
Significant heterogeneity found[a]	47
Random effects model used	30
Investigated sources	20
Removed outliers	11
Discussed sources	3
No significant heterogeneity found	11

a. More than one may apply to each study.

Only 5 of the 73 studies that employed a meta-analysis made no reference to heterogeneity. Fifty-eight studies formally tested for heterogeneity (Table 3). The most common tests used were the *Q* statistic and chi-squared test. Other methods used included testing whether the mean effects variance was null,[31] testing for an interaction with study design,[25] testing whether individual results differed from others,[32] and the I^2 statistic.[33–35]

Out of the 26 studies that synthesised prevalence, 14 found significant heterogeneity in their estimates and 5 discussed heterogeneity without formally testing for it. Of the 40 studies synthesising association only which tested for heterogeneity, heterogeneity was not statistically significant in 8. However, limitations of tests for heterogeneity due to lack of power are well known.[36] Significant heterogeneity was found in all but 2 (Singer and Costello *et al*) of the 16 reviews of prevalence studies in which tests for it were reported.[14,37] In only 4 of these reviews of prevalence studies[34,38–40] was heterogeneity largely eliminated, for example by providing estimates for women only,[39] and by removing outliers that appeared to explain heterogeneity,[34,40] such as high- and low-risk samples, differences in period or point prevalence definition or in the diagnostic measure used.[38] Several reviews carefully grouped studies that could be said to be homogeneous and then performed a formal meta-analysis taking account of error in each estimate.[20,38,39,41,42] Grouping more homogeneous studies in this way appeared to improve precision as reported in three of these reviews.[20,38,39] However, in one review this led to larger confidence intervals, which might be due to the limited number of studies available for inclusion, with the result that the review was less conclusive.[38]

Studies that detected significant heterogeneity made some allowance for this in analysis through the use of random effects models, by removing results that were outliers, or through controlling for moderator variables. Skeem *et al* made a thorough examination of heterogeneity by removing outliers, exploring the effects of moderator variables and using random effects models as well as performing sensitivity analyses.[43]

Thirty-three studies did not perform a meta-analysis; fifteen of these gave heterogeneity between individual studies as the reason for not doing so, and four stated that they were unable to assess heterogeneity and so did not combine the data in a meta-analysis. The remaining articles either did not mention heterogeneity at all or mentioned it but were unclear as to why they did not attempt a meta-analysis.

Table 4 Publication bias, bias and confounding

	Number of studies
Publication bias	
No mention	59
Discussed	18
Tested	29
Method of detecting publication bias	
Funnel plot	14
Fail-safe effect	16
Other	7
Publication bias detected	
No	20
Yes	9
Bias[a]	
No mention	53
Discussed bias	49
Analysis to explore bias	16
Steps taken to limit bias (e.g. exclusion criteria for review)	9
Confounding[a]	
No mention	46
Discussed confounding	40
Adjusted for confounders in analysis	35
Steps taken to limit confounding (e.g. exclusion criteria for review)	7

a. Studies used more than one method, so total exceeds 106. Table does not include data from McGrath *et al*.[18]

Publication bias

Just over half of the studies did not mention publication bias (Table 4). Nine studies did not assess publication bias in any formal way, but stated that it was unlikely that it would have affected their results. The most frequent reasons given for publication bias being unlikely were that unpublished studies had been included (3 studies) and that the question being asked in the review was not the main research question of the papers identified (2 studies), although in fact publication bias is not limited to the primary outcome. Three studies stated that their results were likely to be affected by publication bias, without formally assessing it. Twenty-nine studies assessed publication bias, with the majority using funnel plots,[44,45] or the fail-safe (or 'file drawer') method.[46,47] Other methods of assessing publication bias were the Begg–Mazumdar adjusted rank correlation test,[48,49] the Egger test,[34,50] and looking at the correlation between the variance and log odds ratio.[51] Of the studies that detected publication bias, only four discussed the effects this might have had on their findings.

