Eur J Oral Sci 2003; 111: 85–92
Printed in UK. All rights reserved

European Journal of
Oral Sciences
ISSN 0909-8836

The assessment of systematic reviews in dentistry

A. M. Glenny[1], M. Esposito[2], P. Coulthard[3], H. V. Worthington[1]

[1]Cochrane Oral Health Group, University Dental Hospital of Manchester, UK; [2]Department of Biomaterials and Department of Prosthetic Dentistry/Dental Material Sciences, The Sahlgrenska Academy at Göteborg University, Göteborg, Sweden; [3]Oral and Maxillofacial Surgery, University Dental Hospital of Manchester, Manchester, UK

Glenny AM, Esposito M, Coulthard P, Worthington HV. The assessment of systematic reviews in dentistry. Eur J Oral Sci 2003; 111: 85–92. © Eur J Oral Sci, 2003

The aim of this study was to assess the quality of systematic reviews of effectiveness of interventions in dentistry. The Database of Abstracts of Reviews of Effectiveness and the Cochrane Database of Systematic Reviews were searched to identify systematic reviews examining the effectiveness of interventions for oral, dental and craniofacial disorders and diseases. Sixty-five reviews were identified and assessed independently by two reviewers. The area most frequently evaluated within the reviews was pain relief/prevention (20/65, 31%) followed by caries and oral medicine. The quality assessment of the identified systematic reviews highlighted key areas where improvements could be made. One major weakness of the reviews was that the search strategies employed were not always adequate. Only 12 reviews (19%) demonstrated an attempt to identify all relevant studies. Other areas of weakness include the screening and quality assessment of primary studies, the pooling of data and examination of heterogeneity, and the interpretation of findings. This investigation shows that the quality of systematic reviews in dentistry could be improved. If future clinical decisions are to be based upon systematic reviews, it is imperative that the reviews address clinically relevant, focused questions and follow a 'transparent', well-designed protocol.

Anne-Marie Glenny, Cochrane Oral Health Group, MANDEC University Dental Hospital of Manchester, Higher Cambridge Street, Manchester M15 6FH, UK

Telefax: +44–161–2757815
E-mail: a.glenny@man.ac.uk

Key words: systematic review; dentistry; evidence-based practice

Accepted for publication November 2002

As the importance of evidence-based practice continues to be promoted, so the profile of the systematic review grows. In the late 1980s MULROW (1) highlighted that many traditional, medical reviews were haphazard and biased, often reflecting the opinion of the review's authors. In contrast, systematic reviews follow explicit, well-documented, scientific methodology in order to reduce both systematic errors (biases) and random errors (those occurring by chance) and provide a more objective, comprehensive view of the research literature.

The rationale for systematic reviews has been well documented (2, 3). In addition to the reduction in bias, one of the many advantages of systematic reviews is that they enable us to reduce the increasing torrent of both published and unpublished research literature into manageable portions. Within dentistry alone there are around 500 journals publishing over 43 000 research articles a year. It is not feasible to expect anyone to keep abreast of the emerging research evidence, identifying those articles that are both of a high quality and relevant to their own clinical practice.

In recent years, the interest in systematic reviews, their production, and their publication, has been growing. Yet, despite the well-documented advantages of this scientific technique, some are still doubtful about the usefulness of such reviews. These doubts are often based upon misconceptions regarding the 'history, purpose, methods and uses of systematic reviews' (4).

In order to promote further the role of the systematic review and dispel criticism aimed at them, it is imperative that those undertaking systematic reviews ensure that the reviews are of high quality and follow a rigorous, 'transparent' protocol. In order to assess whether a systematic review has followed such a protocol, the reporting of the review has to be sufficiently detailed. Guidelines for the reporting of systematic reviews of both randomized controlled trials and observational studies have been proposed (5,6).

The Cochrane Collaboration (www.cochrane.org) is an international organization that aims to prepare and maintain rigorous systematic reviews in order to help people make well-informed decisions about health care. The Collaboration has over 50 review groups focusing on different aspects of health care. The reviews produced by each group undergo rigorous peer review, with the protocols of the reviews published in The Cochrane Library, and are open to all for further comments and criticism. The Cochrane Library is a useful resource for all those interested in evidence-based practice. Not only does it include the systematic reviews produced by the Cochrane review groups, it also contains several other databases including the Cochrane Controlled Trials Register (CCTR – a bibliography of controlled trials identified

through electronic and handsearching), the NHS Economic Evaluation Database (NHS EED) and the Database of Abstracts Reviews of Effectiveness (DARE). DARE (also available through the NHS Centre for Reviews and Dissemination homepage at www.york.ac.uk/Inst/crd) contains structured abstracts assessing and summarizing published systematic reviews that have met predefined quality criteria.

Given the continued discussions around evidence-based practice and systematic reviews, the aim of this study was to assess the quality of systematic reviews of effectiveness conducted in the field of dentistry.

Methods

Search strategy

The Database of Abstracts of Reviews of Effectiveness provides a good source of systematic reviews, compiled through rigorous, monthly searches of bibliographic databases, hand-searching of key major medical journals, and by scanning grey literature electronic databases (covering publications such as technical report and discussion papers). DARE was established in 1994 although the database does hold records of some systematic reviews published prior to this date (identified on an *ad hoc* basis). The databases currently searched are:

- Current Contents Clinical Medicine
- Medline
- CINAHL
- ERIC
- Biosis
- Allied and Alternative Medicine
- PsycINFO

In order to identify reviews of interventions in the dental field, the Cochrane Oral Health Group's (COHG) search strategy was applied to DARE (1994 to January 2001) accessed through The Cochrane Library. In addition, all completed systematic reviews registered by the COHG on the Cochrane Database of Systematic Reviews were retrieved.

