This is a repository copy of Change in children's oral health-related quality of life following dental treatment under general anaesthesia for the management of dental caries: a systematic review.

White Rose Research Online URL for this paper:
http://eprints.whiterose.ac.uk/104277/

Version: Accepted Version

**Article:**
Knapp, R. orcid.org/0000-0002-8899-6548, Gilchrist, F. orcid.org/0000-0002-0418-6274, Rodd, H.D. et al. (1 more author) (2017) Change in children's oral health-related quality of life following dental treatment under general anaesthesia for the management of dental caries: a systematic review. International Journal of Paediatric Dentistry, 27 (4). pp. 302-312. ISSN 0960-7439

https://doi.org/10.1111/ipd.12259

This is the peer reviewed version of the following article: Change in children's oral health-related quality of life following dental treatment under general anaesthesia for the management of dental caries: a systematic review, which has been published in final form at http://dx.doi.org/10.1111/ipd.12259. This article may be used for non-commercial purposes in accordance with Wiley Terms and Conditions for Self-Archiving.

**Reuse**
Unless indicated otherwise, fulltext items are protected by copyright with all rights reserved. The copyright exception in section 29 of the Copyright, Designs and Patents Act 1988 allows the making of a single copy solely for the purpose of non-commercial research or private study within the limits of fair dealing. The publisher or other rights-holder may allow further reproduction and re-use of this version - refer to the White Rose Research Online record for this item. Where records identify the publisher as the copyright holder, users can verify any specific terms of use on the publisher's website.

**Takedown**
If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.
Change in children’s oral health-related quality of life following dental treatment under general anaesthesia for the management of dental caries: A systematic review.

Rebecca Knapp\textsuperscript{1}, Fiona Gilchrist\textsuperscript{2}, Helen D Rodd\textsuperscript{3} and Zoe Marshman\textsuperscript{4}

\textsuperscript{1}BDS, PGDip (Public Health). PhD Candidate, Unit of Dental Public Health, School of Clinical Dentistry, Claremont Crescent, Sheffield, S10 2TA, UK

\textsuperscript{2}BDS, MPaed Dent, FDS (Paed Dent), PhD. Senior Lecturer and Honorary Consultant in Paediatric Dentistry, Academic Unit of Oral Health and Development, School of Clinical Dentistry, Claremont Crescent, Sheffield, S10 2TA, UK

\textsuperscript{3}BDS (Hons), FDS(Paed), PhD. Professor and Honorary Consultant in Paediatric Dentistry, Unit of Oral Health and Development, School of Clinical Dentistry, Claremont Crescent, Sheffield, S10 2TA, UK

\textsuperscript{4}BDS, MPH, DDPH, FDS (DPH), PhD. Reader and Honorary Consultant in Dental Public Health, Academic Unit of Dental Public Health, School of Clinical Dentistry, Claremont Crescent, Sheffield, S10 2TA, UK

\textbf{Running title:} Change in children’s OHRQoL following dental GA

\textbf{Correspondence address:} Rebecca Knapp,

Unit of Dental Public Health,

School of Clinical Dentistry,

Claremont Crescent, Sheffield, S10 2TA, UK.

Email:\[\texttt{becky.knapp@sheffield.ac.uk}\]

\textbf{Word count (excluding references & tables):} 4160
Summary
Background: Dental caries has significant impact on children and their families, and may necessitate treatment under general anaesthesia (GA). The use of oral health-related quality of life (OHRQoL) measures enables evaluation of dental treatment from a patient’s perspective.

Objective: This systematic review aimed to assess change in OHRQoL in children following treatment under GA for the management of dental caries.

Methods: A comprehensive search was conducted to identify articles which were assessed against inclusion criteria before data extraction. Studies involving children under 16-years, having treatment for dental caries under GA, were considered eligible. Included studies were quality assessed.

Results: Twenty studies were included, which demonstrated significant heterogeneity. Most studies employed a pretest-posttest design. All but one study relied on proxy reports of OHRQoL. Only half the studies used instruments validated in the study population. While all studies reported improved OHRQoL overall, some subscales showed changes which were not significant or worsened OHRQoL. The scientific quality of the studies varied considerably.

Conclusion: Heterogeneity of included papers limited the conclusions which could be drawn. Treatment under GA appears to result in overall improvements in proxy-reported OHRQoL, however, there is a need for further high quality studies employing validated, child-reported measures of OHRQoL.

