Quality of Meta-Analysis in terms of actual Statistical analysis

Edith Njideka Jude-Eze

A Dissertation Submitted to the
University of Glasgow
for the degree of
Master of Science

School of Mathematics and Statistics

March 2011

Abstract

Meta-analysis is an essential tool that facilitates clinicians, medical experts and decision makers to cope with the information overload in the public and healthcare sectors. The publication of meta-analyses is increasing rapidly every day with less or more of methodological rigour. Clinicians and medical experts depend wholeheartedly on the results and conclusions obtained from analyzing meta-analysis in order to assess the clinical effectiveness of healthcare intervention on a daily basis. Meta-analysis provides a specific estimate of a relationship which may also indicate if there is any need for further research.

But the foundational problem of performing the clinically powerful meta-analysis is the guideline for how similar the studies must be in order to meet the inclusion criteria of the meta-analysis and the reliability of its conclusions [1]. When there are discrepancies in the studies being combined and patient populations being studied, meta-analysis may provide results that are wrong. These may mislead potential users of meta-analysis to give wrong prescriptions to their patients.

One requirement is to prepare a validated and reliable checklist that can assess the quality of meta-analysis in terms of reporting, methodology, science and most especially, the actual statistical analysis. Literature review reveals that existing checklists mainly focus on other aspects of quality with little or no attention to the quality of statistical methodology. Consequently, this thesis attempts to cover this gap.

OBJECTIVE- To construct an appropriate validated quality instrument. Use the instrument to assess the quality of selected meta-analyses in terms of actual statistical analysis. To assess accuracy and consistency of reported estimates using Lee's [2] methods for checking errors in reported relative risks, odds ratios and confidence intervals.

STUDIES - Eligible articles (Meta-analysis of randomised controlled trials) identified in the Cochrane database, Web of Knowledge and Medline databases were used in this project. Eligibility criteria include studies published in English language, as a full report between the periods of 2000-2008, have a comprehensive search strategy and have clear methods of selecting studies for inclusion and performed statistical analysis.

We developed a checklist that measures the quality of meta-analysis in terms of actual statistical analysis and used the instrument to assess papers published in both Cochrane and Non-Cochrane reviews.

RESULTS: A sample size of 100 papers was obtained using an estimated maximum error bound of 0.1. Studies were allocated equally between Cochrane and Non-Cochrane publications and selections were made from electronic databases. Records of meta-analysis of randomised contolled trials published in English, full text and journal articles between the periods of 2000 - 2008 show that there were 515 results out of 5821 records of meta-analysis published in Cochrane library, 507 out of 1434 records and 130 out of 135 records of meta-analysis published in Web of Knowledge and Medline respectively. Simple random sampling, implemented in R statistical package, was used to select random sample of studies from each database. 83 out of the 100 selected studies met the inclusion criteria - 42 studies from Cochrane reviews and 41 from Non-Cochrane reviews.

Reporting and methodology quality are high in the two databases. However, in terms of statistical analysis, both databases are unlikely to explicitly state the design of individual studies combined in the meta- analysis. The Cochrane review is more likely to contact authors of published studies than their Paper-base counterparts. Cochrane reviews are less likely also, to use OQAQ(Overview Quality Assessment Questionnaire) and QUOROM (Quality of Reporting of Meta-analyses)in asssessment of validity of studies than paper reviews.

There was no double counting of some aspects of studies identified among Paper-base

Journals while we discovered four studies in Cochrane reviews that double counted the control arms.

However, there was no simple double counting of studies found in both Cochrane and Non-Cochrane reviews.

Lee's checks were performed on the twenty selected studies to verify errors on reported odd ratios, relative risks and confidence intervals. Some studies included in the meta-analysis reported zero events either in the treatment or control groups or both which led to a disparity between our calculated results and the estimates reported by the authors. The addition of a continuity correction factor of 0.5 to each cell of the studies with zero events took care of the disparities. Mabinary sas macro designed by Weir and Senn [3] was also used to assess and check the validity of reported odd ratios, relative risks and confidence intervals on both reviews. The results obtained using the macro are consistent with the original reported results in most of the studies.

Studies reporting relative risks in both Paper-Base Journals and Cochrane reviews are more likely to disagree with the Lee's requirement on minimum subject size and number of diseased subjects in either exposure groups given the CI, than those reporting odd ratios. These studies also have large outcomes. This seems to suggest that Lee's checks are not reliable for studies reporting relative risks, especially when outcomes are relatively large.

CONCLUSION: Cochrane Handbooks and scales relating to specific interventions were mostly used to assess quality of studies in Cochrane reviews. Results showed no statistically significant difference between the reporting and methodological quality of Cochrane and non-Cochrane publications. More improvement is needed in the reportage of the design of included studies in both Cochrane and non-Cochrane reviews. This will help establish if the combined studies and the statistical method used in combining them are compatible. However, double counting of some aspects of studies was found in some meta-analysis selected from Cochrane reviews. Analysis suggests that studies reporting odd ratios are likely to be

consistent with Lee's checks than those reporting relative risks. We also showed that Peter Lee's checks involving totals cannot be relied on to assess the quality of studies reporting relative risks.

Contents

List of Figures

List of Tables

ACKNOWLEDGEMENT

I give thanks to God for giving me the opportunity, strength, knowledge and understanding to work through to this research.

My heart-felt thanks and appreciation goes to Novartis, Boehringer Ingelheim and Department of Statistics, University of Glasgow who fully funded this research.

I offer my sincere gratitude to my supervisor, Professor Stephen Senn, whose encouragement, direction and support from the beginning to the conclusion of this research exposed me to greater understanding of the subject.

I thank Prof John McColl, my Internal Examiner and Prof James Carpenter my External Examiner for giving constructive comments and advice on this thesis.

I lack appropriate words to express my gratitude to my Sweetheart, Dr Jude Eze whose dedication, love and constant confidence in me has taken loads off my shoulder. I owe him everything. Also my three wise boys who provided me with emotional and physical support.

I would like to thank my Parents and Parents-in-Law, Late Benjamin Nwosu and Mercy Nwosu, Late Patrick Eze and Virginia Eze who had showed me the way and trained me up in the fear of God . My siblings Ifeyinwa, Amaka, Obinna and Okwudili for their supports.

Lastly, I offer my regards and blessings to all of those who contributed to the successful realization of this thesis.

DECLARATION

I have prepared this thesis myself; no section of it has been submitted previously as part of any application for a degree. I carried out all the work reported in it, except where otherwise stated.

E.N Jude-Eze

DEDICATION

TO MY HUSBAND AND CHILDREN

Chapter 1

INTRODUCTION

According to Kulik et al [4], meta-analysis has a long past and a short history. The history of meta-analysis began in 1976 when Glass[5] first used the term in his presidential address to the American Educational Research Association to describe the statistical analysis of findings from a large number of independent studies. Glass described meta-analysis as the "analysis of analyses". He went on to define Meta-analysis more formally as the statistical analysis of a large collection of analysis results from individual studies for the purpose of integrating the findings [5]. But the roots of meta-analysis go as far back as 1904 when Karl Pearson used formal techniques to combine data from different studies when examining the preventive effect of serum inoculations against enteric fever [6].

Fisher was one of the first to devise a means for transforming and combining P values. He noticed that the natural logarithm of a value given by P of a test multiplied by -2 is exactly distributed as a chi square with two degrees of freedom and that the sum of independent chi-square is also distributed as chi-square [7]. Cochran's method of integrating treatment effects was developed to deal with results from a planned series of studies [8]. According to Rosenthal [9], meta-analysis is the use of statistical techniques either to combine or compare either effect size measures for probability levels from two or more studies. Glass [5] also

outlined the following summary of steps which an analyst should undertake, (a) the meta-analyst uses objective methods to find studies for a review (b) Describe the features of studies in quantitative or quasi-quantitative terms (c) Expresses treatment effects of all studies or a common scale of effect size, and (d) uses statistical techniques to relate study features to study outcomes.

Hedge's contribution to meta-analysis was his demonstration that the effect size statistics usually calculated for meta-analyses were biased estimators of an underlying population effect [10].

Chalmers [11] devised a strategy for combining clinical trials to summarize findings and published the first meta-analysis in medicine. This earned him the 1982 annual research award of the Evaluation Research Society for his accomplishment.

The term 'meta-analysis' has been adopted within other disciplines and has proved particularly popular in clinical research [12]. Over the past one and half decades, the number of meta-analyses published annually has increased greatly[13]. Several hundred are now published each year [14].

According to the Cochrane Collaboration [15],"*It is an independent, not-for-profit organisation established in 1993, named in honour of Professor Archibald Leman Cochrane (1909 - 1988), a British epidemiologist who advocated the use of randomised controlled trials as a means of reliably informing healthcare practice. It is also an international network of people assisting healthcare providers, decision makers, policy makers, patients, their advocates and carers make well-informed decisions about human health care by preparing, updating and promoting the accessibility of Cochrane Reviews. Cochrane reviews are scientific investigation that synthesize results of multiple primary investigations using approaches that limit bias and random error. These approaches include a comprehensive search of all potentially relevant studies and the use of explicit, reproducible criteria in the selection of studies for review. Cochrane reviews are published in Cochrane library*".

The Cochrane library is an independent high quality evidence base for health-care decision making. It is an online collection of databases that brings together in one place rigorous and up-to-date research on the effectiveness of healthcare treatments and interventions, as well as methodology and diagnostic tests [16]. Hence, the rich storehouse of Cochrane Collaboration will provide us with limitless information on the quality of statistical content of meta-analysis published over the years [17].

In order to obtain the actual number of Meta-analysis published over the years, we made a simple search through the Web of Knowledge (ISI) by typing the words *Meta-analysis * OR *Meta analysis * OR *Metaanalysis *.The search yielded a total of 71,901 records. Analysing this number of publications by year, choosing a selected field (Publication year) to analyze up to 100000 records with a minimum record count of 10 from the set display options in the ISI Web of Knowledge Database, gave the result represented in the Figure1.1. Figure 1.1 shows the trend of publications of meta-analysis between 1944 and 2008. The first four decades (1944-1987) witnessed a slow but steady increase in the number of Meta–analyses published. However, the last two decades (1988-2008) saw an astronomical rise in the number of Meta-analyses published in various journals per year. A closer scrutiny of the graph show that in 1944 only 6 Meta-analyses were published compared to 10,350 published in 2008.

Cochrane library confirmed the increasing diffusion of meta-analysis with theoretical and methodological advances as well as from empirical research [18].A recent check in the Cochrane database of systematic reviews indicate that the total number of meta-analyses published was nearly 6000.

Citations and uses of meta-analyses in health-related literature have been rising very rapidly over the past decade and can offer a rational and helpful way of dealing with a number of practical difficulties that beset anyone trying to make sense of the effectiveness of research [18]. However, “it seems that Meta-analysis has not always been embraced