Bias and confounding in observational studies

Bias and confounding pose particular challenges for observational studies, and may thus affect the conclusions of meta-analyses and reviews of such studies. Around half of the studies made no mention of any bias (53 studies) or confounding factor (46 studies) that might have affected the results (Table 4). Of those that did mention bias, 9 studies took steps to avoid bias affecting their results through the use of inclusion or exclusion criteria for the papers included in their review, and 16 explored bias by looking at the effect of various factors, including sample type and type of assessment.[14,27,32,41,48,49,52–61] Studies adjusted for confounding factors including age, gender, education, work status, social support, severity and duration of symptoms and disability, and whether methodological factors could account for differences including source of recruitment, sample size, diagnostic criteria

https://doi.org/10.1192/bjp.bp.111.098103 Published online by Cambridge University Press

and type of interview. Several studies also adjusted for publication year and geographic location.

The two-phase review allowed the possibility of noting apparent trends in the quality and characteristics of meta-analyses in psychiatric epidemiology between two time periods. However, few such trends were clear-cut, with perhaps only four being worthy of note. For example, although only around a quarter of reviews (15 of 61 papers) mentioned or discussed confounding in the initial review period, more than half (25 of 45 papers) in the second period either discussed or, at the very least, made a mention of confounding. Reviews in the more recent period were also more likely to give their exact search terms, with few (3 of 45 papers) not giving this information, as opposed to around a quarter (16 of 61) in the earlier period. Reviews in the recent period were also much more likely to state the actual method of abstraction, with around three times as many reviews using two or more independent reviewers to abstract the data compared with reviews in the initial period (20% *v.* 60%). There were around twice as many reviews (23% *v.*11%) in the first period that rated the quality of primary studies compared with the recent, update period.

Discussion

This review found a number of deficiencies in the conduct and reporting of systematic reviews and meta-analyses of observational psychiatric epidemiology studies that could have serious implications for inferences drawn or decisions made on the basis of these reviews. There were frequent omissions of descriptions of method of abstraction, study quality, publication bias, bias and confounding. Many of these deficiencies are simple and potentially remediable. Of the 106 studies examined, 73 performed a meta-analysis but only 58 tested for heterogeneity. There were also some terminological issues, with the most important being the description of the quantitative synthesis method adopted. Thus to the several examples described as meta-analyses by their authors (and yielding weighted average estimates of prevalences or odds ratios) need to be added investigators who calculated weighted averages but did not describe their method as a meta-analysis.[62,63] As indicated above there were some dimensions of meta-analytical practice which appear to be improving between the two periods considered.

In 47 reviews heterogeneity was detected and reported; this needs to be followed by an exploration of sources of and modelling of heterogeneity. In half of the meta-analyses of studies of association with mental disorder in which heterogeneity was tested it occurred for some but not all risk factors; in a minority, heterogeneity was completely absent. All studies that detected heterogeneity made some allowance for this in analysis, through the use of random effects models, removing results that were outliers, or controlling for moderator variables.

Limitations of meta-analysis

The majority of reviews reported pooled mean prevalence estimates, in most cases with narrow confidence intervals, but the range of prevalence estimates across such heterogeneous studies was considerable (prevalence estimates ranging from 5% to over 40% were not unusual),[62,64] with occasionally much lower rates such as in Somers *et al* whose estimates for individual anxiety disorders ranged from less than 1% to around 6%.[21] Use of a single summary estimate in these circumstances in planning and economic projections is not as appropriate or useful as use of a set of more specific estimates reflecting important dimensions of the heterogeneity. Such limitations call for far greater caution in the interpretation of comparisons of prevalence estimates, which are often reported in the scientific literature. Most of these reviews, if carefully studied in detail, provided some additional information on the limitations of the available data; but unless studied by readers with specialist knowledge of the data synthesis methods used, headline overall prevalence estimates could quite easily mislead policy makers and the wider community of those concerned about the burden and cost of mental disorder in the general population. McGrath *et al* and Saha *et al* also cautioned against the use of standard methods of meta-analysis compared with simpler methods of representing prevalence findings from different studies,[15,18] for example using median values,[18] and graphical representations of the variation of estimates around the central value.[15]

Sensitivity analyses have an important role in meta-analysis, for example investigating whether dropping or adding primary studies with (say) slightly non-standard disease definitions makes a difference. A major obstacle facing reviewers, which many acknowledge or discuss, is the wide variation in the use of instruments to measure outcomes and in sampling methods used. Thus quantitative reviews may sometimes be based on combining different measures which should not always be so combined – 'apples and oranges' in a classic of the meta-analytical literature.[65] One review of post-traumatic stress disorder prevalence studies found a consistent threefold difference in estimates between two commonly used diagnostic methods.[64] Systematic reviews are usually useful and valid if well performed, but meta-analyses may be unduly common in this branch of the psychiatry literature given the considerable debate about the validity of meta-analysis in observational studies. Sometimes a single high-quality, well-reported study can be recommended instead of a statistical synthesis of heterogeneous studies.