Inclusion criteria

All systematic reviews examining the effectiveness of interventions for oral, dental and craniofacial disorders and diseases were included in the review.

It was recognized that some reviews do not focus entirely on dental studies. For example, systematic reviews examining the effectiveness of analgesics for postoperative pain relief often include trials of patients that have undergone a variety of surgical procedures. Where such reviews included studies conducted in a dental setting alongside studies conducted in medicine, then the reviews were included in the quality assessment.

Assessment of relevance

The results of the search strategy were divided between the four reviewers. Each title was assessed independently by two reviewers (one clinician and one methodologist) and the full paper article retrieved for those meeting the inclusion criteria. Any discrepancies were resolved through discussion, with the full article being retrieved for clarification, if necessary.

Assessment of validity

A quality assessment checklist was devised and piloted by all four reviewers on a sample of 10 reviews. The items examined on the checklist included are listed in Table 1. In addition to the quality items, data on the area of dentistry examined (e.g. orthodontics, restorative dentistry) (Table 2) and the type of intervention were recorded. The validity assessment form was piloted on 10 reviews by all four reviewers. Areas of ambiguity were discussed and the assessment form revised.

Each systematic review meeting the inclusion criteria was then assessed independently by two reviewers (one clinician and one methodologist) and discrepancies resolved through discussion.

Data analysis

The results of the validity assessment were managed in SPSS software and the number of reviews meeting each quality criteria recorded. Overall quality scores for each review were not produced as such summary scores have been shown to be problematic (7).

Results

A total of 115 titles were identified through the searching of DARE (16 January 2001). After initial assessment of the titles identified through DARE, 65 articles (8–72) were deemed potentially relevant for the review and the full articles were sought. Four of these articles were subsequently excluded (8–11). Reasons for exclusion were: the articles were summaries or abstracts of systematic reviews that did not report the review methodology in sufficient detail for assessment (8–10); the paper focused purely on the validity assessment of papers and did not present the whole systematic review (11). In addition, the full version of one article was unattainable at the time of assessment (12).

In addition to the 60 full reviews identified for inclusion through DARE, a further five reviews were retrieved from the Cochrane Database of Systematic reviews (73–77), thus a total of 65 reviews were included for assessment (13–77). The number of reviews published in each year since 1990 can be seen in Fig. 1.

The piloting of the validity assessment form led to alterations regarding the search strategy used. Initially, the validity assessment form asked 'Do you think all important relevant studies were included in the review?'. It was felt that the answering of such a question would require a sound knowledge of the subject covered in the review. Given the wide spread of topics covered by the systematic reviews, this question was consistently being answered with 'Can't tell'. The question was therefore reworded to 'Do you think the authors attempted to identify all relevant studies?', which did not require such a detailed knowledge of the subject matter, but considered the comprehensiveness of the search.

Table 1

Per cent agreement range of κ *values with asymptotic standard errors (not assuming the null hypothesis is true), assessing the inter-examiner agreement between clinicians and methodologists*

Question *(possible categories)*	Percent agreement	κ values (asymp. SE)
A. Did review address a focused question? (yes, no, can't tell)	82	0.06 (0.13)
B. Did authors look for appropriate papers? (yes, no, can't tell)	72	0.38 (0.11)
C. Do you think authors attempted to identify all relevant studies? (yes, no, can't tell)	69	0.47 (0.09)
D. Search for published and unpublished literature (yes, no, can't tell)	71	0.46 (0.09)
E. Were all languages considered? (yes, no, can't tell)	88	0.81 (0.06)
F. Was any hand-searching carried out? (yes, no, can't tell)	71	0.40 (0.10)
G. Was it stated that the inclusion criteria were carried out by at least two reviewers? (yes, no, can't tell)	72	0.52 (0.09)
H. Did reviewers attempt to assess the quality of the included studies? (yes, no)	88	0.74 (0.09)
I. If so did they include this in the analysis? (yes, no, can't tell, not applicable)	57	0.36 (0.09)
J. Was it stated that the quality assessment was carried out by at least two reviewers? (yes, no, not applicable)	62	0.43 (0.08)
K. Are the results given in a narrative or pooled statistical analysis? (narrative, pooled, not applicable)	88	0.73 (0.08)
L. If the results have been combined was it reasonable to do so? (yes, no, can't tell, not applicable)	57	0.35 (0.09)
M. Are the results clearly displayed? (yes, no, not applicable)	80	0.60 (0.10)
N. Was an assessment of heterogeneity made and reasons for variation discussed? (yes, no, not applicable)	74	0.57 (0.09)
O. Were results of review interpreted appropriately? (yes, no, can't tell, not applicable)	55	0.25 (0.08)

$n = 65$ for all assessments.