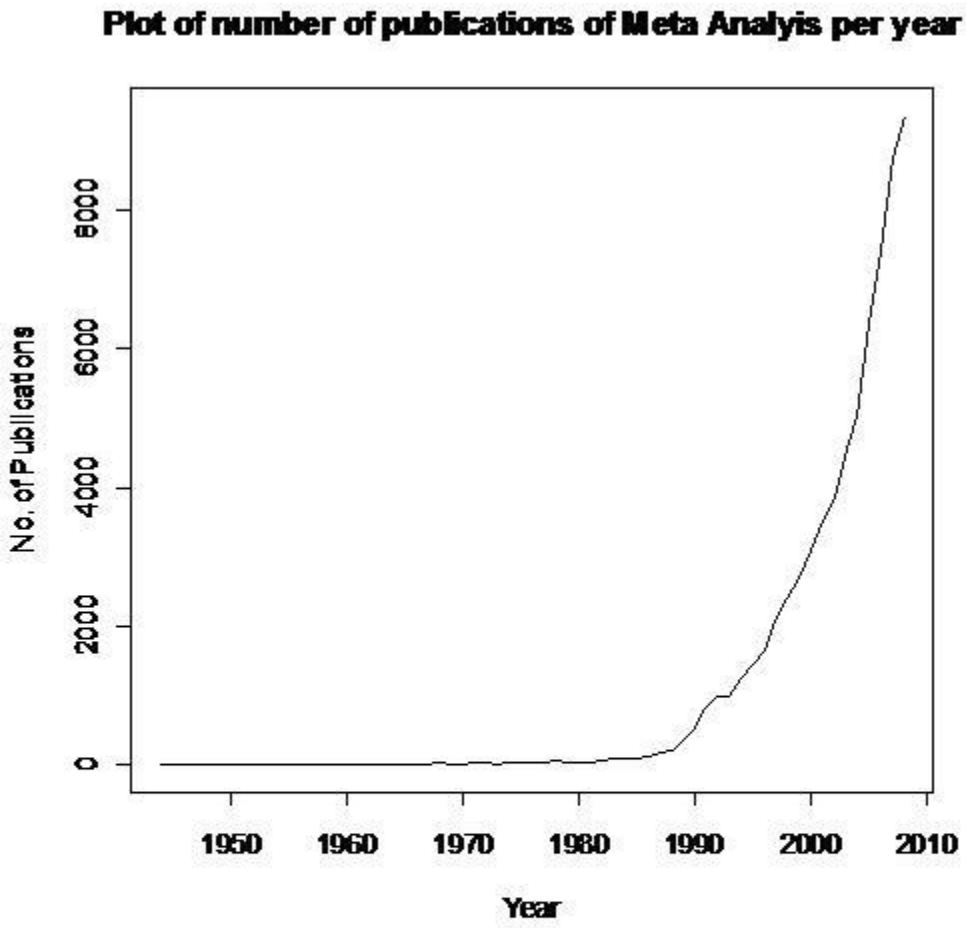


Figure 1.1: Trend in the publication of Meta-analysis

with enthusiasm within the Pharmaceutical industries"[19]. The International Conference Harmony(ICH) E9 [20] while acknowledging the usefulness of meta-analysis as a means of summarising overall drug efficacy also warns that confirmation of efficacy solely from meta-analysis will not usually be accepted as a substitute for confirmation of efficacy from individual trials. This is because the magnitude of the treatment effect is likely to be an important factor in regulatory decision-making. If the treatment effect is smaller than anticipated, then statistical significance may not be reached in the individual trials even if the statistical significance is reached in the meta-analysis. Therefore, it is considered insufficient for approval when the magnitude of treatment effect may not be clinically significant [21].

1.0.1 Three basic types of Meta-analysis

According to Senn [19], meta-analysis can be grouped broadly into three types depending on the nature of the data being summarized:

- Type A: the outcome measure is the same in all trials being analyzed, the analyst has access to all original data and chooses to base the analysis on these data.
- Type B: the outcome measure is the same in all trials being analyzed but the analyst uses summaries from each trialqw as the basis of analysis, the original data are not available.
- Type C: different outcomes have been measured in different trials and analysis has to proceed using unit-free summaries.

1.0.2 Conducting Meta-analysis

DeCoster [22] outlined the following steps for the conduct of meta-analysis.

- Define the theoretical relationship of interest

- Location of studies-: meta-analysis requires a comprehensive search strategy
- Selection of studies based on inclusion criteria stated, selection of quality of studies and selection of specific studies based on the specific subjects.
- Examine the distribution of effect sizes and analyse the impact of moderating variables
- Interpret and report the results

1.1 BACKGROUND

Meta-analyses have been suggested to be the utmost outline of evidence available to clinicians to guide clinical practice in critical care[23]. Given this strategic position, the quality of every Meta-analysis is expected to be of the highest standard [24]. Consequently, several authors have described the sources of bias and errors during the design, conduct, and analysis of meta-analyses, [25],[26],[27] and at least one instrument has been developed according to conventional methodological standards to assess the methodological rigour, reporting quality or scientific quality of review articles [28],[29].

Some of the popular instruments include:

OQAQ (Overview Quality Assessment Questionnaire) this is a validated instrument consisting of 10-items checklist, used to assess the scientific quality of an overview research. Mulrow [30] was first to draw attention to the poor scientific quality of healthcare review articles. Later, Oxman and Guyatt[31] published a guideline to help readers assess reviews in health care. Many authors have used this tool to evaluate and assess the validity of reporting and methodology, Example, Boluyt et al [32] used this tool to evaluate clinical, methodological and reporting aspects of systematic reviews on acute Asthma management in children and found that the methodological quality of both Cochran and Journal reviews seem good, with Cochran reviews being more rigorous. Bereza et al [33] considered this

instrument less than fair-to-good on their findings for assessing the reporting and scientific quality of meta-analyses of randomised controlled trials of treatments for Anxiety Disorders. Empirical evidence suggested that inadequate methodological reporting correlates with bias in estimation of treatment effects.

QUOROM (Quality Of Reporting Of Meta-analyses) consists of an 18-items checklist. A conference was convened to address standards for improving the quality of reporting of meta-analysis of randomised controlled trials, this resulted in the QUOROM statement, a checklist and a flow diagram for reporting on meta-analysis and systematic Reviews [34] . Example, Junehua et al [35] assessed the methodology and reporting quality of systematic reviews/meta-analyses of traditional Chinese Medicine, published in paper-based journals in china and concluded that the methodology and reporting quality are poor in both systematic reviews/meta-analysis of TCM. Schulz [36] stated that faulty reporting appears to portray faulty methods.

However, all these instruments failed to assess the quality of the statistical methods employed in the meta-analysis. It is a well known fact that the use of a wrong statistical method on the best data yields results that are at best misleading [37],[38]. The use of meta-analysis has been criticised on statistical grounds in the past,[39], [40]. Indeed, the potential for the misuse of meta-analysis has been well recognized even by those who have done the most to develop it as a methodology [41];[4].

Assessing the quality of meta-analyses in terms of actual statistical analysis is important and relatively new. Quality gives an estimate indication of the likelihood that the results are valid estimates of the truth and the major problems with the implementation of meta-analyses have been common [42], [43] ,[44],[45] . There has been a wide variety of these, including failure of investigators performing the meta-analysis to understand the basic issues, carelessness in abstracting,summarizing and concluding appropriate papers, failure to consider important covariates, bias on the part of the meta-analyst, and perhaps most often,

overstatements of the strength and precision of the results [46]. It is not unusual to find that two or more meta-analyses done at about the same time by investigators with the same access to the literature, reach incompatible or even contradictory conclusions [47], [48].

The dilemma of missing studies in meta-analysis has received much attention and less attention has been paid to the more serious problem of overstating the evidence [49]. These potential problems with meta-analysis concerning missing studies were due to publication bias, that is, studies with unfavourable outcomes tend to be suppressed , data missing at the study level, example, investigators may not report estimates of treatment effect size and/or study-level covariates in a publication [50] and also data missing at the individual patient level, such as non-response [51].

1.1.1 Problem of overstating the precision of results from meta-analysis

Senn [52] outlined ways in which precision could be overstated in meta-analyses as follows

- Simple double counting of studies
- Double counting of some aspects of studies,
- Accepting implausible claims for the precision of studies
- Imputing data
- Spurious precision of individual trials
- Inappropriate pooling of treatments,
- Numerical slips and poor reporting

1.2 SOURCES OF DATA

It was decided that studies should be selected from the Cochrane collaboration and non-Cochrane databases of systematic reviews using an appropriate sampling plan.

1.3 OBJECTIVES OF THE STUDY

- To evaluate the criteria currently being used in assessing the scientific quality of research reviews.
- To establish an appropriate quality instrument.
- To develop a pilot study to check the instrument's practicability
- To assess the quality of meta-analyses in term of actual statistical analysis.
- To compare the quality of Cochrane and non-Cochrane meta-analyses
- To establish the validity and reliability of the instrument.
- To summarise the results and draw suitable inferences
- To check errors in reported estimates(Relative Risks, Odds Ratios and Confidence Intervals) using Peter Lee's method.

Chapter 2

LITERATURE REVIEW

Our research interest was to develop a validated, comprehensive and reliable checklist that can assess the quality of meta-analysis in terms of actual statistical analysis. Numerous standardized checklists had been developed by various clinicians, editors, research groups and authors to measure and improve research quality on external and internal validity of clinical studies and meta-analyses, to assess the quality of reports of randomised controlled trials, to assess the methodological quality and scientific quality of included trials with particular emphasis on selection, performance, attribution, critical appraisal and detection of bias in a research medical literature. Little consideration has been given to the statistical quality of meta-analysis.

A checklist is a comprehensive list of an important or relevant action or step to ensure consistency or completeness in carrying out a task and it is formatted as list with small checkboxes down the left hand side of the page and use checkmarks or 'yes' or 'no' or 'unsure' to denote if an item on the checklist has been completed satisfactorily or if the characteristic is present [53].

2.1 USES AND APPLICATIONS OF CHECKLIST

A checklist can be applied in medical practice to ensure that clinical practice guidelines are carried out accordingly.

Wikipedia[53] listed the use and applications of checklist as follows:

- It can be used as a human factor aid in aviation safety to ensure that long lists of an items are not forgotten.
- A checklist can be used as a job-aid to help in evaluation and in decision-making.
- It is used to determine the value of process, procedure, decision, solution and outcome of research literature.
- It can be modified to check the level of quality, priority, importance and condition of each item in the list.

2.2 BENEFITS OF CHECKLIST

"Advantages of Checklist"as given by the Mindfire Solutions [54] on Checklist Driven Testing - An Overview are listed below:

- A checklist draws on a range of issues that help in deciding where to concentrate effort
- A checklist is a predefined guideline to quality, which is created over time and is an investment both of time and energy.
- A checklist helps guide developers in reducing errors and preventing errors in assessment
- Checklists work as a reminder to testers. The essential tests, which testers tend to forget , can be entered to checklists to ensure that these tests get executed

2.3 REVIEW OF CHECKLISTS

Early attempt at developing checklists for the evaluation of different aspects of quality of meta-analysis of randomised control trials can be traced to the early 1960s. It does seem that the first documented checklist is a procedure outlined by Mahon and Daniel [55] in 1964 to assess the accuracy of reports of drug trials. They were the first to identify a lack of rigour [56],[57], [58].

Mulrow(1987) [30] used a checklist on "Policy Research Incorporated Literature Review Validation Procedures Manual"[59],[60],[61], developed around 1979. She assessed fifty reviews published during June 1985 to 1986 in four major medical journals. Her results signified that the present medical reviews do not regularly apply scientific methods to identify, assess and synthesize information, then She drew attention to the poor scientific quality of healthcare review articles.

Sacks et al evaluated the quality of 86 meta-analyses of reports of randomized controlled trials in English-language literature, using a scoring method that considered 23 items in six major areas which includes - study design, combinability, bias control, statistical analysis, sensitivity analysis, and application of results. They concluded that an urgent need exists for improved techniques in literature searching, quality evaluation of trials, and synthesizing of the results [62].

2.3.1 (OQAQ)-Overviews of Quality of Assessment Questionnaire (1991)

Following the above assertions, Oxman and Guyatt(1991) [31] published guidelines to help readers assess the quality of reviews in health care and it was used to prepare valid checklist that measures the quality of research overviews. The instrument consists of 9-items checklist and is titled "Overviews of Quality of Assessment Questionnaire"(OQAQ) as itemised below.

The authors believed that this instrument can be productively used by the readers and editors of clinical trials to identify scientifically sound overviews and thus judge the confidence that should be placed in their conclusions [31].

Criteria for assessing the scientific quality of research overviews by Oxman and Guyatt(1991) [31] are

- *Were the search methods reported?*
- *Was the search comprehensive?*
- *Were the inclusion criteria reported?*
- *Was selection bias avoided?*
- *Were the validity criteria reported?*
- *Was validity assessed appropriately?*
- *Were the methods used to combine studies reported?*
- *Were the findings combined appropriately?*
- *Were the conclusions supported by the reported data?*
- *What was the overall scientific quality of the overview?*

2.3.2 Blind Assessment of the Quality of Trials Reports(1996)

From other perspective, the designed instrument failed to measure in detail the statistical quality of the overviews. Jadad et al(1996) [63] suggested that the quality of clinical trials should be assessed by blinded raters to limit the risk of introducing bias in meta-analyses and systematic reviews, and into the peer-reviews process. They assembled a multidisciplinary

panel of judges with an interest in pain research who also had experience in the development of instruments [63]. They generated items to be considered in the instrument termed Jadads'scale for quality of randomised controlled trials around 1996. This instrument (listed below) is used to assess the quality of clinical reports. Fisher et al [64] stated that results suggested that blinded reviewers may provide less biased reviews and unblinded reviewers may be affected by various type of bias.