All systematic reviews found by us, apart from our own,[66] included studies using different instruments and/or definitions of disorder, or failed to specify how outcomes were defined or measured. Heterogeneity was not significant in only seven studies of associations, four of which addressed a relatively reliable risk factor (season of birth, complications of pregnancy and labour, gender).

Need for guidelines

Currently there are no recommended guidelines for good-quality reporting of meta-analyses of observational studies specifically in psychiatric epidemiology, but more general guidelines for meta-analysis of observational studies such as the Meta-analysis of Observational Studies in Epidemiology (MOOSE) guidelines are relevant.[3] Systematic review and meta-analysis in this (psychiatric) field share many issues with the use of these techniques in other fields in which the results of observational as opposed to experimental primary studies are to be synthesised. In particular, although meta-analysis may improve the precision of estimates – of prevalence, or of odds ratios for association – it does so at the potential cost of conflating results of different primary studies subject to different types and degree of bias, rendering greater precision largely worthless.[67,68] Some of these biases may be associated with the use of varying definitions and or measures of outcomes and perhaps exposure variables; in such cases coordinated studies across several centres, using uniform approaches, will almost always be preferable if feasible.[69,70]

Comparing the two periods of our review for trends in the quality and characteristics of meta-analyses in psychiatric epidemiology, few clear-cut trends emerged. Although our review allowed for a comparison of two periods, a possible limitation is that the second phase ran only to 2009. Therefore a secular trend

https://doi.org/10.1192/bjp.bp.111.098103 Published online by Cambridge University Press

of improvement in study execution and reporting since our review period may have taken place and indeed would be hoped for.

Other between-study differences arising because of different populations and contexts between studies may be of interest and importance to identify and quantify. Thus the exploration and explanation of sources of heterogeneity is important here as elsewhere; in these circumstances meta-analyses are better deployed as exploratory tools rather than to establish definitive estimates. Where they are appropriately employed, reference to guidelines and checklists for their implementation should promote quality in their execution.[3,67,68,71,72] Consideration should be given to development of guidelines for systematic review and meta-analysis in psychiatry, developing existing guidelines with more emphasis on the issues of disease definition, measurement instruments used, and population sampling, which are especially important in psychiatry. Initial proposals for guidelines are given in online Appendix DS2. These are modified from a comparable review of the use of systematic reviews and meta- analysis and guidelines on reporting,[2] and on guidelines based on the Quality of Reporting of Meta-analyses (QUOROM) statement,[73] which include features of the MOOSE statement and further modifications making the guidelines more specific to our subject.

Recommendations for primary studies

The recommendation by Fryers *et al* also point to a need, apparently unmet, for recommendations as to the design, conduct, analysis and reporting of the primary studies on which systematic reviews or meta-analyses draw.[66] These could include:

(a) desirability of prospective registration of primary studies, to include specification of key hypotheses and analyses proposed;

(b) use of (including reporting of results in terms of) standard definitions of disease (defined by collaborative groups and networks) as well as the authors' own;

(c) full reporting of factors for which allowance has been made in design (e.g. by restricting samples) or analysis (e.g. by inclusion as regression covariates) – or agreement on a standard list always to be used at least in secondary analyses, as in (b) above;

(d) full quantitative reporting of key results to facilitate meta-analyses, such as numbers at risk and numbers of cases for prevalences, numbers at risk and events in each group for odds ratios, etc., and (adjusted) odds ratios and confidence intervals.