Introduction
Untreated dental caries in children is the tenth most prevalent disease worldwide, affecting 621 million children globally. In the UK alone, nearly one third of children
aged 5-years have experience of caries, rising to almost a half of 8-year-olds. There are widening inequalities, with children from a lower socioeconomic group being more likely to experience, and have greater severity of, dental caries than those from higher socioeconomic groups. The negative impact of dental caries on children and their families is well documented, and is associated with a number of factors which affect quality of life in children, including pain, impaired function and loss of school days.

Delivery of treatment to children can be difficult, especially where extensive treatment is required. Where other behavioural or pharmacological techniques fail, a general anaesthetic (GA) may be required to deliver effective treatment for dental caries. In England alone, there were approximately 42,000 hospital admissions of children under 16-years with a diagnosis of dental caries in 2014-15. The majority of these admissions were in the 5 to 9-year-old age group, with 33,871 cases of children under 10-years requiring tooth extractions under GA due to dental caries, at an estimated cost of £30 million.

Over recent years, there has been a move to evaluate not only clinical outcomes of healthcare interventions, but patient-reported outcomes. Oral health-related quality of life (OHRQoL) is one such measure, which aims to evaluate the impact of oral health on daily life.

A previous systematic review included studies, published 1978-2009, which report OHRQoL in children undergoing dental treatment under GA. However not all of these studies reported change in OHRQoL following treatment. A number of papers have been published since this review which warrant further systematic investigation. In addition, as no quality assessment of included papers was carried out by Jankauskeine and Narbutaite, it is unclear how robust the included studies were.
Aim
This systematic review was conducted to provide an updated appraisal of the recent body of research reporting change in OHRQoL in children following treatment under GA for the management of dental caries. The specific objectives were: to describe changes in OHRQoL reported in these studies; to describe the instruments used to measure OHRQoL and to examine the quality of the studies using a validated quality assessment tool.

Methods
This review was conducted in accordance with published guidelines for undertaking a systematic review.\textsuperscript{10,11}

Inclusion and Exclusion Criteria
Inclusion and exclusion criteria were determined by discussion and agreed by three investigators (RK, ZM and FG) based on the population, intervention, comparator, outcome and study design (PICOS) model\textsuperscript{10} as shown in Table 1.

Search Strategy
To ensure the review was as comprehensive as possible, an attempt was made to identify all relevant studies, regardless of year of publication or language. Database searches were carried out from date of inception to present of MEDLINE (1946-), Scopus (1966-) and Web of Science (1900-) using free text and MeSH terms individually and combined with Boolean operators. The following terms were included in the search strategies: oral health, quality of life, dental treatment, general
anaesthesia, dental care for children. Citation and reference list searching of included studies was carried out to identify other relevant studies. Duplicates were recorded and removed at this stage. The searches were re-run prior to final analyses to identify any further studies.

**Eligible Study Selection**

Titles and abstracts were independently reviewed against inclusion criteria by two investigators (RK and FG). Where titles and abstracts met or appeared to meet the inclusion criteria, the full text was obtained to determine eligibility for inclusion in the review. This process was carried out by two researchers independently. RK reviewed all full text papers, with ZM, FG and HDR each reviewing a third of the papers. Where reviewers disagreed on eligibility, they met to discuss and reach a conclusion. Where agreement could not be reached the opinion of a third reviewer was sought. Studies which did not meet the criteria at this stage were noted along with the reason for exclusion. Where a paper was not written in English, a translation was obtained to enable assessment against the criteria and subsequent data extraction.

**Data Extraction**

Data were extracted using a custom spreadsheet to record the following: author and publication year, study design, sampling and data collection methods, number of participants, caries experience (recorded as dmft/DMFT), demographic details (age, socio-economic status), treatment received (exodontia only or comprehensive dental care involving restorations as well as extractions), instrument (questionnaire) used, measured outcomes (change in OHRQoL and any secondary outcomes) and whether findings were statistically and/or clinically significant.
Initially, the data extraction spreadsheet was piloted using three articles; all reviewed by three investigators independently (RK, ZM, FG). This exercise gave the opportunity to refine the spreadsheet, and any disagreements in the extraction data were resolved by discussion. A final version of the data extraction sheet was produced following these discussions. Subsequently, three teams of two investigators (RK/ZM, RK/FG and RK/HDR) independently carried out the data extraction for each paper. Where there were discrepancies, these were resolved by discussion. Where agreement could not be reached the opinion of a third reviewer was sought.