The designed instrument by Jadad et al(1996) [63] are

- *Was the study described as randomized?*
- *Was the study described as double-blind?*
- *Was there a description of withdrawals and drop outs?*
- *Were the objectives of the study defined?*
- *Were the outcome measures defined clearly?*
- *Was there an explicit description of the inclusion and exclusion criteria?*
- *Was the sample size justified (e.g., power calculation)?*
- *Was there a clear description of the interventions?*
- *Was there at least one control (comparison) group?*
- *Was the method used to assess adverse effects described?*
- *Were the methods of statistical analysis described?*

2.3.3 (QUOROM)-QUality Of Reporting Of Meta-analysis of randomised controlled trials (1999)

-

In 1996, Authors with different backgrounds summoned a conference in order to address issues on present reporting standard of meta-analysis and systematic reviews and need for improvement. The conference resulted in the production of an instrument termed Quality Of Reporting Of Meta-analysis (QUOROM) which consists of an 18-items checklist. The checklist put through the best way to present the abstract, introduction, methods, result and discussion. The checklist was published in 1999 by Moher et al [65] .

Criteria for improving the quality of reporting by Moher et al(1999) [65] are Title
Abstract

- *Objectives*
- *Data sources*
- *Reviews methods*
- *Result*
- *Conclusion*

  Introduction

  Methods

- *Searching*
- *Selection*
- *Validity assessment*
- *Data abstraction*
- *Study characteristics*
- *Quantitative data synthesis*

  Results

- *Trial flow*
- *Study characteristics*
- *Quantitative data synthesis*

discussion

2.3.4 (MOOSE)-Meta-Analysis of Observational Studies in Epidemiology(2000)

Blettner et al [66] were concerned about inherent biases in the original studies. Similarly, the intense diversity of study designs and populations in epidemiology makes the interpretation of simple summaries problematic at best. Consequently, methodologic and reporting issues,such as publication bias, could have an impact when combining results of observational studies [67],[68]. In order to address this challenge, a workshop was convened in 1997 and guideline regarding reporting of Meta-analysis Of Observational Study in Epidemiology (MOOSE) was established to help understand and measure sources of differences in results across studies [69].

A proposed Reporting Checklist of observational studies(MOOSE) by Stroup et al(2000) [67] Reporting of Background

- *Hypothesis definition, description of study outcomes,*
- *Study population and designs used.*

  Reporting of search strategy
- *of searchers, search strategy,*
- *Databases and registries searched*

  Reporting of methods

- *Description of relevance of studies, Rationale for the selection and coding of data,*
- *Documentation of how data were classified and coded,*
- *Assessment of study quality, provision of appropriate tables and graphics*

  Reporting of results
- *Graphic summarizing individual study estimates and overall estimate,*
- *Table giving descriptive information for each study included,*
- *indication of statistical uncertainty of findings*

  Reporting of discussions
- *Quantitative assessment of bias, justification for exclusion,*
- *Generation of the conclusions, Guidelines for future research,*
- *Disclosure of funding source*

2.3.5 (CONSORT)-CONsolidated Standard Of Reporting Trials(2001)

The quality of reporting is very important in the medical literature reviews as it provides the clinicians, experts in medical fields complete and valuable information and guideline to their daily clinical practice. When there is little or no detailed information in literature reviews, researchers and potential users would lack knowledge and revelation of fundamental content of the reviews and the current clinical issues that would be of help in healthcare services and assessment of the primary studies will be difficult.

Moher et al[65] assert that a lack of sufficiently reported randomization has been associated with bias in evaluating the effectiveness of interventions. To assess the strengths and limitations of randomised controlled trials, readers need and deserve to know the quality of its methodology.

Given this limitation on quality of reporting in clinical trials, experts comprising of medical journal editors, clinical trialists, epidemiologists and methodologists organised a workshop in 1993 with the aim of developing new instrument that can assess the quality of reporting on randomised controlled trials . This workshop gave birth to a checklist in 1996. But still, considerable evidence revealed that only one item in the checklist addressed the reporting of safety and suggests that reporting on harm-related issues needs improvement.

In 1999, another meeting was held in order to revise and modify the original CONSORT checklist with intention to standardize, improve and publish reporting of randomised controlled trials and it was actualized in 2001. It consists of a 22-item checklist. The Table below shows the revised checklist.

Revised Checklist for CONSORT (CONsolidated Standard Of Reporting Trials) by Moher et al[65]

Paper Section and Topic Title and abstract Introduction

- *Background*

  Methods

- *Participants*

- *Interventions.*

- *Objectives*

- *outcome*

- *Sample size*

  Randomization

- *Sequence generalization*

- *Allocation concealment*

- *Implementation*
- *Blinding (masking)*
- *Statistical methods*

  Results
- *Participant flow*
- *Recruitment*
- *Baseline data*
- *Numbers analyzed*
- *Outcome and estimation*
- *Ancillary analyses*
- *Adverse events*

  Discussion
- *Interpretation*
- *Generalizability*
- *Overall evidence*

2.3.6 (REMARK)-REporting recommendations for tumor MARKer prognostic studies(2005)

This was a case study of research reports on tumour makers in oncology where several studies of the same or related makers produced inconsistent conclusions or stand in direct contradiction to the promising results [70]. Attempts have been made to identify reasons why multiple

studies of the same makers can yield different conclusions. Diversity of methodological problems have been cited to solve this differences but the reports obtained from study of tumour makers were not carried out in a rigorous manner and the published articles usually lack necessary information that can be of help for adequate assessment of study quality . The first international conference meeting on Cancer diagnostics was held in 2000 and there was a recommendation on development of guidelines for the reporting of tumour makers in order to encourage transparency and complete reporting so that important information should be available for clinicians and other potential users [71]. This led to the production of the REMARK checklist given below.

Reporting recommendations for Tumor Maker prognostic studies (REMARK) by McShane et al [70]

INTRODUCTION

MATERIALS AND METHODS

- *Patients*
- *Specimen characteristics*
- *Assay methods*
- *Study design*
- *Statistical analysis methods*

  RESULTS

- *Data*
- *Analysis and presentation*

  .

DISCUSSION

2.3.7 STARLITE (Sampling strategy,Type of study,Approaches,Range of years,Limits,Inclusion and exclusions,Terms used,Electronic sources)(2006)

The use of checklists make it easier for readers to assess the quality of reviews and for researcher to replicate their methods[72]. These checklists were developed to improve the quality of reporting, methodological quality and scientific quality of quantitative systematic reviews, but Booth [73] stated that no standard have been published for reporting literature searches, the systematic reviews of qualitative research are limited by poor quality of reporting of search method and that the criteria for reporting literature search must acknowledge the demands of both qualitative and quantitative systematic reviews.

They designed structure for reporting the quality of literature searches based on the empirical findings from their reviews which conveyed using the mnemonic STARLITE (Standards for Reporting Literature searches).

Representing the elements of STARLITE by Booth (2006) [73]
Element

- S: *Sampling strategy*
- T: *Type of Studies*
- A: *Approaches*
- R: *Range of years (start date-end date)*
- L: *Limits*
- I: *Inclusion and exclusions*
- T: *terms Used*
- E: *Electronic sources*

2.3.8 (AMSTAR)Assessment of Multiple Systematic Reviews(2007)

This instrument was developed from the combination of items drawn from two available checklists- constructed by Oxman et al [74] and Sacks et al[62]. Validated checklist was developed to measure the methodological quality of systematic reviews. Again, more item were added in the checklist to determine if the language restriction and publication bias were applied in selecting studies for systematic reviews [75]. The checklist was designed with 31-item assessment tools. The Table shows the checklist.

AMSTAR was developed by Shea et al

- *Was an 'a priori' design provided?*
- *Was there duplicate study selection and data extraction?*
- *Was a comprehensive literature search performed?*
- *Was the status of publication(i.e, grey literature) used as an inclusion criteria?*
- *Was a list of studies (included and extracted) provided?*
- *Were the characteristics of the included studies provided?*
- *Was the scientific quality of included studies assessed and documented?*
- *Was the scientific quality of the included studies used appropriately in formulating conclusions?*
- *were the methods used to combine the findings of studies appropriate?*
- *Was the likelihood of publication bias assessed?*
- *Was the conflict of interest stated?*

2.3.9 The PRISMA Checklist(2009)

Moher et al who developed The QUOROM statement (Quality Of Reporting Of Meta-analyses), which focused on the reporting of meta-analyses of randomised controlled trials was revised, expanded and renamed (PRISMA) "Preferred Reporting Items for Systematic reviews and Meta-Analyses"in 2009. The aims- to address numerous conceptual and practical advances in the science of systematic reviews and used as a basis for reporting systematic reviews of other types of research, especially evaluations of interventions [76]. It consists of a 27-items checklist. The table below shows the checklist of items included when reporting meta-analysis or systematic reviews.

The PRISMA Statement developed by Moher et al [76].

Section/topic

TITLE

- *Title*

  ABSTRACT

- *Structure summary*

  INTRODUCTION

- *Rationale*

- *Objectives*

  METHODS

- *Protocol and registration*

- *Eligibility criteria*

- *Information sources*

- *Search*

- *Study selection*
- *Data collection process*
- *Data items*
- *Risk of bias in individual studies*
- *Summary measures*
- *Synthesis of results*
- *Risk of bias across studies*
- *Additional analyses*

RESULTS

- *Study selection*
- *Study characteristics*
- *Risk of bias within studies*
- *Results of individual studies*
- *Syntheses of results*
- *Risk of bias across studies*
- *Additional analyses*

DISCUSSION

- *Summary of evidence*
- *Limitations*

- *Conclusions*

  FUNDING

- *Funding*

In healthcare systems, there is an increase in health care decisions made on research-base evidence rather than clinical experience and expert opinions. Meta-analyses were formulated to represent the rigorous procedure of compiling scientific research evidence to solve problems concerning healthcare issues on treatment of patients, its interventions and remedy to the services. However, clinicians and medical experts need to select and read medical literature reviews of related topic under study in order to be abreast and up-to-date with health care issues. The review articles under study are being published continuously at high rate. Meta-analyses and systematic reviews attempt to minimize confounding bias by comprehensive and reproducibility of search for, article selection for review, and methodological quality.

Criteria were outlined as a guideline and were used to prepare a checklist that can be used to evacuate all the factors affecting selection and inclusion of studies in meta-analysis.

There are about 25 checklists developed so far from authors of different background of meta-analyses to tackle the problems of poor face validity of study quality. They organised and convened workshops, conferences and meetings at difference places to suggest, recommend and propose a validated instruments that can measure the trials' quality based on meta-analysis, systematic reviews and clinical studies

From literature reviews of nine validated checklists listed above, we discovered that most checklists were only focused on methodological quality, reporting quality, scientific quality and inherent bias as criteria for quality assessment in meta-analysis and clinical study. Less emphasis was placed on statistical methodological quality. Given these facts, we propose to design and develop a checklist that can assess the quality of meta-analysis in terms of actual statistical analysis.

2.3.10 The summary of historical checklist

It appears that Mahon and Daniel [55] were the first to notify that there was a lack of rigour during literature search in clinical trials. In 1964, they developed a checklist as a guide to obtaining accurate reports. Between the year 1985/86, Mulrow [30] and Sack et al [62] commented during their reviews on medical research that there was inconsistency in the use of scientific method to assess healthcare information. This alertness of poor scientific methods drew attention of Oxman and Guyatt[31] in 1989 and led to publication of a checklist (OQAQ) that measures scientific quality of reviews in 1991. Jadad et al(1996)[63] suggested that there might be a potential risk of bias in meta-analysis when using this checklist without blinding the raters. Consequently, they introduced their own checklist (Blind Assessment of the Quality of Trials Reports). Moher et al (1999)[65] had noticed that the present standard of reporting of meta-analysis of randomised controlled trials was not sufficient and that it needs improvement. They developed checklist called QUOROM .

Blettner et al (2000)[66] were much concerned on the methodological issues related to meta-analysis such as publication bias and thus they developed guideline (MOOSE). Moher et al (2001) [65] designed checklist (CONsolidated Standard Of Reporting Trials) to examine the quality of reporting in randomised controlled trials. There was a report that multiple studies of the same markers gave different conclusions as a result of lack of necessary information or not carried out in a rigorous manner. In order to encourage transparency, McShane et al (2005) [70] designed a checklist, REMARK, to serve as a guideline for the reporting tumour markers. Booth et al (2006)[73] complained that the present checklists failed to highlight the criteria / standard for reporting literature search of qualitative and quantitative systematic reviews and they proposed a framework based on the experimental findings from the reviews (STARLITE). Oxman et al [74] and Sack et (2007) [75] were apprehensive that language restriction and publication bias might be a potential risk that will affect the assessment of the methodological quality of systematic reviews and thereby designed a vali-

dated checklist (AMSTAR). Finally, as a result of continuous insufficient quality of reporting in the present checklist, QUOROM statement was revised and renamed PRISMA by Moher et al (2009) [76] in order to addressed more the various formations and practical advances in the science of systematic reviews and also used as a base for reporting of evaluations of interventions.