Reviewers should also consider whether and when a quantitative synthesis method such as meta-analysis is the correct approach to studies using heterogeneous methods; a minority of reviewers reported their decision not to use such quantitative synthetic methods and employed alternative methods. In the case of systematic reviews of treatment studies for which meta-analytic methods are not appropriate, effects may often still be examined to provide a systematic assessment of the evidence available.[72] Some combinations of research objectives, evidence types, contexts and resources may be better matched by alternative approaches.[74] Boaz *et al* indicated wide-ranging interest in synthesis methods of different types in different fields,[74] and psychiatric epidemiologists might want to explore which would be useful in the challenging area of the study of mental disorders in populations.

Funding

This project was funded by the European Commission (contract QLG5-1999-01042), Fondo de Investigación Sanitaria (FIS 00/0028-02), the Ministry of Science and Technology, Spain (SAF 2000-1800-CE), the Piedmont Region, Italy, other local agencies and by an unrestricted educational grant from GlaxoSmithKline. The European Study of the Epidemiology of Mental Disorders (ESEMeD) survey is carried out in conjunction with the World Health Organization World Mental Health Survey Initiative.

Acknowledgements

We thank the World Mental Health coordinating staff both for their assistance with instrumentation and for their consultation on field procedures (more information is available at http://www.epremed.org). We acknowledge the support of the European Study of the Epidemiology of Mental Disorders (ESEMeD)/MHEDEA 2000 investigators (Jordi Alonso, Matthias Angermeyer, Herbert Matschinger, Sebastian Bernert, Ronny Bruffaerts, Traolach S. Brugha, Giovanni de Girolamo, Ron de Graaf, Koen Demyttenaere, Isabelle Gasquet, Josep Maria Haro, Steven J. Katz, Ronald C. Kessler, Viviane Kovess, Jean Pierre Lépine, Johan Ormel and Gemma Vilagut). Mary Edmunds Otter, Department of Health Sciences, University of Leicester, UK, provided library information support.

Traolach S. Brugha, MD (NUI), **Ruth Matthews**, MSc, **Zoe Morgan**, MSc, **Trevor Hill**, MSc, Division of Psychiatry, Department of Health Sciences, University of Leicester, Leicester, UK; **Jordi Alonso**, PhD, Health Services Research Unit, Institut Municipal d'Investigació Mèdica (IMIM-Hospital del Mar) and Centre of Biomedical Research Network in Epidemiology and Public Health (CIBERESP), Barcelona, Spain; **David R. Jones**, PhD, Centre for Biostatistics and Genetic Epidemiology, Department of Health Sciences, University of Leicester, Leicester, UK

Correspondence: Professor T. Brugha, Department of Health Sciences, Division of Clinical Psychiatry, Leicester General Hospital, Leicester LE5 4PW, UK. Email: tsb@leicester.ac.uk

First received 26 Feb 2009, final revision 5 Oct 2011, accepted 30 Nov 2011

Appendix 1

Table A1 Search terms used to identify systematic reviews

To identify	Search terms
Studies of psychiatric illness	Mental disorders (exploded term) OR mental illness OR mental disease OR psychiatr*
AND	
Observational studies	Birth cohort OR longitudinal study OR cohort analysis OR epidemiologic methods OR follow up studies OR follow-up studies OR prospective studies OR incidence stud* OR ep* OR epidemiology OR epidemiological studies
AND	
Systematic reviews or studies using synthetic methods	Meta analys* OR quantitative review* OR quantitative synthes* OR review synthe* OR research synthes* OR ''Systematic review'' (Keyword) OR quantitative overview

References

1 Derry CJ, Derry S, McQuay HJ, Moore RA. Systematic review of systematic reviews of acupuncture published 1996–2005. *Clin Med* 2006; **6**: 381–6.

2 Peters JL, Sutton AJ, Jones DR, Rushton L, Abrams KR. A systematic review of systematic reviews and meta-analyses of animal experiments with guidelines for reporting. *J Environ Sci Health B* 2006; **41**: 1245–58.

3 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. Meta-analysis of Observational Studies in Epidemiology (MOOSE) group. *JAMA* 2000; **283**: 2008–12.