Table 2

Areas of dentistry addressed in the systematic reviews

Area of dentistry	Frequency	(%)
Caries	7	11
Fluorosis	1	2
Oral maxillofacial surgery	5	8
Oral health promotion	3	5
Oral medicine	7	11
Orthodontics	3	5
Pain	20	31
Periodontics	6	9
Restorative dentistry	4	6
Sleep apnoea	2	3
Temporomandibular disorders	5	8
Other	2	3
Total	65	

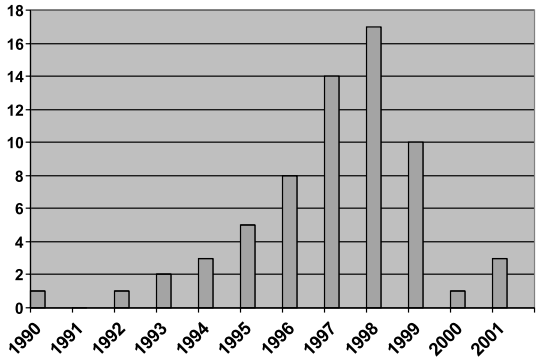

Fig. 1. Number of systematic reviews published since 1990 by year.

In addition, the initial form required a decision to be made on whether or not the review had searched for non-English language articles. There was much variation in the number of non-English languages considered within the reviews therefore the question was re-worded to 'Were all languages considered?'.

Two questions regarding the application of the inclusion criteria and quality assessment carried out within the systematic reviews also required modification. These were initially worded as 'Were (inclusion/validity) criteria applied independently by at least two reviewers?'. The wording was changed to 'Was it stated that…' (see Table 1).

Inter-rater agreement

There were four raters, two clinicians and two methodologists. At least one clinician and one methodologist

assessed each review. When studies were assessed by more than one either clinician or methodologist, the assessment used for analysis was selected at random. The per cent agreement was generally high, ranging across all assessments from 55% to 88%, with a median of 72%. Overall κ values ranged from 0.06 to 0.81, with a median value of 0.46. The results of this assessment are given for each question on the assessment form in Table 1. Using the classification of kappa suggested by LANDIS and KOCH (78) one item only achieved slight agreement ('Did the review address a focused question?'), whereas one-third of items achieved substantial agreement, with the remaining items achieving fair to moderate agreement.

Areas of dentistry covered by the systematic reviews

Of the 65 systematic reviews including studies in the field of dentistry, 20 (31%) focused on pain (Table 2). Only three of these reviews focused on dental pain alone (13–15), the remainder of the reviews pooled studies set in both dental and medical settings. The other areas most frequently evaluated within the systematic reviews were periodontics, caries, oral medicine, and oral maxillofacial surgery.

Type of intervention

Forty-five of the systematic reviews focused on the treatment of a disease or disorder (69%) (Table 3). Seven of these reviews also evaluated the associated side-effects of the intervention (11%) and three provided information on both treatment and prevention (5%). Prevention alone, diagnosis and adverse events alone were covered in 11 (17%), five (8%) and four (6%) of the reviews, respectively.

Results of quality assessment

The results of the validity assessment can been seen in Table 4 (owing to restrictions in space, a table providing the results of the quality assessment of the individual reviews is available from the authors on request). The results for the quality assessment of the 20 pain reviews are also presented separately to examine whether reviews conducted in medical/dental settings differed in quality from those conducted in dental settings alone. The percentage of pain reviews meeting each of the 15 quality criteria was higher for 12 of the criteria; however, when comparing the pain reviews with the other reviews there was only a statistically significant difference for 'Did the authors look for appropriate papers' ($\chi^2 = 10.2$, 1 df, $P < 0.001$).

Table 3

Type of intervention

Intervention	Frequency	(%)
Treatment	35	54
Prevention	11	17
Diagnosis	5	8
Adverse events	4	6
Treatment and prevention	3	5
Treatment and adverse events	7	11
Total	65	

Fifty-six (86%) of all 65 reviews addressed a clearly focused question, in which it was apparent what types of participants, interventions and outcome measures were included in the review. Seventy-four per cent of the reviews looked for studies of an appropriate design to answer their questions. Randomized controlled trials (RCTs) were included in 42 of the 65 reviews (65%). The number of RCTs included in these reviews ranged from two to 90 RCTs, with a median of 14 RCTs in each review. Of these 42 reviews, 33 included only RCTs, and a further seven only RCTs and controlled clinical trials (CCTs).

For the other 23 reviews, CCTs were the highest level of evidence included in six of the reviews, and cohort studies in three reviews. Other studies used lower levels of evidence or were unclear about the study designs included. In one review it was also unclear how many studies were included. For the remaining 64 reviews, the total number of studies included in reviews ranged from three to 116, with a median of 15 studies per review.

When examining how comprehensive the search strategies were within the reviews, it was felt that only 12 reviews (19%) demonstrated that they had attempted to identify all relevant studies. All 12 of these reviews were published between 1996 and 2001.

The most frequently searched electronic database was Medline, which was searched by 63 of the systematic reviews (97%). Embase was searched in 17 reviews (26%), the CCTR in 10 reviews (15%) and BIDS searched in only four reviews (6%). Other databases were also searched in 33 of the reviews included. One review mentioned three unspecified databases.

Only eight reviews searched for published and unpublished literature and considered all languages. Of these eight reviews only five had carried out hand-searching in addition to electronic database searching.

Inclusion criteria were often clearly described. However, in only 24 of the 65 reviews (37%) was it explicitly stated that the inclusion criteria were applied independently by at least two reviewers. Validity assessment of the primary studies included in the systematic reviews was undertaken in 39 reviews (60%). Of these 39 reviews, only 24 stated that the assessment had been conducted independently by at least two reviewers, and only 25 included the results of the quality assessment exercise in the review's analysis.