Chapter 3

DEVELOPMENT OF CHECKLIST

We developed a new checklist to assess the quality of meta-analysis in terms of actual statistical analysis. The current checklists do not alert users of potential problems and rarely have any item about the statistical analysis. Our checklist consists of 21 items and was designed to assess the methodological quality, reporting quality and statistical quality of studies. A pilot study was also carried out to evaluate the effectiveness of the new checklist using some randomly selected studies.

We articulated the following checklist for the assessment of quality. Item 1 to item 8 were obtained from Oxman et al(1996)[31]

1. Is the objective/aim/hypothesis of the study clearly described?
2. Are the main outcomes to be measured clearly described in the introduction or methods section?
3. Were comprehensive search methods used to locate relevant studies?
4. Was the validity of the primary studies assessed?
5. Were the inclusion criteria reported?

6. Were explicit methods used to determine which articles to include in the review?
7. Was bias in the selection of studies avoided?
8. Were the criteria used for assessing the validity of the included studies reported?
9. Are the statistical methods used in analysing the data specified in detail?
10. Are the statistical methods appropriate?
11. Are assumptions underlying the statistical tools met by the data?
12. Is the statistical method used in combining studies in harmony with the design of the clinical trials?
13. Do studies being combined have adequate information to merit inclusion in meta-analysis?
14. Is the measure of precision available (or can it be derived) for each study being combined?
15. Is there double-counting of studies?
16. Are studies being combined compatible?
17. Are appropriate weights used in combining studies?
18. Is the Statistical software used in obtaining the results stated?
19. Are the actual probability values reported?
20. Are conclusions based on hypothesis test valid?
21. Can results be reproduced by an independent checker

In order to assess the workability of the instrument, we need to select some studies from published sources. To do this, we first determine our population of interest and define the sampling plan and inclusion criteria. Cochrane and non-Cochrane publications published between 2000 and 2008 on meta-analysis of RCTs were chosen as the populations to be investigated. We have outlined below our inclusion criteria and sampling procedure.

3.1 INCLUSION CRITERIA

All meta-analyses of randomised controlled trials published from 2000 to 2008 in each database. Also, Meta-analysis of RCT's published in full text, in English language and journal article were included. Studies that preformed a comprehensive search of literature, and clearly explain the statistical methods used. Studies published in Cochrane database of systematic reviews, Medline and Web of knowledge database that state the characteristics of included studies. We excluded observational, cohort and case control studies, meta-analysis of discriminant capacity trials and studies without statistical analysis.

3.2 ESTIMATION OF SAMPLE SIZE

Our interest is on published articles on meta-analysis of randomised controlled trials contained in the Cochrane database of systematic reviews and paper-base journal databases published between the period of 2000 and 2008. The first approach is to establish a sample size that would guarantee adequate error bounds on estimates of proportions obtained from samples.

The variance of a proportion attains its maximum when $p = 0.5$. A conservative approach is to choose the value of n that will guarantee a sufficiently narrow error bound for the worst case ($p = 0.5$).

On the other hand one has to consider resource constraints. From the pilot survey we conducted, about 25 papers could be reviewed per week. On average, it may be possible to assess about 100 papers in a month. Using this number of samples and the proportion of publications on meta-analysis from the database $p = 0.5$, the estimated maximum error bound is about 0.1.

3.3 SAMPLING PLAN

A check on Cochrane database of systematic reviews, web of knowledge and Medline database of reviews indicate valid number of publications on meta-analyses between the period of 2000 - 2008. We estimated the sample size to be 100 and allocated 50 papers each to Cochrane and non-Cochrane database (Web of Knowledge and Medline). The 50 papers allocated to the non-Cochrane database were distributed proportionally to Web of Knowledge and Medline.

3.3.1 Sampling Frame

All eligible studies published in the Cochrane and non-Cochrane libraries on Meta-analysis of Randomised Controlled Trials between 2000 and 2008 constitute our target population.

A search through the Electronic databases showed that in Cochrane library, there are 4646 results out of 5821 records for Meta-analysis of Randomised Controlled Trials published between 2000 and 2008 in Cochrane database of systematic reviews. In non-Cochrane library there are 1434 results of Meta-analysis of RCT's published in database of Web of knowledge of Figure 1.1 and 135 results of Meta-analysis of RCT's published in database of Medline within the same period.

3.3.2 Studies meeting Inclusion Criteria

Further comprehensive search was performed for studies on meta-analysis of Randomised Controlled Trials published in ENGLISH, FULL TEXT and are JOURNAL articles. Search results were obtained directly from electronic database of Cochrane, Web of Knowledge and Medline as shown below.

Cochrane Library There were 515 results out of 5821 records for: "Meta-analysis of Randomised Controlled Trials* OR Metaanalysis of randomised controlled trials* or Meta analysis of Randomised Controlled Trials* published in English language and full text and journal article, from 2000 to 2008 in Cochrane Database of Systematic Reviews"

‘

Web of Knowledge and Medline There were 507 and 130 results of Meta-analysis of Randomised Controlled Trials OR Metaanalysis of Randomised Controlled Trials OR Meta analysis of Randomised Controlled Trials published in full text, in English and journal published in Web of Knowledge and Medline databases respectively, Timespan = 2000-2008

3.3.3 Choice of database

The choice of Cochrane Library and Medline databases for this research is informed by the fact that they contain the most up to date, peer reviewed, high quality, and strong scientific-based journals published in meta-analysis. Medline is a powerful tool to efficiently access the voluminous amount of medical literature. It is a premier database of biomedicine and health sciences [77]. The use of Web of Knowledge is to enable us have access to scientific journals published elsewhere outside the Cochrane library and Medline. Web of Knowledge provides easy links to full texts and has analysis tool that helps to refine searches by authors, publication year, source title or subject category [77]. It makes it easy to spot developing trends which give insight into tracing the history of particular field of study.

3.3.4 Sample Size Allocation

Basically, the estimated sample size of 100 were equally distributed among Cochrane and non-Cochrane (Medline and Web of Knowledge) libraries. Fifty papers were allocated to Cochrane database and the other fifty papers were allocated proportionally to Medline and Web of Knowledge.

The proportional allocation for the sample size of 50 papers between Web of Knowledge and Medline is calculated below as;

Web of Knowledge (WOK) = 507 results of meta-analysis of RCT's meeting the inclusion criteria.

Medline = 130 results of meta-analysis of RCT's meeting the inclusion criteria

These give a total of $507 + 130 = 637$

To obtain the actual sample size for WOK and Medline, the proportion of each is calculated as shown below

$WOK = 507/637 * 50 = 40$

$Medline = 130/637 * 50 = 10$

This shows that the sample size of 40 papers is allocated to Web of Knowledge database and sample size of 10 papers to Medline database. From each database, random sample of publications were selected that met the inclusion criteria until the sample size is attained.

Fig 3.1 reveals the flowchart of the procedure used to arrive at selected studies from Cochrane and Non-Cochrane database. 17 studies were excluded as they did not meet the inclusion criteria. We included 42 eligible studies obtained from Cochrane database, 31 and 10 eligible studies obtained from Web of knowledge and Medline respectively.

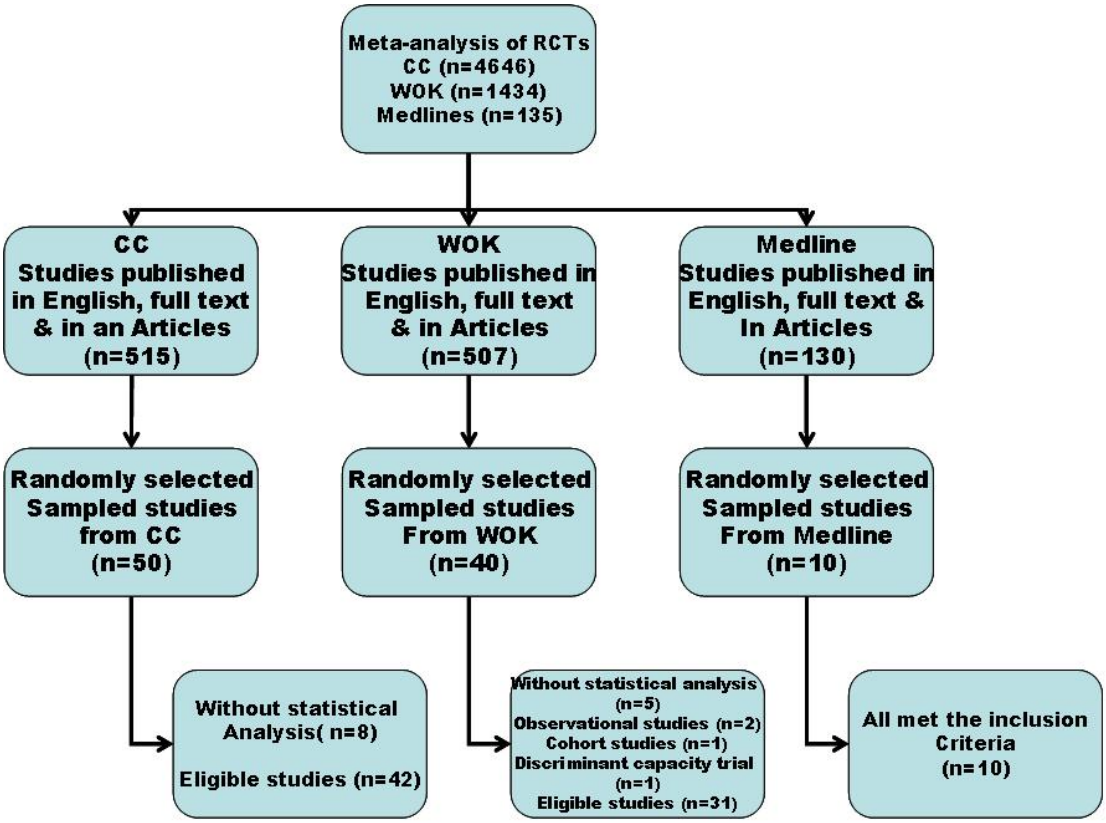


Figure 3.1: Flowchart of studies: Search for Meta-analysis and reasons for inclusion.

3.3.5 Sample selection from Databases

Studies were searched using field labels. Example, “Metaanalysis of RCTs* OR “Meta-analysis of RCTs* OR “Meta analysis of RCTs* published in English, full text and Journal articles. The search was restricted to Cochrane database of Systematic Reviews and date ranging from 2000 to 2008, likewise in Web of Knowledge and Medline. The actual number of search results appeared along with the long list of published studies consisting of record title, name of the author(s), year of publication, source, volume, issue and pages. Search results obtained from the Cochrane Library website were subsequently assigned with serial numbers. Then, using the random numbers generated in R package in the previous section, we selected the study that has the corresponding number. We then downloaded and saved it for further study.

Chapter 4

APPLICATION OF THE NEW CHECKLIST TO SELECTED STUDIES

4.1 PILOT STUDY

To assess the quality of meta-analysis in terms of actual statistical analysis and methodology quality, a pilot study was first conducted to verify the completeness and reliability of the instrument designed in the previous chapter. Studies were retrieved from the Cochrane database and Paper-based journals (Web of Knowledge and Medline) published between 2000 and 2008. We assessed 17 papers from a group of 50 papers randomly selected from the Cochrane library, 12 trials from group of 40 studies randomly selected from Web of Knowledge and 10 studies from Medline using our newly developed 21-item checklist. This enabled us to assess the completeness, adequacy and the ability of the instrument to capture the salient information needed to effectively assess studies. This led to a slight modification of the instrument.

4.2 MODIFICATION OF CHECKLIST

The initial 21-item checklist was subjected to a pilot study. This exposed the need to examine more thoroughly the content of the checklist and also concentrate more on the statistical quality aspects of the reviews. Consequently, we made some modifications on the checklist by generating more items and expressed them in more explicit form. We added four extra items (as listed below) and deleted item 10-(Are the statistical methods appropriate?) from the checklist. We now have an instrument with 24 items, (as shown in Figure 4.1), that has the potential to assess the statistical and overall quality of studies.

The four items added to the checklist is listed below.

- Was the aim of using the statistical method stated?
- Was the use of statistical methods justified?
- Were the study design of individual trials stated in meta-analysis? E.g Parallel, crossover, factorial etc.
- Are there double counting of some aspects of studies?