4 Sutton AJ, Jones DR, Abrams KR, Sheldon TA, Song F. Systematic reviews and meta-analysis: a structured review of the methodological literature. *J Health Serv Res Policy* 1999; **4**: 49–55.

https://doi.org/10.1192/bjp.bp.111.098103 Published online by Cambridge University Press

5 Brugha TS, Bebbington PE, Jenkins R. A difference that matters: comparisons of structured and semi-structured diagnostic interviews of adults in the general population. *Psychol Med* 1999; **29**: 1013–20.

6 Kessler RC, Abelson J, Demler O, Escobar JI, Gibbon M, Guyer ME, et al. Clinical calibration of DSM-IV diagnoses in the World Mental Health (WMH) version of the World Health Organization (WHO) Composite International Diagnostic Interview (WMHCIDI). *Int J Methods Psychiatr Res* 2004; **13**: 122–39.

7 Thompson C, Thompson C. *The Instruments of Psychiatric Research*. Wiley, 1988.

8 EPREMED Consortium. *European Policy Information Research for Mental Disorders: Final Report of the Project SANCO 2004123*. EU Commission, 2008.

9 White VJ, Glanville JM, Lefebvre C, Sheldon TA. A statistical approach to designing search filters to find systematic reviews: objectivity enhances accuracy. *J Information Sci* 2001; **27**: 357–70.

10 Sutton AJ, Abrams KR, Jones DR, Sheldon TA, Song F. *Methods for Meta-analysis in Medical Research*. Wiley, 2000.

11 Egger M, Davey Smith G, Schneider M. Systematic reviews of observational studies. In *Systematic Reviews in Health Care: Meta-analysis in Context* (eds M Egger, G Davey Smith, D Altman): 211–27. BMJ Books, 2001.

12 Saha S, Chant DC, Welham JL, McGrath JJ. The incidence and prevalence of schizophrenia varies with latitude. *Acta Psychiatr Scand* 2006; **114**: 36–9.

13 Saha S, Welham J, Chant D, McGrath J. Incidence of schizophrenia does not vary with economic status of the country: evidence from a systematic review. *Soc Psychiatry Psychiatr Epidemiol* 2006; **41**: 338–40.

14 Singer GHS. Meta-analysis of comparative studies of depression in mothers of children with and without developmental disabilities. *Am J Ment Retard* 2006; **111**: 155–69.

15 Saha S, Chant D, McGrath J. Meta-analyses of the incidence and prevalence of schizophrenia: conceptual and methodological issues. *Int J Methods Psychiatr Res* 2008; **17**: 55–61.

16 McGrath J, Saha S, Welham J, El Saadi O, MacCauley C, Chant D. A systematic review of the incidence of schizophrenia: the distribution of rates and the influence of sex, urbanicity, migrant status and methodology. *BMC Med* 2004; **2**: 13.

17 Saha S, Chant D, Welham J, McGrath J. A systematic review of the prevalence of schizophrenia [see comment]. *PLoS Med* 2005; **2**: e141.

18 McGrath J, Saha S, Chant D, Welham J. Schizophrenia: a concise overview of incidence, prevalence, and mortality. *Epidemiol Rev* 2008; **30**: 67–76.

19 Waraich P, Goldner EM, Somers JM, Hsu L. Prevalence and incidence studies of mood disorders: a systematic review of the literature [see comment]. *Can J Psychiatry* 2004; **49**: 124–38.

20 Goldner EM, Hsu L, Waraich P, Somers JM. Prevalence and incidence studies of schizophrenic disorders: a systematic review of the literature [see comment]. *Can J Psychiatry* 2002; **47**: 833–43.

21 Somers JM, Goldner EM, Waraich P, Hsu L. Prevalence and incidence studies of anxiety disorders: a systematic review of the literature. *Can J Psychiatry* 2006; **51**: 100–13.

22 Roy MA, Maziade M, Labbe A, Merette C. Male gender is associated with deficit schizophrenia: a meta-analysis. *Schizophr Res* 2001; **47**: 141–7.

23 Scott J, McNeill Y, Cavanagh J, Cannon M, Murray R. Exposure to obstetric complications and subsequent development of bipolar disorder. Systematic review. *Br J Psychiatry* 2006; **189**: 3–11.

24 Cantor-Graae E, Selten JP. Schizophrenia and migration: a meta-analysis and review [see comment]. *Am J Psychiatry* 2005; **162**: 12–24.