The majority of the reviews conducted a meta-analysis (46/65, 71%), with 16 reviews pooling data narratively (25%). In two reviews it was unclear how the data had been pooled. Data pooling was classed as appropriate in 37 of the systematic reviews. For pooling to be classed as appropriate, the review needed to consider heterogeneity between studies and give a clear presentation of the characteristics of the primary studies included in the review.

The assessment of the interpretation of the reviews' findings is subjective, and is reflected in the wide range in

Table 4

Results of the quality assessment for the 65 reviews

	All reviews ($n = 65$)		Can't tell (%)	Not applicable (%)	Pain reviews ($n = 20$)		Can't tell (%)	Not applicable (%)
	Yes (%)	No (%)			Yes (%)	No (%)		
A. Did review address a focused question?	56 (86)	7 (11)	2 (3)		19 (95)	1 (5)		
B. Did authors look for appropriate papers?	48 (74)	9 (14)	8 (12)		20 (100)			
C. Do you think authors attempted to identify all relevant studies?	12 (19)	45 (69)	8 (12)		4 (20)	12 (60)	4 (20)	
D. Search for published and unpublished literature?	17 (26)	39 (60)	9 (14)		6 (30)	11 (55)	3 (15)	
E. Were all languages considered?	17 (26)	26 (40)	22 (34)		7 (35)	6 (30)	7 (35)	
F. Was any hand searching carried out?	17 (26)	41 (63)	7 (11)		6 (30)	10 (50)	4 (20)	
G. Was it stated that the inclusion criteria were carried out by at least two reviewers?	24 (37)	41 (63)			6 (30)	12 (60)	2 (10)	
H. Did reviewers attempt to assess the quality of the included studies?	39 (60)	26 (40)			13 (65)	7 (35)		
I. If so did they include this in the analysis?	25 (64)*	13 (33)*	1 (3)*		8 (62)¶	4 (31)¶	1 (8)¶	
J. Was it stated that the quality assessment was carried out by at least two reviewers?	24 (62)*	15 (39)*			9 (69)¶	4 (31)¶		
K. Are the results given in a narrative or pooled statistical analysis?	17 (26)a	46 (71)b		2 (3)	7 (35)a	13 (65)b		
L. If the results have been combined was it reasonable to do so?	37 (57)	10 (15)	12 (19)	6 (9)	13 (65)	1 (5)	5 (25)	1 (5)
M. Are the results of the primary studies clearly displayed?	41 (63)	17 (26)	1 (2)	6 (9)	11 (55)	8 (40)	1 (5)	
N. Was an assessment of heterogeneity made and reasons for variation discussed?	29 (45)	24 (37)		12 (19)	10 (50)	8 (40)	2 (10)	
O. Were results of review interpreted appropriately?	33 (51)	2 (3)	7 (11)	23 (35)	13 (65)		6 (30)	1 (5)

*$n = 39$.
¶$n = 13$.
a = narrative, b = statistical.

per cent agreement (Table 1). Reviews in which the data pooling was inappropriate or unclear were marked 'not applicable'. Of the remaining 42 reviews, 33 were thought to have interpreted the results appropriately. In two reviews the interpretation of the results was clearly inappropriate, and in a further seven reviews it was not possible for a judgement to be made.

Discussion

It is encouraging to see that there are already over 60 systematic reviews in the field of dentistry, covering a wide range of topics. It is interesting that the publication rate has risen well from 1990 to 1999. The fall in the number of reviews published since 1999 identified in this study is perhaps a reflection of the nature of our search rather than a true drop in the number of systematic reviews being published. The search strategy focused on systematic reviews with abstracts published on DARE. The process undertaken to identify and appraise systematic reviews before they appear on DARE is lengthy. Inevitably, there is a time delay between systematic reviews being identified, appraised and an abstract being published on DARE. The time delay is just over 2 yr, which means that systematic reviews published in 2001 and 2000, and some of those published in 1999, were still awaiting appraisal and will not have been included in our assessment. However, it was felt that the assessment undertaken reflects the quality of systematic reviews that are easily accessible to dentists, given that DARE is freely available on the Internet and does not require complex searching to identify the reviews.

The only reviews included in our assessment that have been published since 1999 are those registered with the Cochrane Oral Health Group (COHG). At present, the COHG has 16 completed reviews (11 published following our initial assessment), 29 published protocols and a further 30 registered titles/protocols. Considering Cochrane reviews alone, it is envisaged that the number of systematic reviews in the field of dentistry will continue to increase.

The quality assessment of the systematic reviews identified highlighted key areas where improvements could be made in the conduct of the reviews. One of the major weaknesses of the systematic reviews was that the search strategies reported were not always adequate. The aim of a systematic review is to provide a comprehensive, unbiased summary of current research evidence. In order to achieve this aim, a systematic review must use a transparent and inclusive search. Ideally, the search strategy for a systematic review should aim to identify all published and unpublished data, irrespective of language. There is good evidence that research findings showing statistically significant results are more likely to be submitted and accepted for publication (79,80), and more likely to be published in English language journals (81). By excluding unpublished and non-English lan-

guage articles, the results of a systematic review are susceptible to such publication biases. It was disappointing to find that only eight (12%) of the systematic reviews included searched for both published and unpublished data with no language restriction. Similarly, only 17 (26%) of the reviews carried out hand-searching. Incomplete indexing and the lack of 'key words' in the titles and abstracts can make published articles difficult to identify through the searching of electronic databases such as Medline. A recent study comparing the results of a Medline search and hand-search of five major orthodontic journals (1991–2000) showed that hand-searching increased the yield of controlled clinical trials by 145% (HARRISON JE & BICKLEY SR, 2002, PEF-IADR, Cardiff, UK, unpublished data). It is reassuring to note that the reviews included in the assessment felt to have presented a comprehensive search strategy (12/65, 19%) were those published most recently (post-1996).