4.3 RELIABILITY CHECK

We propose to evaluate the effectiveness of the modified instrument and use it to measure the quality of meta-analysis in terms of actual statistical analysis. A re-assessment of the statistical and methodological quality of all the reviewed study showed that it was easier, faster and less time consuming using the modified checklist.

The reliability of the 24-item checklist was established by assessing all the studies in two separate times. Results obtained in the two assessments were identical and highly consistent. Cronback's alpha was used to check the internal consistency of the checklist. This yielded

Checklists for measuring Quality of Meta-analysis in terms of actual Statistical Analysis	
1	Is the objective/aim/hypothesis of the study clearly described?
2	Are the main outcomes to be measured clearly described in the introduction or methods section?
3	Were comprehensive search methods used to locate relevant studies?
4	Was the validity of the primary studies assessed?
5	Were the inclusion criteria reported?
6	Were explicit methods used to determine which articles to include in the review? Randomization , double blinding
7	Were the Authors of unpublished data contacted?
8	Was bias in the selection of studies avoided?
9	Were the criteria used for assessing the validity of the included studies reported? Eg QUOROM,QOAQ
10	Were the statistical methods used in analysing the data specified?
11	Was the aim of using the statistical method stated?
12	Was the use of the staistical method justified?
13	Were the group designs of individual trials stated in the meta-analysis? Eg Parallel,Crossover,Factorial & others
14	Were studies being combined compatible? Eg in units, designs, dosages,ages
15	Is the statistical method used in combining studies in harmony with the design of the clinical trials?
16	Do studies being combined have adequate information to merit inclusion in meta-analysis?
17	Is the measure of precision available (or can be derived) for each study being combined?
18	Are there double-counting of studies?
19	Are there double counting of some aspects of studies?
20	Are appropriate weights used in combining studies?
21	Is the Statistical software used in obtaining the results stated?
22	Are the actual probability values reported?
23	Are conclusions based on hypothesis test valid?
24	Can results be reproduced by an independent checker?

Fig 4.1 Modified Checklist

very high correlation value of 0.92. This gave us the confidence to apply our checklist to some selected studies.

4.4 ASSESSMENT OF STUDIES

All the 83 studies were assessed using the modified checklist. Needed information were then extracted and analysed.

4.4.1 Data Extraction

Data for analysis were extracted by assessing the quality of studies in the Cochrane and Journal reviews using the modified checklist. The studies from Non-Cochrane library (Web of Knowledge and Medline) were merged together to give a single group.

From each trial, data were extracted on:

- Meta-analysis of randomised controlled trials
- Full detail of the background of study
- The comprehensive search strategy
- Stating the explicit methods used to identify trials for reviews
- The statistical method used, the aim of using it and its justification
- The reporting of statistical software and actual Probability value
- Whether studies and some aspects of studies were double counted.
- Avoidance of Selection bias and publication bias
- The type of designs used for each study and how they were combined

- Validation of primary study assessment
- Whether inclusion criteria were stated and method of determining included articles

We used the statistical package SAS 9.1 version and R package to perform the Fisher Exact Test in order to determine if there are non-random associations between the databases [78]. Some sample sizes were small with expected cell frequencies less than five in some cases. We judged that it was unsafe to rely on the chi-squared test and used Fisher's exact test instead. Two-sided Fisher's exact test was used to analyse the responses to the 24 items in the checklist as shown in Table 4.2. We computed the proportions that describe the statistical quality between Cochrane and Paper-base journals. Each item is regarded as being statistically significant if the P-value < 0.05. There is no adjustment for multiple testing therefore results should be interpreted cautiously in the light of this.

An example of the SAS procedure used in carrying out the Fisher Exact test is given below: A=31, B=23 C=11 D=18 data Qua;

input $ paper $ objective count;

datalines;

cc	yes	31
cc	no	11
jn	yes	23
jn	no	18

run;

$proc freq data = Qua$;

tables paper*objective / fisher;

weight count;

title 'quality of meta-analysis in terms of actual statistical analysis';

run;

paper Frequency Percent Row Pct Col Pct	objective no	 yes	 Total
Cochran	11	31	42
	13.25	37.35	50.60
	26.19	73.81	
	37.93	57.41	
Journal	18	23	41
	21.69	27.71	49.40
	43.90	56.10	
	62.07	42.59	
Total	29	54	83
	34.94	65.06	100.00

Table 4.1: Fisher's Exact Test

Quality of meta-analysis in terms of actual statistical analysis

The FREQ Procedure

Table of paper by objective

Left-sided $Pr = 0.0716$,Right-sided $Pr = 0.9731$, Table Probability $(P) = 0.0448$, Two-sided Pr $\leq$P = 0.1102, Sample Size = 83

Example of R procedure in fisher exact test is

database =matrix (c(42,41,0,0),nr=2,dimnames=list(database=c("cochrane","journal"), outcomes=c("yes","no"))) fisher.test(database)

Fisher's Exact Test for Count Data data: database

p-value = 1

alternative hypothesis: true odds ratio is not equal to 1

95 percent confidence interval:

The numbers of studies reviewed	All reviews ($n = 83$)	CC Reviews ($n = 42$)	Paper Reviews ($n = 41$)	P-values P<0.05
1 Objective/aim/hypothesis clearly described?	83	42(100%)	41(100%)	1.0
2 Main outcomes clearly described?	83	42(100%)	41 (100%)	1.0
3 Comprehensive search methods used	83	42(100%)	41 (100%)	1.0
4 Validity of the primary studies assessed?	79	42(100%)	37 (90%)	0.055
5 Inclusion criteria reported?	80	42(100%)	38 (93%)	0.12
6 Explicit methods used to determine articles?	79	42(100%)	39 (95%)	0.24
7 Authors of unpublished data contacted?	54	31(74%)	23 (56%)	0.11
8 Bias in the selection of studies avoided?	80	42(100%)	38 (93%)	0.12
9 Criteria for validity assessment reported?	11	1(2.4%)	10 (24%)	0.003
10 Statistical methods used specified?	82	42(100%)	40 (98%)	0.31
11 Aim of using the statistical method stated?	76	38(90%)	37 (90%)	1.0
12 Use of the statistical method justified?	66	34(81%)	32 (78%)	0.89
13 Design of individual trials stated ?	20	13(31%)	7 (17%)	0.2
14 Studies being combined compatible?	22	12(29%)	10 (24%)	0.8
15 Statistical method agree with the study design?	19	12(29%)	7 (17%)	0.3
16 Studies combined have adequate information ?	76	39(93%)	37(90%)	0.71
17 Measure of precision available?	81	41(98%)	40 (98%)	1.0
18 Is there double counting of studies?	0	0	0	1.0
19 Double counting of some aspects of studies?	4	4(10%)	0 (0)	0.24
20 Appropriate weights used in combining studies?	72	39(93%)	33 (80%)	0.12
21 Software used in obtaining the results stated?	58	31(74%)	27 (66%)	0.48
22 Are the actual probability values reported?	76	40(95%)	36 (88%)	0.27
23 valid conclusions based on hypothesis test?	83	42(100%)	41 (100%)	1.0
24 Results reproducible by independent checker?	78	40 (95%)	38 (93%)	0.68

Table 4.2: Checklists for measuring Quality of Meta-analysis in terms of actual Statistical Analysis

0 Inf

sample estimates:

odds ratio = 0

Studies	Cochrane	Journal
1 - 5	9	8
6 - 10	9	18
11 - 15	7	4
16 - 20	4	3
21 - 25	4	2
26 - 30	5	1
31 - 35	2	1
36 - 40	1	0
41 - 45	0	2
46 - 50	1	0
51+	0	2
TOTAL	42	41

Table 4.3: Number of Trials in each meta-analysis

4.5 RESULTS

4.5.1 Methodological Characteristics

The number of authors involved in conducting meta-analysis of randomised controlled trials ranged from 2 to 21. However, the number of trials in each meta-analysis included in the reviews varied from 2 to 52. Most of the reviews included were meta-analysis on randomised controlled trial. We evaluated the number of included trials and authors of meta-analysis obtained from the Cochrane and Journal database using R package to plot Stem and Leaf . We also represent the number of trial and authors of Cochrane and Paper-base Journal reviews in a frequency table as shown in Tables 4.3 and Table 4.4.

4.5.2 Stem and leaf plot

According to Weisstein[79], a stem and leaf plot is a diagram that summaries the data while maintaining the individual data point . It is used to show the shape and distribution of numbers. Stem is the digit in the tens place and leaf is the number in the ones place.

The basic command use in R is

Authors	Cochrane	Journal
1 - 3	23	13
4 - 6	15	20
7 - 9	3	4
10 - 12	1	0
13 - 15	0	1
16 - 18	0	2
19 - 21	0	1
TOTAL	42	41

Table 4.4: Number of Authors

stem(Journal.Author, scale=2)

stem(Cochrane.Author)

Figure 4.2: Number of Trials in the selected Cochrane

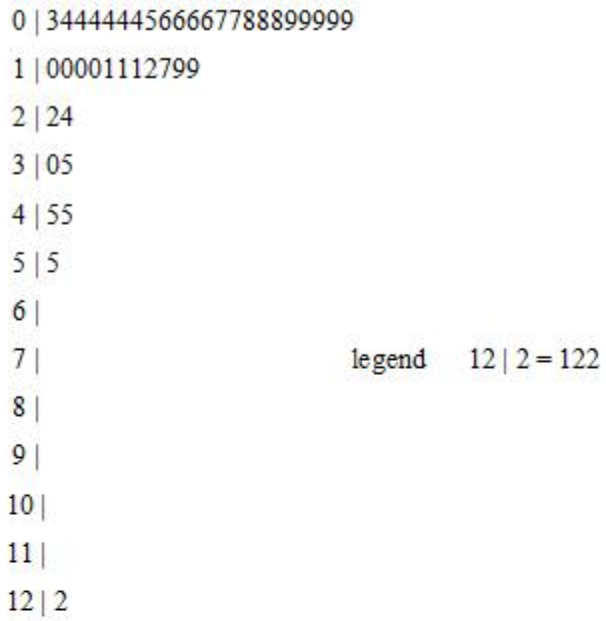
No of trials in Cochrane

0 | 22223344
0 | 556778999
1 | 00122344
1 | 568
2 | 01114
2 | 66688
3 | 14
3 |
4 | 0
4 | 7

Legend
3 | 14 = 31, 34

We used Stem and Leaf plot to demonstrate the distribution of the number of authors and the trials in the reviewed studies from Cochrane and Non-Cochrane databases as shown in Fig 4.2,4.3, 4.4 and 4.5 which display positive skewness. This is more prominent in the non-Cochrane journals. The paper-base journal has the highest number of authors and trials compared with Cochrane library.

Figure 4.3: Number of Trials in the selected non-Cochrane studies

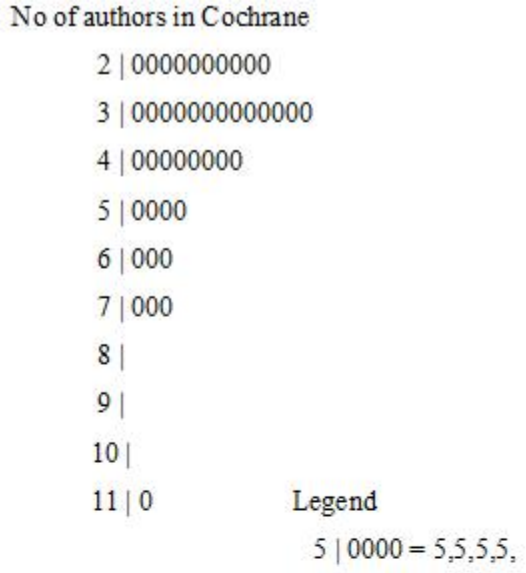


There were no statistically significant difference between the Cochrane Collaboration and Paper-Base journal in most of the items except for item number 9 with $p = 0.003$. This item showed that the paper based journals more often use the instrument developed by Oxman et al (1988)[80] to assess the validity of included studies than the Cochrane Collaboration based papers.

4.5.3 Methodological Quality

From our review, we found that both Cochrane and paper-base journals adequately reported the objectives/hypothesis, described the main outcomes in the introduction and reported comprehensive search methods. This finding may be indicative of the fact that most authors follow the stipulations in existing checklists. This is an important development in that reviewers can have a bird's eye view of a publication without having to go through the whole publication. Thus, selection of eligible studies for inclusion in reviews is enhanced and faster because decision for inclusion or otherwise can be reached by reading a well-detailed abstract.