**Quality Assessment**

The same teams of two reviewers then independently assessed the quality of included studies using the Quality Assessment Tool for Studies of Diverse Design (QATSDD), which has shown good reliability and validity for use with a range of study designs. This tool includes 16 items, which are scored between 0 and 3. Two of the items were not evaluated as they were only relevant to qualitative studies, giving a total possible score of 42 from 14 items. Total scores for each paper and the mean score for each criterion met by the included papers were calculated. Disagreements between the reviewers over the quality assessment were resolved by discussion, with a third reviewer invited to resolve issues where necessary.

**Results**

The search strategy yielded 325 records. Following removal of duplicates, 121 titles and abstracts were screened against the inclusion and exclusion criteria. In all, 28 full text articles were obtained and screened against inclusion criteria by two reviewers.
independently, at which point a further six were excluded. Twenty studies, which had been reported in 22 different articles, were included in the final review (see Figure 1).

Description of the studies
Most of the included studies were prospective longitudinal studies (n=18). One study was a randomised controlled trial, but randomisation groups were created to measure the effect of the pretest questionnaire rather than treatment under GA itself\textsuperscript{13}. One study carried out secondary analysis of data retrospectively\textsuperscript{14}. The majority of the prospective studies employed a single group pretest-posttest study design, with just one study including a cross-matched control group\textsuperscript{5}. However, OHRQoL was only measured at one time point in this control group, limiting its value in allowing comparison with the intervention group, where change in OHRQoL was measured.

The studies were conducted in 14 different countries, with the majority based in a hospital setting and the remaining four studies conducted in a community clinic\textsuperscript{13–16}. Only one study\textsuperscript{13} used random sampling, with the other studies using convenience sampling (n=9) or consecutive sampling (n=10). Nineteen studies were published in English, with just one study published Mandarin which was subsequently translated by a dental colleague\textsuperscript{17}.

Data Collection
The method of data collection varied across the studies and across time points within those studies as shown in Table 2. The majority of the studies used self-completed questionnaires on the clinic as the primary method of data collection. For the posttest questionnaire only seven studies used this method in isolation, with a further seven using a combination of methods. One study used a combination of self-completion on
clinic and self-completion by post, depending on which arm of the study a participant
had been randomly assigned to. In the remaining six studies, the researchers
attempted to use self-completed questionnaires on clinic for the posttest time point,
but then conducted structured interviews by telephone or self-completed
questionnaires by post if participants failed to attend their follow up
appointment.

In most cases, it was unclear whether the same parent/caregiver completed
both the pre- and posttest questionnaires. Only five studies specifically documented
that it was the same person in both cases. In three studies it was reported that a
percentage of the questionnaires were completed by different people, ranging from
1.6% to 9.2% of instances.

There were marked differences between the studies in the timing of the
completion of questionnaires. In ten studies, the pretest questionnaire was completed
on the day of the GA itself. In one study the questionnaire was completed the day before the GA, and in two studies it was between one and two
weeks prior to the GA. In the remaining seven studies it was unclear how far in
advance of treatment the questionnaires were administered.

In the majority of studies, just one posttest questionnaire was administered.
Seven studies carried this out four weeks after treatment. Six studies
administered the posttest questionnaire earlier than this, between one and four
weeks. Two studies collected data at between four and eight weeks, one study at three months, one study at six months and one study at between six
and nine months. In two studies it was unclear when the posttest questionnaire was
carried out. Just two studies administered a second posttest questionnaire, in both
cases this was three months after treatment.
Response Rates

Loss to follow up was reported in fourteen of the studies, with figures ranging from 0%\textsuperscript{5,22} to 47.8%\textsuperscript{23}, and a mean loss to follow up of 18.8%. In only seven studies were the characteristics of those lost-to-follow-up participants considered, but these studies found no difference in characteristics between the groups\textsuperscript{5,16,26,19,20,23,18}.