25 Geddes JR, Verdoux H, Takei N, Lawrie SM, Bovet P, Eagles JM, et al. Schizophrenia and complications of pregnancy and labor: an individual patient data meta-analysis. *Schizophr Bull* 1999; **25**: 413–23.

26 Mojtabai R. Duration of illness and structure of symptoms in schizophrenia [see comment]. *Psychol Med* 1999; **29**: 915–24.

27 McLeod BD, Weisz JR, Wood JJ. Examining the association between parenting and childhood depression: a meta-analysis. *Clin Psychol Rev* 2007; **27**: 986–1003.

28 Cole MG, Dendukuri N. Risk factors for depression among elderly community subjects: a systematic review and meta-analysis. *Am J Psychiatry* 2003; **160**: 1147–56.

29 Robertson E, Grace S, Wallington T, Stewart DE. Antenatal risk factors for postpartum depression: a synthesis of recent literature [see comment]. *Gen Hosp Psychiatry* 2004; **26**: 289–95.

30 Wohl M, Gorwood P. Paternal ages below or above 35 years old are associated with a different risk of schizophrenia in the offspring. *Eur Psychiatry* 2007; **22**: 22–6.

31 Cole MG, Bellavance F, Mansour A. Prognosis of depression in elderly community and primary care populations: a systematic review and meta-analysis. *Am J Psychiatry* 1999; **156**: 1182–9.

32 Dickens C, McGowan L, Clark CD, Creed F. Depression in rheumatoid arthritis: a systematic review of the literature with meta-analysis. *Psychosom Med* 2002; **64**: 52–60.

33 King M, Semlyen J, Tai SS, Killaspy H, Osborn D, Popelyuk D, et al. A systematic review of mental disorder, suicide, and deliberate self harm in lesbian, gay and bisexual people. *BMC Psychiatry* 2008; **8**: 70.

34 Ali S, Stone MA, Peters JL, Davies MJ, Khunti K. The prevalence of co-morbid depression in adults with Type 2 diabetes: a systematic review and meta-analysis. *Diabet Med* 2006; **23**: 1165–73.

35 Gilbody S, Lightfoot T, Sheldon T. Is low folate a risk factor for depression? A meta-analysis and exploration of heterogeneity. *J Epidemiol Commun Health* 2007; **61**: 631–7.

36 Sutton A, Duval S, Tweedie R, Abrams K, Jones D. Empirical assessment of effect of publication bias on meta-analyses. *BMJ* 2000; **320**: 1574–7.

37 Costello EJ, Erkanli A, Angold A. Is there an epidemic of child or adolescent depression? *J Child Psychol Psychiatry* 2006; **47**: 1263–71.

38 Gaynes BN, Gavin N, Meltzer-Brody S, Lohr KN, Swinson T, Gartlehner G, et al. Perinatal depression: prevalence, screening accuracy, and screening outcomes. *Evid Rep Technol Assess (Summ)* 2005; **119**: 1–8.

39 Fazel S, Danesh J. Serious mental disorder in 23000 prisoners: a systematic review of 62 surveys [see comment]. *Lancet* 2002; **359**: 545–50.

40 Gavin NI, Gaynes BN, Lohr KN, Meltzer-Brody S, Gartlehner G, Swinson T. Perinatal depression: a systematic review of prevalence and incidence. *Obstet Gynec* 2005; **106**: 1071–83.

41 Abrams RC, Horowitz SV, Zweig RA. Personality disorders after age 50: a meta-analytic review of the literature. In *Personality Disorders in Older Adults: Emerging Issues in Diagnosis and Treatment* (eds E Rosowsky, RC Abrams, RA Zweig): 55–68. Erlbaum, 1999.

42 DiMaggio C, Galea S. The behavioral consequences of terrorism: a meta-analysis. *Acad Emerg Med* 2006; **13**: 559–66.

43 Skeem JL, Edens JF, Camp J, Colwell LH. Are there ethnic differences in levels of psychopathy? A meta-analysis. *Law Human Behav* 2004; **28**: 505–27.

44 Light RJ, Pillemar DB. *Summing Up: The Science of Reviewing Research*. Harvard University Press, 1984.

45 Mendelson T, Rehkopf DH, Kubzansky LD. Depression among Latinos in the United States: a meta-analytic review. *J Consult Clin Psychol* 2008; **76**: 355–66.