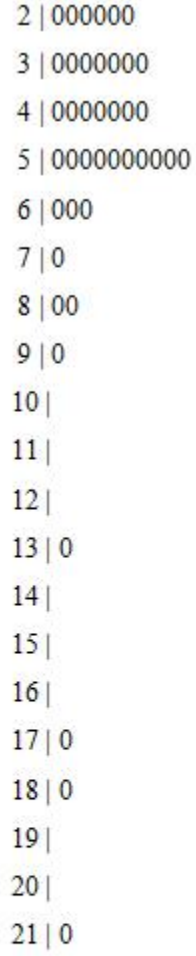
Figure 4.4: Number of Authors in the selected Cochrane



From our investigation, Cochrane is more likely (74%) to contact authors of unpublished articles than paper-base journals (56%). Also the Cochrane groups are slightly more rigorous 42(100%) in stating steps taken to avoid selection bias compared with 38(93%) of the paper-base journals.

We selected two checklists - Overview Quality Assessment Questionnaire (OQAQ) and Quality Of Reporting Of Meta-analysis (QUOROM) which we included in our checklist as guide for assessing the scientific quality, methodological quality and reporting quality. Surprisingly, only 1(2.4%) study in Cochrane review explicitly stated the use of QUOROM or OQAQ to assess quality against 10(24%) of paper-base journals that explicitly stated the use of the checklists. The results showed a statistical significant difference with P = 0.003. However, further investigation indicate that Cochrane library prefer the use of checklists prepared by Cochrane Renal Group, Cochrane approach - Jadad scale (1996) and other Cochrane Review checklists related to interventions under investigation.

Figure 4.5: Number of Authors in the selected non-Cochrane studies



Study design of a clinical trial could be parallel, crossover or factorial where participants are randomized or blinded to given treatments over a period of time, and outcomes are measured. In meta-analysis, statistical techniques used to combine studies should be consistent with the design. Thus, we advocate that meta-analysts should state the design of each study included in analysis in order to prevent a mismatch of studies.

For instance in Cochrane review, the meta-analysis on tropical antibiotics without steroids for chronically discharging ears by Macfadyne et al (2005 [81] reported the study designs of individual trials included in the study. He had 14 trials - 13 parallel groups and 1 crossover group. He went further to state that he used results of the first period before participants were crossed over to the alternative treatment. One other author combined several study designs which includes parallel, crossover, comparative, prospective study and randomised controlled trials. Nevertheless, from our assessment, Cochrane reviews were more likely 31% than paper-base journal 17% to report the types of study designs that were combined in the meta-analysis.

It was difficult to adequately assess items 14 and 15 of the checklist which has to do with compatibility of combined studies given their individual design and if the statistical method used in combining the studies is in harmony with the study design. This is because of the failure of most authors to explicitly state the design of individual studies combined in meta-analysis (item 13). There is need for authors to specify the design of individual studies combined. The consequence of combining studies with different design background is that interpretation of simple summaries becomes problematic. For instance, combining (a) parallel study with crossover study, (b) prevalence study with incidence study, (c) double with single blinded, (d) retrospective study with prospective study etc, may yield confusing and misleading results.

The presence of statistical heterogeneity, clinical and methodological diversity could impact on the compatibility of studies. Therefore, combining studies with different measured

effects or studies with widely differing participants, interventions and outcomes may lead to misleading conclusions. We have included the item on compatibility of studies into our checklist to stimulate and provoke deep consideration of this all important issue among meta-analysts. There is every need to reflect on the design of individual studies being combined and ensure that only compatible studies are used in the same meta-analysis.

Assessment of adequacy of statistical methods is easier when study designs are clearly stated. Where information on study design is lacking, appraisal of appropriateness of statistical method is hampered.

This checklist is designed with the aim of encouraging study authors to include as much statistical information as is relevant in the published works. Rich statistical information content of individual studies makes easy the work of the meta-analyst in the identification and combination of compatible studies.

Given that our interest is to assess the quality of meta-analysis in terms of actual statistical analysis, studies were reviewed to find out whether the statistical methods used were specified, whether the aim of using it was stated and if the use of the method was justified. Most authors that published their articles in Cochrane library mostly used statistical methods such as odd ratio, relative risk and weighted mean deviation. In Cochrane 34(81%) justified the methods used. In journals reviews, 98% reported the statistical methods used, 37(90%) stated the aim and 78% stated the justification of the methods used.

Sample size is the main factor in determining the weight for a trial. The larger the sample size the more weight assigned and smaller the variance. Likewise, less weight is assigned to the smaller sample size and with large variance. We identified that Cochrane reviews had 39(93%) that reported appropriate weight used to combine studies, 33(80%) reported in journals with p=0.2

There is a need to report the actual probability values, rather than only saying that a result is or is not statistically significant, in this way the reader can compare the probability

value to the significance level[82]. We found that Cochrane reviews are more likely to state the actual probability value and/or confidence intervals in their analyses than journals. We identified 40(95%) and 36(88%) studies in Cochrane and journal reviews respectively that reported the actual probability values and/or confidence interval in the meta-analysis. However, most of these analyses (9 out of 10) used confidence intervals and needed not state the p-value.

Cochrane authors mostly used different versions of RevMan Statistical software in data analysis. We noticed that paper-base journal did not only use RevMan statistical software but also used other softwares like Statdirect, Statview, Micro strata and winBUGS. We obtained 31(74%) studies reporting the use of statistical software in Cochrane and 27(66%) in the journals.

Finally, we identified that the possibility of studies reviewed being reproduced by potential users is 95% and 93% in Cochrane database and journal database respectively. Also, both databases made valid conclusions that was based on hypothesis test.

4.5.4 Double counting of some aspects of studies

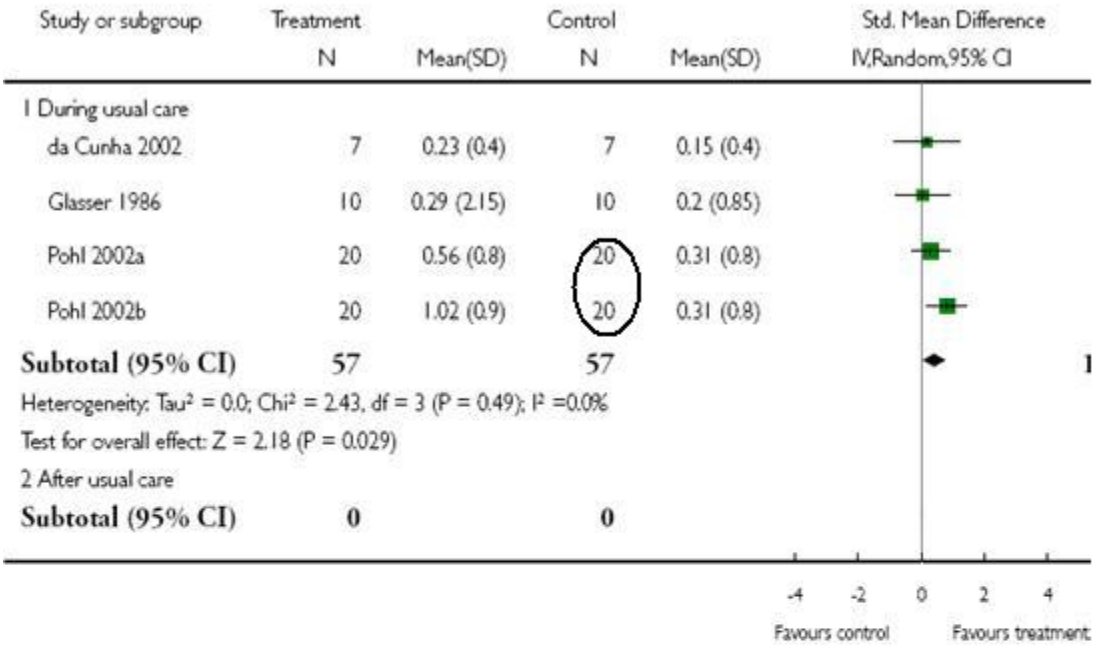
There were no simple double counting of studies observed during the assessment in Cochrane and Non-Cochrane databases, but we identified 4(10%) meta-analyses that contain double counting of some aspects of studies in Cochrane reviews.

Here are the examples

Example 1: Figure 4.6 shows the evidence of double counting of some aspects of studies identified in meta-analysis from Cochrane reviews. Saunder et al[83] in their meta-analysis on the effect of physical fitness training on dependency, death and disability after stroke, included a trial published by Pohl [84] who included two separate treadmill training intervention groups sharing the same control group. In this case, control arm was counted twice

in the analysis of meta-analysis conducted by Sander et al.

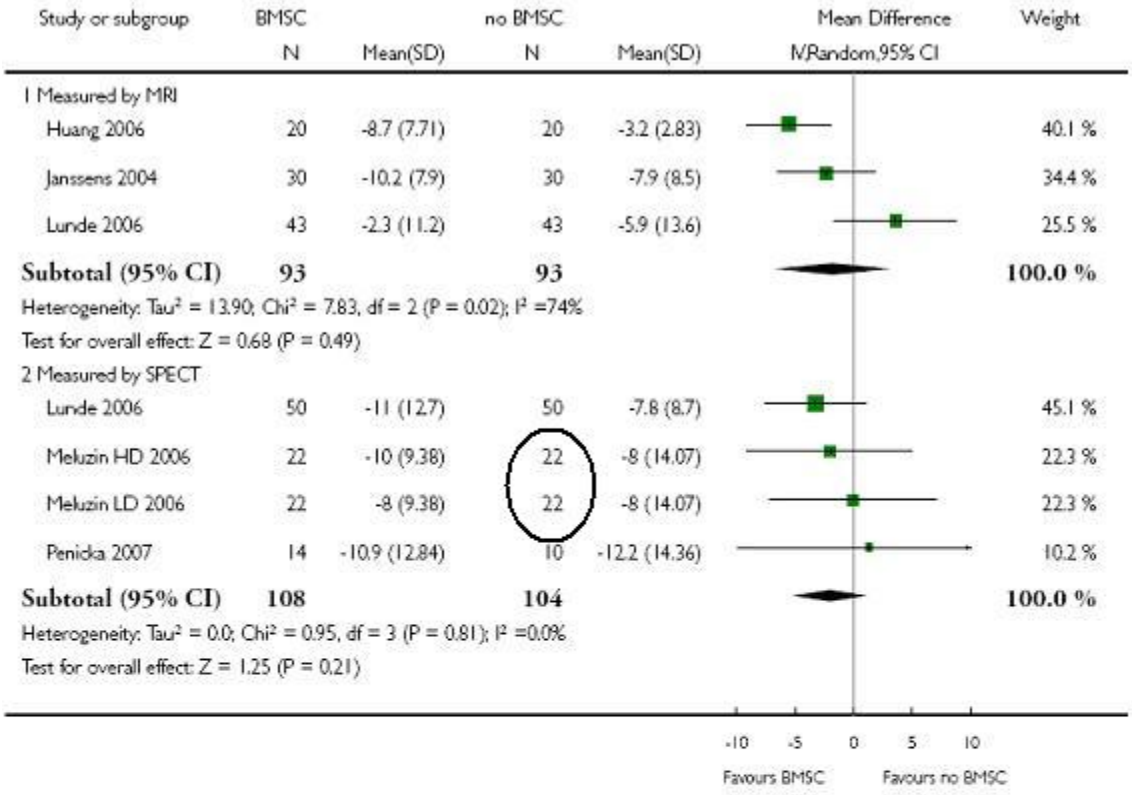
Figure 4.6: Physical fitness training for stroke patients by Sander et al(2004)



Example 2: Review published in Cochrane library is the meta-analysis by Rendon et al [85]. He added the trial by Meluzin [86] with three-arm comparison, which had 22-patients assessed in each arm. The two treatment arms were compared with the same control arm. This led to double counting of control groups. This is shown in Figure 4.7.