Participant Characteristics

The number of participants in the studies ranged from 28\textsuperscript{17} to 352\textsuperscript{30} (median: 88, interquartile range: 68, 140). The age of the children undergoing treatment in the included studies ranged from 2.3 years\textsuperscript{30} to 15.1 years\textsuperscript{19}, with the mean age across all the studies being 4.6 years. In one study\textsuperscript{18} children received either comprehensive care or exodontia only treatment, with the remaining studies all involving comprehensive care treatment only. Just ten studies recorded caries experience\textsuperscript{5,15,25,30,26,17,24,28,20,23}. Caries experience was recorded as the total number of primary and secondary decayed, missing and filled teeth (dmft/DMFT). At baseline, mean dmft/DMFT in the studies ranged from 6.9\textsuperscript{28} to 13.3\textsuperscript{24}. Within individual studies, the caries experience of individuals varied considerably. For example, in the study by Anderson and colleagues, the baseline dmft/DMFT in the study sample ranged from 1 to 18\textsuperscript{25}.

Oral Health-Related Quality of Life

A range of instruments were employed to measure OHRQoL (Table 3). Two studies\textsuperscript{21,26} designed their own questionnaires with the remainder employing pre-existing questionnaires. However, of these, only nine of the included studies used instruments which had been previously validated in the study population, or included
validation of the instrument as part of their study\textsuperscript{5,7,14,16,26,27,28,18}. The most commonly used instrument was the Early Childhood Oral Health Impact Scale (ECOHIS), used in nine of the studies. One study used the Child Perceptions Questionnaire (CPQ), which was the only study to use a child-reported measure of OHRQoL\textsuperscript{23}. All the other studies relied solely on parent/caregiver reported outcomes, in the form of the ECOHIS or the Parental-Caregiver Perceptions of child oral health-related quality of life (P-CPQ) questionnaire.

There was significant heterogeneity in how the studies reported change in OHRQoL, and therefore a summary of the findings of each paper, ordered by ascending year of publication, is given in Table 3. In all the included studies an overall improvement in OHRQoL was seen, however, improvements were not found across all subscales in some studies.

All but one study\textsuperscript{29} applied statistical tests to determine whether there were significant differences in OHRQoL following treatment. The majority of the studies found a significant change in both overall and subscale scores. However, some studies found that within the subscales there was not always a significant change in score.\textsuperscript{15,24}

Interestingly, some studies found an increase in some subscale scores, i.e. worse OHRQoL, following the dental GA. For example, two studies\textsuperscript{24,28} found an increase in mean score for the ECOHIS ‘child self-image and social interaction’ subscale and another study\textsuperscript{15} found an increase in mean P-CPQ ‘social well-being’ subscale score, albeit not statistically significant.

As well as considering if the change was statistically significant, three studies\textsuperscript{16,19,18} also looked at whether the change was clinically significant by calculating minimally important difference (MID). Two studies found 63% of the population showed or exceeded the MID for the P-CPQ, but only 40% did so for the
Family Impact Scale (FIS)\textsuperscript{16,19}, whereas de Souza and colleagues found 54\% of the population showing or exceeding the MID for the P-CPQ and 65\% for the FIS\textsuperscript{18}.

Eleven studies included a measure of effect size, the results of which are also given in Table 3. Large to moderate effect sizes were seen for overall changes and in all subscales, with the exceptions being the small effect sizes seen in the ECOHIS ‘child psychology’\textsuperscript{27} and ECOHIS ‘child self-image and social interaction’ subscales\textsuperscript{27,20}. Only half of the studies asked a global transition judgement (GTJ) question, however, seven of the ten studies which included a GTJ did not then correlate this to the change in OHRQoL scores as recommended by the COSMIN group.

Secondary Outcomes
Nine studies also reported secondary outcomes. Three studies used the Dental Subscale of the Children's Fear Survey Schedule (CFSS-DS) to measure change in dental anxiety\textsuperscript{13,15,28}. In the studies by Klaassen and coworkers no significant difference was found between pretest and posttest anxiety scores\textsuperscript{13,15}. In contrast, Cantekin and colleagues found a statistically significant decrease in CFSS-DS score post-treatment, indicating an increase in dental anxiety\textsuperscript{28}. Thomas and Primosch recorded change in weight, but found no significant difference 18 months post-treatment\textsuperscript{29}. The remaining four studies included ‘parental satisfaction’ as a secondary measure, all of which developed their own questionnaires to measure this change. No statistical tests were applied to these data, but all four studies found high levels of parental satisfaction, with 80-100\% of parents reported as being ‘satisfied’ with the treatment\textsuperscript{25,30,27,20}. Two studies also looked at clinical outcomes, in the form of ongoing caries experience\textsuperscript{30,17}. El Batawi and colleagues\textsuperscript{30} found 59\% of participants
had new carious lesions within two years of treatment and Xiao and coworkers\textsuperscript{17} found that 37\% of participants had new carious lesions after six months.