When devising search strategies, future systematic reviews should consider a combination of the following: structured searches of electronic databases (including The Cochrane Library) using a combination of free text and controlled vocabulary (e.g. MeSH terms); hand-searching of relevant journals or conference abstracts; screening of reference lists of all studies included and relevant review articles; contacting appropriate manufacturers/pharmaceutical companies and experts in the field.

In order to avoid selective inclusion of studies that support the reviewer's opinion, screening of search results to identify potentially relevant articles should, ideally, be undertaken independently by at least two reviewers. This process was explicitly stated in only 24 reviews (37%).

Only 39 (60%) of the reviews carried out a quality assessment of the primary studies included. Of these 39 reviews, 24 (62%) stated that the process was conducted independently by at least two reviewers, thus limiting bias. It is disappointing that so many reviews failed to undertake some form of quality assessment, as it has been shown that low quality primary studies can distort the findings of a systematic review. In addition, some of the reviews undertaking validity assessment failed to incorporate the findings of the assessment when analysing the results of the review. In order to establish the robustness of a review's findings, it is important that the validity of the studies included is assessed and a sensitivity analysis undertaken to evaluate the effect of including/excluding low quality trials.

The pooling of data within systematic reviews is not always appropriate. Systematic reviews can, but do not have to, use meta-analysis when combining data from primary studies. Indeed, it is not always appropriate to use meta-analysis, particularly when data are sparse or when the primary studies are heterogeneous (either clinically or statistically). Prior to conducting a meta-analysis, a reviewer should examine the consistency of the treatment effect across the primary studies. A statistical test of heterogeneity should be undertaken. If heterogeneity is considerable, consideration ought to be given to why the differences exist between the studies (examining the individual trial characteristics) and whether or not it is appropriate to pool the data statistically. If a meta-analysis is undertaken where heterogeneity exists, a random effects model should be used as this will incorporate an estimate of between study variation into the calculation of the common effect (82). Over 70% of the systematic reviews examined did conduct a meta-analysis. When deciding whether the data had been pooled appropriately, the clear presentation of the characteristics of the primary studies, and a test of heterogeneity were considered. The pooling of data was deemed appropriate in 37 (57%) of the reviews.

One of the strengths of a systematic review process is objectivity. However, when it comes to the interpretation of the review's findings, objective conclusions can often be difficult to make, and reviewers may, consciously or not, interpret the results according to their own preconceptions. Indeed, even if the reviewer has no preconceptions, ambiguous results can still be interpreted in many different ways. A study undertaken on the interpretation of the results of a systematic review of fat-modifying diets highlights this issue (HOOPER C, HIGGINS J, GLENNY A, 2000, 3rd Symposium of systematic reviews: beyond the basics, Oxford, UK). Several of the results from the review hovered around the significant/non-significant border. Numerical data from the review were disguised by removing details of the interventions and outcomes, and circulated on two e-mail discussion lists (the Cochrane Collaboration list, CCINFO, and Mailbase list, Evidence Based Health). Members of the discussion lists were asked to vote on the most accurate interpretation of the data. Responses included 'no significant evidence of effect on any outcome', 'a trend towards protection', and 'significant effect on disease outcomes'. The difficulty in interpreting results strengthens the argument that a systematic review should be a transparent process, with the reader of the review being able to identify what has been done at each stage. A clear presentation of the review methodology and characteristics of the primary studies included allows the reader the opportunity to establish whether or not they feel the interpretation of the results is fair.

We have to be aware that the quality assessment undertaken in this paper assesses only what has been reported in the reviews and may not always reflect how the review was actually conducted. Both the QUOROM statement (5) and MOOSE (6) provide guidelines on how to improve the reporting of systematic reviews/meta-analyses. The former focuses on meta-analyses of RCTs, while the latter discusses the reporting of meta-analyses of observational studies in epidemiology. Items covered in these checklists include those covered in Table 1. However, both checklists are more detailed and, between them, cover additional items such as the provision of a structured abstract, a description/justification of excluded studies, the qualification of the searcher, rationale for any a priori sensitivity or subgroup analyses, quantitative assessment of bias (for example, publication bias), trial flow, and disclosure of funding. It is envisaged that such guidelines will improve not only the reporting of systematic reviews/meta-analyses, but also, by

increasing the awareness of what is required in a good-quality review, the actual conduct of the review itself. It should be noted, however, that many of the criteria listed in these checklists, although now recognized as being important in the conduct of a systematic review, may not have been so when some of the earlier reviews included in this assessment were undertaken.

Systematic reviews have an important role to play in evidence-based dentistry. However, their usefulness is totally dependent upon their validity. This investigation shows that the quality of systematic reviews in dentistry could be improved. If future clinical decisions are to be based upon systematic reviews, it is imperative that the reviews address clinically relevant, focused questions and follow a 'transparent', well-designed protocol.