Example3: Meta-analysis conducted by Urquahart et al [87] titled Antidepressants for non-specific low back pain published in Cochrane library. Publication by Atkinson et al [88] were added in meta-analysis. Atkinson et al compared Paroxetine (n=22) and Maprotiline (n=20) drugs with the same active placebo (n= 32). The potential problem was that active placebo was being counted more than once. Figure 4.8 represents the double counting of some aspects of studies of Atkinson and his colleagues included in the meta-analysis by

Figure 4.7: Stem cell treatment for acute myocardial infarction by Rendon et al(2008)

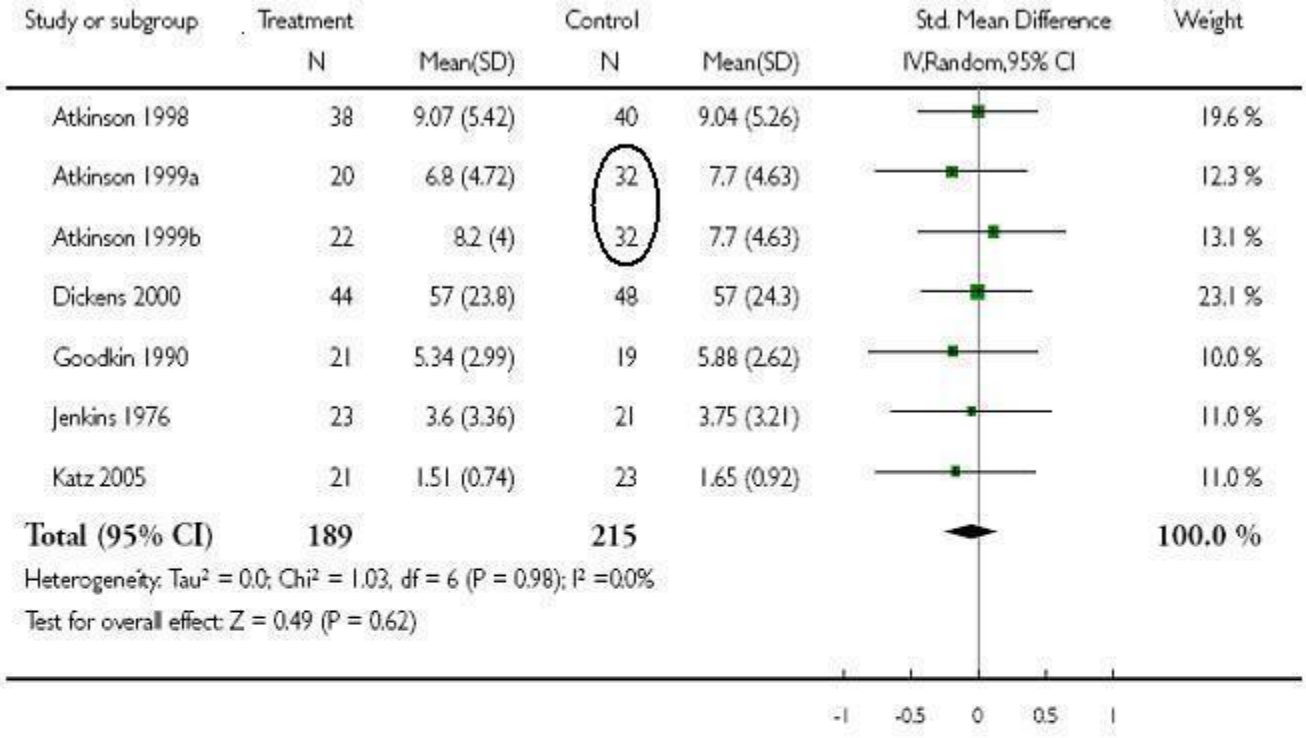


Urquahart et al.

Lastly, meta-analysis on Target-controlled infusion versus manually-controlled infusion of propofol for general anaesthesia or sedation in adults by Leslie et al [89]. One of its included studies named Struys et al [90] included two Manually-controlled infusion groups and one Target-controlled infusion group. Two comparisons were made: TCI vs MCI1 and TCI vs MCI2. This means that the placebo group was double counted. Figure 4.9 represents the meta-analysis of Leslie et al [89].

There was no double counting of some aspects of studies identified in Paper - Base Journal but we noticed that in a article published in BMJ by Arrol et al [91] titled "Are antibiotics effective for acute purulent rhinitis? Systematic review and Meta-Analysis of placebo Controlled Randomised Trials" included the publication of Taylor et al [92] who

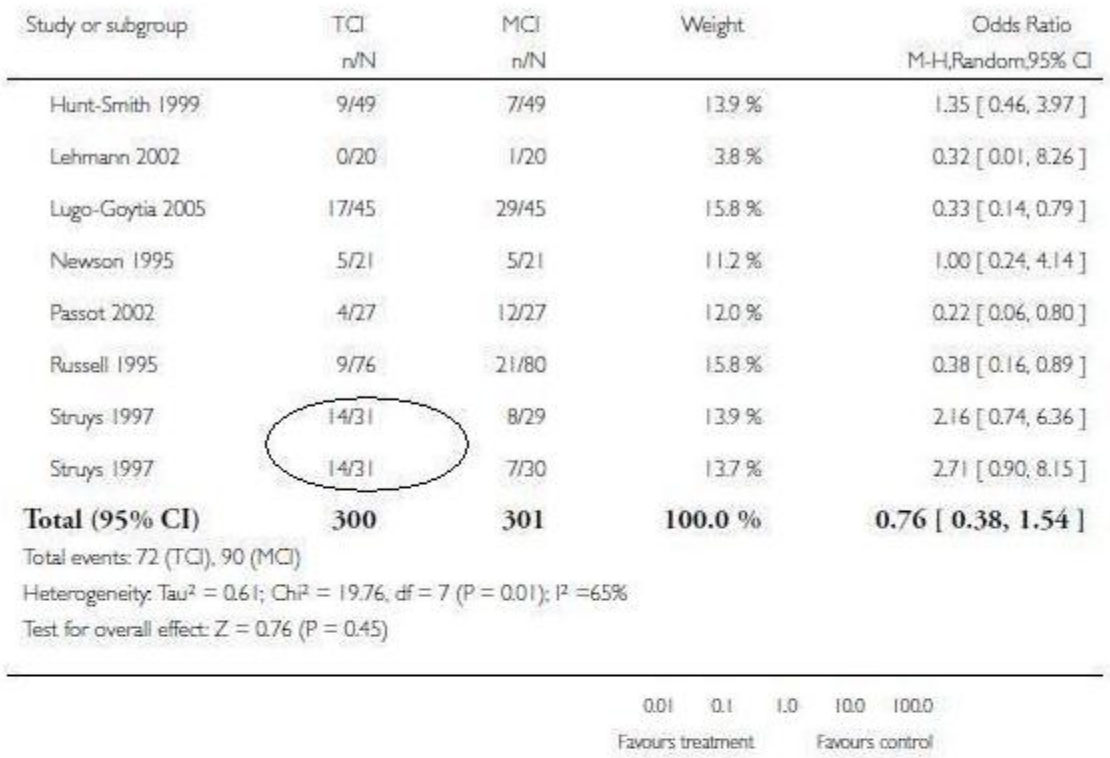
Figure 4.8: Double counting of some aspects of studies of Atkinson et al(1999)



compared more than two arms which include two treatment arms and one placebo arm. He included 188 patients (Amoxillim = 54 patients, Cotrimoxazole = 75 patients and placebo = 59). What the authors did was to divide the placebo into two section and assigned 30 and 29 of placebo group to each treatment group for analysis, and this technique prevented the control arm being counted twice .

We cite other examples of paper-based journal that escaped the problem of double counting of some aspects of studies. The meta-analysis of comparisons of Traditional digital Block and Single Subcutaneous Palmar Injection Block published in Journal of Hand Surgery by Yin et al [93]. In this meta-analysis, they included study published by Hung et al [94] who compared three block techniques - one treatment trial (transthecal) and two control trials (traditional block and subcutaneous palmar block), in that case, treatment arm was

Figure 4.9: Double counting of some aspects of studies of Struys et al (2008)



Study or subgroup	TCI n/N	MCI n/N	Weight	Odds Ratio M-H,Random,95% CI
Hunt-Smith 1999	9/49	7/49	13.9 %	1.35 [0.46, 3.97]
Lehmann 2002	0/20	1/20	3.8 %	0.32 [0.01, 8.26]
Lugo-Goytia 2005	17/45	29/45	15.8 %	0.33 [0.14, 0.79]
Newson 1995	5/21	5/21	11.2 %	1.00 [0.24, 4.14]
Passot 2002	4/27	12/27	12.0 %	0.22 [0.06, 0.80]
Russell 1995	9/76	21/80	15.8 %	0.38 [0.16, 0.89]
Struys 1997	14/31	8/29	13.9 %	2.16 [0.74, 6.36]
Struys 1997	14/31	7/30	13.7 %	2.71 [0.90, 8.15]
Total (95% CI)	**300**	**301**	**100.0 %**	**0.76 [0.38, 1.54]**

Total events: 72 (TCI), 90 (MCI)

Heterogeneity: $Tau^2 = 0.61$; $Chi^2 = 19.76$, df = 7 (P = 0.01); $I^2 = 65\%$

Test for overall effect: Z = 0.76 (P = 0.45)

divided into two comparison groups, transthecal versus traditional block and transthecal versus subcutaneous palmar block.

Another example was "Effect of iron supplementation on mental and motor development in children: systematic review of randomised controlled trials"by Sachdev et al published in Public Health Nutrition journal. In this meta-analysis, there are some studies with two or more iron intervention groups and a single control group , the authors divided equally the sample size of the control group between the number of intervention groups while retaining the same value for the change in outcome and its standard deviation [95]. This was done to avoid multiple counting of control group.

4.6 DATA IMPUTATION

In the course of review, we came across studies that reported the imputation of data during their statistical analysis as a result of missing data. Some authors of the unpublished data / missing data were contacted but there were no responses. The only alternative was to impute the missing or unreported data. The studies include - in Cochrane review,

- (i) a review on Propofol for sedation during Colonoscopy by Singh et al [96] stated that all studies where the standard deviation were not reported, was calculated from other measures of variance using the methodology suggested by Cochrane collaboration.

- (ii) review on Fluoride mouth rinses for preventing dental caries in children and adolescents by Marinho et al (2003) [97] narrated how the missing main outcome data was handled, the standard deviation for caries increments was imputed through the linear regression of log standard deviation on log mean caries increments.

- (iii) review on Non-steroidal anti-inflammatory drugs (NSAIDs) for treating lateral elbow pain in adults by Green et al [98] explained that where necessary standard

deviation was imputed from the range by division by 4.

Moreover, in Paper-base journal- a meta-analysis on Effects of calcium supplementation on bone density in healthy children:

- (i) meta-analysis of randomised controlled trials by Winzenberg et al(2006) [99] published in Medline database, imputed endpoint using baseline plus change for the mean and using standard deviation of the baseline data for the endpoint standard deviation where data for the analysis were not available.

The imputation of standard deviation using confidence intervals, standard errors and P-value when they are reported in the same study gives direct estimates. However, imputing standard deviation from studies within the same meta-analysis or from studies reported in another meta-analysis was found to yield accurate estimates of standard deviation [100]. However, the use of this approach calls for a careful appraisal of the how closely distributed the data from individual studies are to the study whose standard deviation is being imputed. Range imputation is very easy to use but can only be used in studies where the maximum and minimum responses are reported. It may not be a good approximation of the standard deviation as it may under estimate the actual value of standard deviation.

The use of linear regression method for imputation tends to underestimate variability in the data. Also, the method assumes linear relationship between the dependent and independent variables. Hence linearity must be established for this method to be effective.

The imputation of the endpoint standard deviation with baseline standard deviation could yield approximately good results when the scale of measurement, the duration and degree of measurement errors are the same in both the baseline and endpoint. But when this condition fails, misleading results could be obtained.

It is our submission that the problem and need for imputation can be minimized if study authors are sensitized enough on the need to include as much statistical information in

published individual studies. Item 16 in our checklist is intended to address this.

Chapter 5

ASSESSMENT OF PRECISION OF REPORTED STUDIES

Analysis of data from epidemiological studies rely on the estimation of the odds ratio, relative risks and their 95% confidence interval which must be reported by the authors [2]. Before incorporating studies in meta-analysis, it is expedient that authors must check the accuracy of relative risks, confidence intervals and odds ratios against the data source where available.

Peter Lee [2] developed a simple method for checking errors in relative risks, odds ratios and confidence intervals in meta-analysis which comprised of checks on calculated statistics reported in published papers after obtaining data for each study. The checks include:

- Consistency of Odds Ratios and Confidence Intervals in 2x2 table
- Internal consistency of OR and CI
- Minimum number of total subjects given the CI
- Minimum number of subjects in any two groups combined given the CIs
- Minimum number of subjects in any group combined given the CI

- Minimum number of total diseased subjects
- Minimum number of diseased subjects in either exposure groups

Lee [2] used these methods to assess the computed statistics of several studies-(E.g, Wu-Williams et al [101], Geng et al [102], Hackshaw et al [103] and found that some of the reported estimates were wrongly calculated. In this thesis, we used the same Lee simple methods to assess claims of precision of individual studies.