Quality Assessment of the Studies

Study quality varied considerably, and out of a total possible QATSDD score of 42, scores for the individual studies ranged from 7\textsuperscript{29} to 32\textsuperscript{16}. The average score was 22(±7). Table 4 shows the mean score for each of the 14 criteria of the quality assessment. A mean score of 0 indicates none of the papers met any of the components of the criteria, with a total possible score of 3 indicating all the papers fully met the criteria.

Some quality criteria were well addressed by the included studies, in particular the fit between the research question and method of data collection and analysis. However, none of the included papers had evidence of user involvement in the design and there was a lack of explicit theoretical framework underpinning the majority of the studies. Two other areas less well addressed were the approaches taken to estimate the sample size and assessment of reliability and validity of the measurement tools used.

Inter-reviewer Reliability

The two independent reviewers agreed on 115/121 (95\%) abstracts when screened against the inclusion and exclusion criteria, and the disagreements were resolved by discussion. Independent reviewers then agreed on 26/28 (93\%) full text articles screened against the inclusion and exclusion criteria, and again the disagreements were resolved by discussion.
At the data extraction stage, independent reviewers agreed on 605/684 (88.5%) data extraction criteria, with the areas of disagreement resolved through discussion.

For the quality assessment, there was overall agreement between independent reviewers of 176/280 (62.9%) of quality criteria scores. However, because quality assessment was an ordered variable, a weighted kappa was also carried out to establish relative concordance between reviewers. It was assumed that the differences between individual quality scores were equal. The inter-rater agreement (kappa with linear weighting) was 0.65 (95% CI, 0.59- 0.72) indicating substantial agreement.

**Discussion**

This systematic review examined 20 studies, reported across 22 papers. It was clear that all of the studies reported an overall improvement in OHRQoL in children following dental treatment under GA. Within studies, however, there were differences in the change score for individual subscales. Interestingly, in some cases, results suggested that some aspects of OHRQoL may worsen following dental treatment under GA. There could be a number of reasons for this. Many of the studies carried out the posttest questionnaire within 4 weeks following treatment, at which point children may still be experiencing discomfort from extraction sites or difficulty eating due to the number or difficulty of extractions. Further research is indicated to add to this body of evidence. In particular, future work should explore whether the actual number of extractions impacts on OHRQoL; one might expect that children who have higher numbers of extractions are more likely to experience the negative side effects of post-treatment discomfort or impaired function.
In contrast with the previous systematic review\(^9\), the majority of studies involved instruments which had been used in other studies, with just two studies developing their own questionnaires\(^{21,26}\). However, less than half of the included studies used instruments that had been previously validated for the study population, or included validation of the instrument as part of their study. Importantly, it has been shown that the properties of quality of life instruments should be evaluated when used in a different context to the one in which they were developed\(^{31}\). There is, therefore, still a need for further research using validated instruments to evaluate change in OHRQoL following a dental GA, and for longitudinal validation of OHRQoL instruments.

Despite recommendations made in previous studies\(^{9,15,22,19}\) only one study\(^{23}\) to date has employed a child-reported measure of OHRQoL, with the other studies relying on proxy reports of OHRQoL. Caution should therefore be exercised when interpreting some of the findings as it has been shown that parents/caregivers generally have a low to moderate overall agreement with their child’s ratings\(^{32,33}\). A systematic review of parent and child reports of health-related quality of life (HRQoL) by Eiser and Morse\(^{34}\) revealed greater agreement between proxy and child ratings in some subscales (e.g. physical HRQoL) than other, less-observable, subscales (e.g. emotional or social HRQoL). This highlights the need for child-reported measures to be used in future OHRQoL research.