References

1. Mulrow C. The medical review article: State of the science. *Ann Intern Med* 1987; **106**: 485–488.
2. Mulrow C. Rationale for systematic reviews. *BMJ* 1994; **309**: 597–599.
3. Chalmers I, Altman D. *Systematic Reviews*. London: BMJ Publishing Group 1995.
4. Petticrew M. Systematic reviews from astronomy to zoology: Myths and misconceptions. *BMJ* 2001; **322**: 98–101.
5. Moher D, Cook D, Eastwood S, Olkin I, Rennie D, Stroup D, for the Quorom group. Improving the quality of reports of meta-analyses of randomised controlled trials: The QUOROM statement. *Lancet* 1999; **354**: 1896–1899.
6. Stroup D, Berlin J, Morton S, Olkin I, Williamson G, Rennie D, Moher D, Becker B, Sipe T, Thacker S, for the Meta-analysis of Observational Studies in Epidemiology (MOOSE) Group. Meta-analysis of observational studies in epidemiology. A proposal for reporting. *JAMA* 2000; **283**: 2008–2012.
7. Juni P, Witschi A, Bloch R, Egger M. The hazards of scoring the quality of clinical trials for meta-analysis. *JAMA* 1999; **282**: 1054–1060.
8. NHS Centre For Reviews and Dissemination. Dental restorations. What type of filling? *Effective Health Care Bull* 1999; **5**: 1–12.
9. Chadwick B, Dummer P, Dunstan F, Gilmour A, Jones R. What type of filling? Best practice in dental restorations. *Qual Health Care* 1999; **8**: 202–207.
10. Hickner J. Antibiotics for acute maxillary sinusitis in adults. *J Fam Pract* 1998; **46**: 281–282.
11. Rosted P. The use of acupuncture in dentistry: a review of the scientific validity of published papers. *Oral Dis* 1998; **4**: 100–104.
12. Flynn K. *Positron emission tomography: Systematic review: PET as a diagnostic test in head and neck cancer. Technology Assessment Program PET Rep.;* Boston, USA: Veterans Affairs Medical Center, Health Services Research and Development Service, Management Decision and Research Center, 1996.
13. Ahmad N, Grad H, Haas D, Aronson K, Jokovic A, Locker D. The efficacy of nonopioid analgesics for postoperative dental pain: a meta-analysis. *Anesth Prog* 1997; **44**: 119–126.
14. Ernst E, Pittler M. The effectiveness of acupuncture in treating acute dental pain: a systematic review. *Br Dent J* 1998; **184**: 443–447.
15. Rosted P. The use of acupuncture in dentistry: a systematic review. *Acupunct Med* 1998; **16**: 43–48.
16. Wilkie D, Savedra M, Holzemer W, Tesler M, Paul S. Use of the McGill pain questionnaire to measure pain: a meta-analysis. *Nurs Res* 1990; **39**: 36–41.
17. Hayes C, Antczak-Bouckoms A, Burdick E. Quality assessment and meta-analysis of systemic tetracycline use in chronic adult periodontitis. *J Clin Periodontol* 1992; **19**: 164–168.
18. Antczak-Bouckoms A, Joshipura K, Burdick E, Tulloch J. Meta-analysis of surgical versus non-surgical methods of treatment for periodontal disease. *J Clin Periodontol* 1993; **20**: 259–268.
19. Llodra J, Bravo M, Delgado-Rodriguez M, Baca P, Galvez R. Factors influencing the effectiveness of sealants – a meta-analysis. *Community Dent Oral Epidemiol* 1993; **21**: 261–268.
20. Helfenstein U, Steiner M. Fluoride varnishes (Duraphat): a meta-analysis. *Community Dent Oral Epidemiol* 1994; **22**: 1–5.
21. Meunier F, Paesmans M, Autier P. Value of antifungal prophylaxis with antifungal drugs against oropharyngeal candidiasis in cancer patients. *Oral Oncol Eur J Cancer* 1994; **30B**: 196–199.
22. Smith M, Holcombe J, Stullenbarger E. A meta-analysis of intervention effectiveness for symptom management in oncology nursing research. *Oncol Nurs Forum* 1994; **21**: 1201–1210.
23. Liedberg J, Panmekiate S, Petersson A, Rohlin M. Evidence-based evaluation of three imaging methods for the temporomandibular disc. *Dentomaxillofac Radiol* 1995; **25**: 234–241.
24. McQuay H, Carroll D, Jadad AR, Wiffen P, Moore A. Anticonvulsant drugs for management of pain: a systematic review. *BMJ* 1995; **311**: 1047–1052.
25. McQuay H. Pre-emptive analgesia: a systematic review of clinical studies. *Ann Med* 1995; **27**: 249–256.
26. Schmidt-Nowara W, Lowe A, Wiegand L, Cartwright R, Perez-Guerra F. Oral appliances for the treatment of snoring and obstructive sleep a review. *Sleep* 1995; **18**: 501–510.
27. van Rijkom H, Verdonschot E. Factors involved in validity measurements of diagnostic tests for approximal caries-a meta-analysis. *Caries Res* 1995; **29**: 364–370.
28. Bollen C, Quirynen M. Microbiological response to mechanical treatment in combination with adjunctive therapy: a review of the literature. *J Periodontol* 1996; **67**: 1143–1158.
29. Kay E, Locker D. Is dental health education effective? A systematic review of current evidence. *Community Dent Oral Epidemiol* 1996; **24**: 231–235.
30. McQuay H, Tramer M, Nye B, Carroll D, Wiffen P, Moore R. A systematic review of antidepressants in neuropathic pain. *Pain* 1996; **68**: 217–227.
31. Sher A, Schechtman K, Piccirillo J. The efficacy of surgical modifications of the upper airway in adults with obstructive sleep apnoea syndrome. *Sleep* 1996; **19**: 156–177.
32. Sprod A, Anderson R, Treasure E. *Effective oral health promotion. Literature review: Health Promotion Wales; Report No. Techn Rep 20*. Cardiff: Cardiff Dental Public Health Unit, Health Promotion Wales, 1996; 1–67.
33. van Rijkom H, Truin G, van t Hof M. A meta-analysis of clinical studies on the caries-inhibiting effect chlorhexidine treatment. *J Dent Res* 1996; **75**: 790–795.
34. Williamson I, Whelan T. The clinical problem of Bell's palsy: Is treatment with steroids effective? *Br J Gen Pract* 1996; **46**: 743–747.
35. Zhang W, Li Wan Po A. Analgesic efficacy of paracetamol and its combination with codeine and caffeine in surgical pain – a meta-analysis. *J Clin Pharm Ther* 1996; **21**: 261–282.
36. de Bock G, Dekker F, Stolk J, Springer M, Kievit J, van Houwelingen J. Antimicrobial treatment in acute maxillary sinusitis: a meta analysis. *J Clin Epidemiol* 1997; **50**: 881–890.
37. Dodson T, Kaban L. Recommendations for management of trigeminal nerve defects based on a critical appraisal of the literature. *J Oral Maxillofac Surg* 1997; **55**: 1380–1386.
38. Elter J, Offenbacher S, Beck J. Meta-analysis of the effect of systemic metronidazole as an adjunct to scaling and root planing for adult periodontitis. *J Periodontal Res* 1997; **32**: 487–496.
39. Jackson J, Gibbons R, Meyer G, Inouye L. The effect of treating herpes zoster with oral acyclovir in preventing postherpetic neauralgia: a meta-analysis. *Arch Int Med* 1997; **157**: 909–912.
40. Kassir R, Rassekh C, Kinsella J, Segas J, Carrau R, Hokanson J. Osteosarcoma of the head and neck: meta-analysis of nonrandomized studies. *Laryngoscope* 1997; **107**: 56–61.
41. Kay E, Locker D. Effectiveness of oral health promotion: a review. *Health Promot Effectiveness Rev* 1997; **7**: 1–88.
42. Linde K, Clausius N, Ramirez G, Melchart D, Eitel F, Hedges L, Jonas W. Are the clinical effects of homeopathy