46 Rosenthal R, Rubin DB. Comparing significance levels of independent studies. *Psychol Bull* 1979; **86**: 1165–8.

47 Maag JW, Reid R. Depression among students with learning disabilities: assessing the risk. *J Learn Disabil* 2006; **39**: 3–10.

48 Bennett HA, Einarson A, Taddio A, Koren G, Einarson TR. Prevalence of depression during pregnancy: systematic review. *Obstet Gynec* 2004; **103**: 698–709 [erratum 1344].

49 Chida Y, Hamer M, Steptoe A. A bidirectional relationship between psychosocial factors and atopic disorders: a systematic review and meta-analysis. *Psychosom Med* 2008; **70**: 102–16.

50 Dragovic M, Hammond G. Handedness in schizophrenia: a quantitative review of evidence. *Acta Psychiatr Scand* 2005; **111**: 410–9.

51 Lorant V, Deliege D, Eaton W, Robert A, Philippot P, Ansseau M. Socioeconomic inequalities in depression: a meta-analysis [see comment]. *Am J Epidemiol* 2003; **157**: 98–112.

52 Aleman A, Kahn RS, Selten JP. Sex differences in the risk of schizophrenia: evidence from meta-analysis. *Arch Gen Psychiatry* 2003; **60**: 565–71.

53 Pinquart M, Sorensen S. Differences between caregivers and noncaregivers in psychological health and physical health: a meta-analysis [see comment]. *Psychol Aging* 2003; **18**: 250–67.

54 Kim D. Blues from the neighborhood? Neighborhood characteristics and depression. *Epidemiol Rev* 2008; **30**: 101–17.

55 Fazel S, Doll H, Langstrom N. Mental disorders among adolescents in juvenile detention and correctional facilities: a systematic review and metaregression analysis of 25 surveys. *J Am Acad Child Adolesc Psychiatry* 2008; **47**: 1010–9.

56 Breh DC, Seidler GH. Is peritraumatic dissociation a risk factor for PTSD? *J Trauma Dissoc* 2007; **8**: 53–69.

57 Starr LR, Davila J. Excessive reassurance seeking, depression, and interpersonal rejection: a meta-analytic review. *J Abnorm Psychol* 2008; **117**: 762–75.

58 Swinnen SGHA, Selten J-P. Mood disorders and migration: meta-analysis. *Br J Psychiatry* 2007; **190**: 6–10.

https://doi.org/10.1192/bjp.bp.111.098103 Published online by Cambridge University Press

59 Orth U, Wieland E. Anger, hostility, and posttraumatic stress disorder in trauma-exposed adults: a meta-analysis. *J Consult Clin Psychol* 2006; **74**: 698–706.

60 Woodberry KA, Giuliano AJ, Seidman LJ. Premorbid IQ in schizophrenia: a meta-analytic review. *Am J Psychiatry* 2008; **165**: 579–87.

61 Porter M, Haslam N. Predisplacement and postdisplacement factors associated with mental health of refugees and internally displaced persons: a meta-analysis. *JAMA* 2005; **294**: 602–12.

62 Folsom D, Jeste DV. Schizophrenia in homeless persons: a systematic review of the literature. *Acta Psychiatr Scand* 2002; **105**: 404–13.

63 Kleintjes S, Flisher AJ, Fick M, Railoun A, Lund C, Molteno C, et al. The prevalence of mental disorders among children, adolescents and adults in the Western Cape, South Africa. *S Afr Psychiatry Rev* 2006; **9**: 157–60.

64 Friedl MC, Draijer N, de Jonge P. Prevalence of dissociative disorders in psychiatric in-patients: the impact of study characteristics. *Acta Psychiatr Scand* 2000; **102**: 423–8.

65 Myers JE, Thompson ML. Meta-analysis and occupational epidemiology. *Occup Med (Lond)* 1998; **48**: 99–101.

66 Fryers T, Brugha T, Morgan Z, Smith J, Hill T, Carta M, et al. Prevalence of psychiatric disorder in Europe: the potential and reality of meta-analysis. *Soc Psychiatry Psychiatr Epidemiol* 2004; **39**: 899–905 (erratum *Soc Psychiatry Psychiatr Epidemiol* 2005; **40**: 85).