A limitation of some of the included studies was that, in some instances, different individuals completed the pretest and posttest questionnaires. These change scores were included in the final analysis, despite this discrepancy in rater potentially impacting on the scores. Future studies should ensure a consistency in respondents for all time points in the study.
The use of convenience samples and lack of controls needs consideration. A consecutive sample would be preferable as it would better represent the whole population. The majority of studies highlight the issues in obtaining a suitable control for this population, where withholding treatment would be unethical. Where random allocation is not possible, it may be possible to improve the validity of inferences by using statistical techniques to adjust for potential confounders. The disadvantage of this option is that to adequately adjust for confounders, all potential confounding variables must be identified and accurately measured. Inadequate identification and measurement of confounding factors has been identified as a deficit in observational studies. A clear framework underpinning the research is important for understanding which factors may impact the outcome of interest; something which is lacking in all the included studies. To improve the quality of future research, the underpinning theoretical framework should be clearly stated.

While all the studies stated that children were undergoing a GA for the treatment of dental caries alone, only ten studies recorded the level of caries experience of their sample. Recording caries experience and number of decayed teeth, e.g. as dmft/DMFT, would have been useful to give an indication of the burden of disease in the study population, and also because this may influence changes in OHRQoL following treatment. Interestingly, even in those studies which reported caries experience using dmft/DMFT, there was no reference to this in the subsequent analysis or discussions. It might have been useful for caries experience to have been taken into account and perhaps correlated against pretest OHRQoL scores.

Only one study considered difference in change in OHRQoL according to treatment approach, and found no significant difference between extraction only and comprehensive care (i.e. including restorations) groups. However, the sample size
was relatively small and more extensive studies should be carried out to validate this finding.

**Study strengths and limitations**

**Strengths**
The present review employed a comprehensive search strategy and should therefore fully represent the current literature base. The inter-reviewer reliability assessments show a substantial level of agreement, adding to the reliability of the findings. This is the first time a quality assessment of the included studies has been carried out, which has shown there is significant variability in the quality of the studies reporting OHRQoL changes. This assessment also highlighted key areas for improvement in quality, which is of use to those planning future research. The findings have also highlighted areas of discrepancy in the current literature, for example variation within subscales of the measures following intervention.

**Limitations**
There was significant heterogeneity between the studies so it was not possible to carry out a meta-analysis of the findings, which also limits the conclusions that can be drawn. During the literature search, hand-searching of journals was not undertaken thereby potentially omitting some articles. However, it was felt that through thorough database searching, citation searching and reference searching that the search strategy should have been exhaustive, without the need for hand searching. The inclusion of all papers, regardless of language, presented a difficulty with obtaining an accurate translation. Although it was possible to extract all of the relevant data for inclusion in the review itself, it was not possible to calibrate the translator to use the QATSDD
tool and therefore the quality assessment for this paper was not completed. Quality assessment may have an element of subjectivity and the QATSDD tool does not weight individual criteria by importance or degree of impact on quality. The total scores should therefore be used with caution. Subjectivity was, however, reduced by having more than one reviewer complete the quality assessment for all the studies.

**What this paper adds**

- Consolidation of the evidence for improvements in proxy-reported child OHRQoL following dental treatment for caries under GA.
- Evidence for the need for future research to use child-reported measures of OHRQoL.
- Evidence for variable quality in articles reporting change in OHRQoL in children undergoing treatment for dental caries under GA, highlighting areas for improvement in future studies.

**Why this paper is important to paediatric dentists**

- It provides a summary of the evidence to date which would justify the use of GA in the treatment of dental caries due to overall proxy-reported improvements in OHRQoL.
- It highlights the need for future research to compare the impact of different treatment approaches and to examine the long term impact of treatment, which in turn will better inform clinical practice and provide justification for treatment options. This may have implications for future public policies and commissioning of GA services in Paediatric dentistry.
Acknowledgements

The authors would like to thank Albert Yeung, Consultant in Dental Public Health, NHS Lanarkshire, for his translation services.

Conflict of Interest

The authors declare that they have no conflicts of interest in the publication of this paper.

References

5  Baghdadi ZD. Children’s oral health-related quality of life and associated


28 Cantekin K, Yildirim MD, Cantekin I. Assessing change in quality of life and


36 von Elm E, Altman DG, Egger M, Pocock SJ, G??tzsche PC, Vandenbroucke

Figure legends

Figure 1

PRISMA diagram to show the process of study selection

Tables

Table 1
Inclusion and exclusion criteria used to assess eligibility of the articles

Table 2
Methods of data collection used in the included studies at baseline and follow-up

Table 3
Summary of the main findings from the included studies

Table 4
Mean score for each quality criteria against which the papers were assessed