placebo effects? A meta-analysis of placebo-controlled trials. *Lancet* 1997; **350**: 834–843.
43. Moore A, Collins S, Carroll D, McQuay H. Paracetamol with and without codeine in acute pain: a quantitative systematic review. *Pain* 1997; **70**: 193–201.
44. Moore R, McQuay H. Single-patient data meta-analysis of 3453 postoperative patients: oral tramadol versus placebo, codeine and combination analgesics. *Pain* 1997; **69**: 287–294.
45. Picard P, Tramer M, McQuay H, Moore R. Analgesic efficacy of peripheral opioids (all except intra-articular): a qualitative systematic review of randomised controlled trials. *Pain* 1997; **72**: 309–318.
46. Smeele L, Kostense P, van der Waal I, Snow G. Effect of chemotherapy on survival of craniofacial osteosarcoma; a systematic review of 201 patients. *J Clin Oncol* 1997; **15**: 363–367.
47. Trpkova B, Major P, Prasad N, Nebbe B. Cephalometric landmarks identification and reproducibility: a meta analysis. *Am J Orthod Dentofacial Orthop* 1997; **112**: 165–170.
48. Zhang W, Li Wan Po A. Do codeine and caffeine enhance the analgesic effect of aspirin? A systematic overview. *J Clin Pharm Ther* 1997; **22**: 79–97.
49. Carroll D, Seers K. Relaxation for the relief of chronic pain: a systematic review. *J Adv Nurs* 1998; **27**: 476–487.
50. Dolovich L, Addis A, Vaillancourt J, Power J, Koren G, Einarson T. Benzodiazepine use in pregnancy and major malformations or oral cleft: meta-analysis of cohort and case-control studies. *BMJ* 1998; **317**: 839–843.
51. Fishbain D, Cutler R, Rosomoff H. Do antidepressants have an analgesic effect in psychogenic pain and somatoform pain disorder: a meta analysis. *Psychosom Med* 1998; **60**: 503–509.
52. Kowanko I, Long L, Hodgkinson B, Evans D. The effectiveness of strategies for preventing and treating chemotherapy and radiation induced oral mucositis in patients with cancer. *Joanna Briggs Inst Evidence Based Nurs Midwifery* 1998; **2**: 1–84.
53. Kreulen C, Creugers N, Meijering A. Meta-analysis of anterior veneer restorations in clinical studies. *J Dent* 1998; **26**: 345–353.
54. Li Wan Po A, Zhang W. Analgesic efficacy of ibuprofen alone and in combination with caffeine in post-surgical pain: a meta-analysis. *Eur J Clin Pharmacol* 1998; **53**: 303–311.
55. Linde K, Melchart D. Randomized controlled trials of individualized homeopathy: a state-of-the-art review. *J Altern Complement Med* 1998; **4**: 371–388.
56. Lindh T, Gunne J, Tillberg A, Molin M. A meta-analysis of implants in partial edentulism. *Clin Oral Implants Res* 1998; **9**: 80–90.
57. Locker D. Patient-based assessment of the outcomes of implant therapy: a review of the literature. *Int J Prosthodont* 1998; **11**: 453–461.
58. Scurria M, Bader J, Shugars D. Meta-analysis of fixed partial denture survival: Prostheses and abutments. *J Prosthet Dent* 1998; **79**: 459–464.
59. Sellick S, Zaza C. Critical review of five nonpharmacological strategies for managing cancer pain. *Cancer Prevent Control* 1998; **2**: 7–14.
60. Tong D, Rioux K, Drangsholt M, Beirne O. A review of survival rates for implants placed in grafted sinuses using meta-analysis. *Int J Oral Maxillofac Implants* 1998; **13**: 175–182.
61. Tramer M, Williams J, Carroll D, Wiffen P, Moore R, McQuay H. Comparing analgesic efficacy of non steroidal anti inflammatory drugs given by different routes in acute and chronic pain: a qualitative systematic review. *Acta Anaesthesiol Scand* 1998; **42**: 71–79.
62. van Rijkom H, Truin G, van t Hof M. A meta-analysis of clinical studies on the caries-inhibiting effect fluoride gel treatment. *Caries Res* 1998; **32**: 83–92.