67 Sutton AJ, Abrams KR, Jones DR, Sheldon TA, Song F. Systematic reviews of trials and other studies. *Health Technol Assess* 1998; **2**: 1–276.

68 Egger M, Smith GD, Altman DG. *Systematic Reviews in Health Care: Meta-analysis in Context*. BMJ Books, 2001.

69 Alonso J, Angermeyer MC, Bernert S, Bruffaerts R, Brugha TS, Bryson H, et al. Sampling and methods of the European Study of the Epidemiology of Mental Disorders (ESEMeD) project. *Acta Psychiatr Scand Suppl* 2004; **420**: 8–20.

70 Demyttenaere K, Bruffaerts R, Posada-Villa J, Gasquet I, Kovess V, Lepine JP, et al. Prevalence, severity, and unmet need for treatment of mental disorders in the World Health Organization World Mental Health Surveys. *JAMA* 2004; **291**: 2581–90.

71 Centre for Reviews and Dissemination. *Strategies for Searching for Observational Studies*. University of York CRD, 2006.

72 Cochrane Collaboration. *The Cochrane Policy Manual*, issue 1, 2007 (http://cochrane.org/admin/manual.htm).

73 Moher CD, Eastwood S, Olkin I, Rennie D, Stroup DF. Improving the quality of reports of meta-analyses of randomised controlled trials: the QUOROM statement. Quality of Reporting of Meta-analyses. *Lancet* 1999; **354**: 1896–900.

74 Boaz A, Ashby D, Denyer D, Egan M, Harden A, Jones DR, et al. A multitude of syntheses: a comparison of five approaches from diverse policy fields. *Evid Policy* 2006; **2**: 479–52.

EXTRA CONTENT ONLINE

psychiatry in music

Psychiatry rocks

James McDonald

Modern popular music abounds with references to madness, insanity and mania, most commonly as an analogy for a romantically or chemically altered state of mind. The genres of heavy metal and punk, however, have taken particular glee in juxtaposing lyrics inspired by the more shocking psychiatric associations such as asylums, padded cells, electroconvulsive therapy and psychosurgery with loud, fast aggressive music. Songs such as Black Sabbath's 'Paranoid' and Black Flag's 'Depression' are landmarks of their respective genres. No band is as rich in references to psychopathology, however, as original 1970s New York punks The Ramones. Songs such as 'Psychotherapy', 'Gimme Gimme Shock Treatment', 'I Wanna Be Sedated' and 'Teenage Lobotomy', although playing fast and loose with DSM criteria, are punk rock classics, mini case-vignettes with a savage, knuckle-headed wit lying behind the buzz saw guitar attack. This lyrical preoccupation was no coincidence. Lead singer Joey Ramone (real name Jeffry Hyman) had Marfan's syndrome. As a teenager, he experienced a psychotic episode and was treated in hospital, his experience informing many of The Ramones' songs. Later in life Joey was diagnosed with severe obsessive–compulsive disorder (OCD), which, although not directly referenced in song, led to conflicts within the already famously dysfunctional group. Band mates were often left waiting for hours outside the singer's apartment before leaving on tour as he completed a series of rituals inside. Joey's struggle with OCD would sadly indirectly lead to his death. On returning home from a doctor's appointment, Joey was plagued by an intrusive thought that he had not closed the door at the surgery properly and set out across town to remedy this. Unfortunately, he slipped on ice on the street and fractured a hip, complicating the treatment of previously diagnosed lymphoma. He did not recover and passed away in April 2001, aged 49. Two other original members of the band, Johnny and Dee Dee, died soon afterwards. The Ramones remain one of the most influential and best loved rock'n'roll bands of all time, their enduring appeal to those disenfranchised by mainstream society best summarised by the intro to the band's signature anthem 'Pinhead': 'Gabba, Gabba, we accept you, one of us'.

The British Journal of Psychiatry (2012) 200, 453. doi: 10.1192/bjp.bp.111.105981

https://doi.org/10.1192/bjp.bp.111.098103 Published online by Cambridge University Press