63. Crider A, Glaros A. A meta-analysis of EMG biofeedback treatment of temporomandibular disorders. *J Orofac Pain* 1999; **13**: 29–37.
64. Ernst E, White A. Acupuncture as a treatment for temporomandibular joint dysfunction: a systematic review of randomizedtrials. *Arch Otolaryngol Head Neck Surg* 1999; **125**: 269–272.
65. Forssell H, Kalso E, Koskela P, Vehmanen R, Puukka P, Alanen P. Occlusal treatments in temporomandibular disorders: a qualitative systematic review of randomized controlled trials. *Pain* 1999; **83**: 549–560.
66. Ismail A, Bandekar R. Fluoride supplements and fluorosis: a meta-analysis. *Community Dent Oral Epidemiol* 1999; **27**: 48–56.
67. Ismail A, Sohn W. A systematic review of clinical diagnostic criteria of early childhood caries. *J Public Health Dent* 1999; **59**: 171–191.
68. Kellstein D, Waksman J, Furey S, Binstok G, Cooper S. The safety profile of nonprescription ibuprofen in multiple-dose meta-analysis. *J Clin Pharmacol* 1999; **39**: 520–532.
69. Kim J, Viana M, Graber T, Omerza F, Begole E. The effectiveness of protraction face mask therapy: a meta-analysis. *Am J Orthod Dentofacial Orthop* 1999; **115**: 675–685.
70. Kropmans T, Dijkstra P, Stegenga B, De Bont L. Therapeutic outcome assessment in permanent temporomandibular joint disc displacement (review). *J Oral Rehabil* 1999; **26**: 357–363.
71. McQuay H, Carroll D, Moore R. Injected morphine in postoperative pain: a quantitative systematic review. *J Pain Symptom Manage* 1999; **17**: 164–174.
72. Randall R, Wilson NH. Glass-ionomer restoratives: a systematic review of a secondary caries. *J Dent Res* 1999; **78**: 628–637.
73. Chan E-Y, Thornhill M, Zakrzewska J. Interventions for treating oral lichen planus (Cochrane review). In: *The Cochrane Library. Issue, 1, 2000*. Oxford: Update Software, 2000.
74. Clarkson J, Worthington H, Eden O. Interventions for preventing oral mucositis or oral candidiasis for patients with cancer receiving chemotherapy (excluding head and neck cancer) (Cochrane review). In: *The Cochrane Library. Issue, 1, 2001*. Oxford: Update Software, 2001.
75. Harrison J, Ashby D. Orthodontic treatment for posterior crossbites (Cochrane review). In: *The Cochrane Library. Issue, 1, 2001*. Oxford: Update Software, 2001.
76. Poulsen S, Errboe M, Hovgaard O, Worthinton H. Potassium nitrate for dentine hypersensitivity (Cochrane review). In: *The Cochrane Library. Issue, 1, 2001*. Oxford: Update Software, 2001.
77. Needleman I, Gledrys-Leeper E, Tucker R, Worthington H. Guided tissue regeneration for periodontal infra-bony defects (Cochrane review). In: *The Cochrane Library. Issue, 1, 2001*. Oxford: Update Software, 2001.
78. Landis J, Koch G. The measurement of observer agreement for categorical data. *Biometrics* 1977; **33**: 159–174.
79. Ioannidis J. Effect of the statistical significance of results on the time to completion and publication of randomized efficacy trials. *JAMA* 1998; **279**: 281–286.
80. Scherer R, Dickersin K, Langenberg P. Full publication of results initially presented in abstracts. A meta-analysis. *JAMA* 1994; **272**: 158–162.
81. Egger M, Zellweger-Zahner T, Schneider M, Junker C, Lengeler C, Antes G. Language bias in randomised controlled trials published in English and German. *Lancet* 1997; **350**: 158–162.
82. Deeks J, Altman D, Bradburn M. Statistical methods for examining heterogeneity and combining results from several studies in meta-analysis. In: Egger M, Davey Smith G, Altman D, eds. *Systematic reviews in health care. Meta-analysis in context*. London: BMJ Publishing Group, 2001.