Senn [52] alerted in his article titled "Overstating the evidence"published in BMC, that users should be aware of any potential problems in the summary of reported estimates and more serious problems of double counting of studies, double counting of some aspects of studies, imputing data, poor reporting, accepting implausible claims for the precision of individual studies, etc. He advised that authors must constantly check analysed results before publishing and suggested that important standard for judgement is CHECKABILITY which will improve the effectiveness and validity of meta-analysis. He therefore designed a tool for checkability consisting of five points as are listed below

- Be vigilant about double counting
- Make results checkable
- Describe approaches to analysis in detail
- Judge the meta-analysis not the analyst
- Create a culture of correction

Weir and Senn [3] developed SAS MACRO versions 8.2 in June/August 2008 named mapeterlee and mabinary SAS Macros which we used to carry out checks as suggested by Lee [2]. We used these macros in this project to check accuracy of reported estimates.

5.1 SAS MACRO

According to Cohen (*A tutorial on SAS macro language, AsrtZeneca*)"SAS Macro language is another language that rests on top of regular SAS code. It can make programme easier and save repetition and tedious effort. It consists of macro variables, macro language statements, regular SAS program statements and macro functions which are contained within %MACRO and a %MEND. The %MACRO statement includes a name and the macro is called using the macro's name preceded by a %"[104] .

Weir and Senn [3] developed different types of SAS macro program but in this project, we will make use of mapeterlee.sas and mabinary.sas in conducting checks on published estimates.

(a)Mapeterlee.sas macro was used to calculate the minimum number of total subjects given confidence intervals, the value that any two elements should be equal or greater than and that which any one element should be equal or greater than. The dataset needed for the use of this macro include study name and standard error. (b)Mabinary.sas macro was used to check consistency of odds ratio or relative risks and confidence intervals using raw data of included studies. The input data for this macro consist of study name, treat-event, control-event, number in the treatment group (Nt), and number in the control group (Nc).

To do this, we selected ten studies each from Cochrane database and Paper-based Journals. The dataset input needed to execute the checks consist of study name, the number of patients that received treatment, number of subjects that received placebo, the total number of subjects in treatment and placebo, the point estimates(OR/RR) and their 95 or 99 percent Confidence intervals. In this case, different methods were applied to effect error checks on the estimated Odds ratio and Relative risk in a 2x2 table, where raw data are available. Lee [2] checks are detailed in the following subsections:

5.1.1 Check 1: The consistency of Odds ratio/Relative risks and Confidence Interval given a 2 x 2 table

The first check was performed to verify the reliability of Odds ratio or Relative risks and Confidence intervals in a 2x2 table where data are available. The cell frequencies in the 2x2 contingency table is represented by *a= number of exposed cases, b = number of unexposed cases, c = number of exposed control and d = number of unexposed control*, the sketch of 2x2 table is shown in Table 5.1.

	Exposed	Unexposed	Total
Cases	a	b	$a+b$
Control	c	d	$c+d$
Total	$a+c$	$b+d$	$a+b+c+d$

Table 5.1: The Contigency Table

Equations (1) and (2) represent the estimation of Odds ratio and Relative risks respectively using available datasets from the selected studies.

Odds ratio = $\hat{r}$ = ad/bc ——————————————————(1)

Relative Risk = $\hat{r} = a(b+d)/b(a+c)$ ——————————————————- (2)

and the variance of $ln\hat{r}$ for both Odds ratio and Relative risks were estimated from 2x2 table as given in Equations (3) and (4), the variance estimates obtained were used to calculate upper and lower confidence limits of the OR / RR for each study using equation (5). The results derived were used to check for discrepancies on the reported values of lower and upper confidence limits in the included studies.

OR : $Var(lnr) = V_1 = 1/a + 1/b + 1/c + 1/d$ ——————————————————(3)

RR : $Var(lnr) = V_2 = 1/a + 1/b - 1/(a+c) - 1/(b+d)$ ——————————————————(4)

$ln\hat{r}_L$ $ln\hat{r}_u = ln\hat{r} \pm 1.96\sqrt{Var}$ ——————————————————-(5)

5.1.2 Check 2: Internal consistency of OR/RR and CI

This method is used to check the internal consistency of Odds ratio and confidence intervals when the original data are not accessible. Error check was performed on evenness of confidence intervals and odds ratio/Relative risk using equation(6)

$OR^2/RR^2 = r^2 = r_u \times r_l$ ——————————(6)

where r_u and r_l = Upper and Lower Confidence Limits

5.1.3 Check 3: Minimum number of total subjects given the CI in a case control study

Checks can also only be based on the minimum size of the study implied by confidence interval, when the lower and upper confidence interval (r_u and r_L) are given, they are known to be 3.92 for (95%) or 5.16 for (99%) standard errors apart on the logarithmic scale. Standard error was calculated using Lee' method shown in equation (7) and results obtained were used as input in Mapeterlee sas macro. Mapeterlee sas.macro input data consists of study name and standard error with a maximum of eight characters at. Alternatively, checks on minimum number of total subjects in the group for odds ratio can be performed using formular (8). The check is based on fact that the total number of subjects in the study should be greater than or equal to this value estimated using (8).

$V_1 = [ln(r_u/r_l)/3.92]^2 = Q^2/15.3664$ ——————————(7)

$StandardError = \sqrt{variance}$

Where $Q = ln(r_u/r_l)$

$N \geq 16/V_1$ or $N \geq 245.86/Q^2$ ——————————(8)

where $V_1 = 1/a + 1/b + 1/c + 1/d$

N = total number of subjects in the study

5.1.4 Check 4: Minimum number of subjects in any two groups combined given the CIs in a case control study

The addition of any two elements of the 2x2 contigency table should be greater than or equal to this value, this implies that

$V_1 \geq \frac{1}{m} + \frac{1}{n}$ ——————————(9)

where m and n are any two different elements of the 2x2 table.

therefore $(m + n) \geq 61.47/Q^2$ ——————————-(10)

where 61.47 is constant

5.1.5 Check 5: Minimum number of subjects in any group given the CI in a case control study

Any one element of 2x2 table should be greater than or equal this value which implies that,

$V_1 \geq \frac{1}{m}$, where m is any element of the 2x2 table

then, $m \geq 15.3664/Q^2$ ——————————-(11)

5.1.6 Check 6: Minimum number of total diseased subjects in a cohort study

Here, we concentrate on the number of diseased subjects in both exposed and unexposed groups. These groups must be greater than or equal to this estimated value from equation (14). Equation (2) can be rewritten as

$V_2 = \frac{a(a+c)}{c} + \frac{b(b+d)}{d} \geq 61.47/Q^2$ ——————————(12)

$a + b > a + \frac{bc(b+c)}{d(a+c)} >= 61.47C/Q^2(a + c)$ ——————————-(13)

$a + b \geq 61.47S/Q^2$ ——————————————————(14)

where $S = c/(a + c)$, is survival rate in the group at higher risk, for a rare disease $S \approx 1$

and for a relatively common disease in the exposed group, for example $S = 0.9$

5.1.7 Check 7: Minimum number of diseased subjects in either exposure group in a cohort study

Equation (11) will be applied where m is either a or b

5.1.8 Application of Peter Lee's Methods on selected studies

We used Peter Lee's [2] simple methods to assess the consistency of reported odd ratios, relative risks, confidence intervals and minimum numbers of subjects sizes both in exposure and unexposed groups in some selected studies. Peter Lee's checks include simple direct calculation of odds ratio and standard errors when all four frequencies for a four-fold table are available. These lead to equalities that must be satisfied. However his checks also provide inequalities that should be satisfied when the individual entries are not available but only certain marginal totals. Obviously, where the checks using individual frequencies can be done there is no point in doing the checks involving marginal totals since the marginal totals are determined by the cell frequencies. However, in order to illustrate the method we decided to perform these checks anyway. This gave a surprising result. There were some studies that passed the the checks involving individual entries but not the checks using totals. The only possible explanation is that not all of Lee's checks are correct. It seems that the inequalities that he provides for analyses based on odds ratios are correct but that those based on relative risk are not correct. We will discuss this point later in this chapter.

Data from individual studies included in published meta-analysis in Cochrane and Non-Cochrane reviews which are used for these checks are given in the first four columns (left hand side) of Tables 5.2 - 5.21 below.

We selected ten studies of meta-analyses from Cochrane and Non-Cochrane databases

respectively. These studies have information on Study names, treatment groups size, placebo size and estimates of MH(OR/RR) with their 95 or 99% CI which we presented in tables in sub-section 5.2.9. The tables contain two columns - the left-hand side contains reported data and estimates from the authors of meta-analysis and the right-hand columns display our results of estimated Standard Errors and results of Lee's checks.

5.1.9 The sparse events and the use of continuity correction factor of (0.5)

Using Peter Lee's simple methods and mapeterlee.sas macro, we calculated estimates of odds ratio, relative risks and CIs. Some calculated CIs and RR/OR differ from the reported CIs and RR/OR and the data did not meet requirements of minimum number of subjects sizes for given CI.

A close observation of the data in the individual studies indicate that one prominent factor that contributed to the discrepancies between reported estimates and calculated estimates is zero outcomes in either the treatment or control group or both. To bridge this gap, we introduced a continuity correction factor to all studies that reported zero outcomes. That is, we added continuity correction factor of 0.5 to each cell. Results obtained were more efficient in terms of matching the reported estimates of effect size and of CIs. Our analysis tend to suggest that studies that reported odds ratio are more likely to meet the minimum subject size requirement than studies reporting relative risks.

We calculated standard errors of individual studies using data reported by the authors. The SEs were used as input dataset in the mapeterlee.sas.macro.

We employed each of the formula stated above (see Checks (1),(2),(3),(4),(5),(6) and (7)) to assess error in the published studies. Studies reporting estimates of RR or OR were treated differently using relevant formula designed for the method (see subsections 5.2.3 -

5.2.7). The results obtained from these formulas were used to verify consistency of statistical estimates reported by the authors.

Abbreviated terms with their meaning

CI ——— Confidence intervals

RR/OR —— Relative risks/Odds ratio

Nt/Nc —— Total number of treatment/control group

(P) ——— Pass

(F) ——— Fail

SE ——— Standard Error Check(1) - Consistency of Odds Ratios and Confidence Intervals with 2x2 table.

Check(2) - Internal consistency of OR and CI.

Check(3) - Minimum number of total subjects given the CI(TS).

Check(4) - Minimum number of subjects in any two groups combined given the Confidence intervals(Two ele).

Check(5) - Minimum number of subjects in any group combined given the CI(One-ele).

Check(6) - Minimum number of total diseased subjects(TDS).

Check(7) - Minimum number of diseased subjects in either exposure groups(DS).

5.1.10 Examples of procedures used for checking errors in reported estimates

Example: A study by Augustine et al 2004 (see the study in the first row of table 5.2) published in Meta-analysis by Rabindranath et al [105] in Cochrane library is used here to illustrate how the results presented in the right-hand side columns of Tables 5.2 - Tables 5.21 were obtained. The data as reported by the authors is given as

∗**Augustine 2004**

$n_1 = 40$, $n_2 = 40$,

$b = 28$, $d = n_2 - b$, $d = 12$ Frequencies in four-fold table

Total subjects = 80

confidence limits:-

$LCL = 0.72$, $UCL = 1.30$

Relative Risks = 0.96 ,

Given the fact the authors reported relative risk, we shall use checks (1), (2) formula to assess correctness of reported estimates.

(i) STANDARD ERRORS - it is estimated based on frquencies

SE = $\sqrt{V_1}$

Where the variance of Relative risks based on frequencies is estimated by

$V_1 = 1/A + 1/B - 1/(A + C) - 1/(B + D)$

$V_1 = 1/27 + 1/28 - 1/(27 + 13) - 1/(28 + 12)$

$V_1 = 1/27 + 1/28 - 1/40 - 1/40$

$V_1 = 0.037 + 0.036 - 0.025 - 0.025$

$V_1 = 0.023$ ——————————(1)

Therefore, Standard error = $\sqrt{0.023} = 0.15$

(ii) Check (1) - Consistency of Relative Risks and Confidence Intervals with 2x2 table

Relative risks = $A(B + D)/B(A + C)$

$RR = 27(28 + 12)/28(27 + 13)$

$RR = 27(40)/28(40) = 1080/1120$

$RR = 0.96$

The $CI[LCL, UCL]$ is then approximated by noting that

substitute value of $V_1 = 0.023$ from eqn (1) to eqn(2)

$lnLCL, lnUCL = lnRR \pm 1.96(\sqrt{V_1})$ ——————————(2)

$lnLCL, lnUCL = ln(0.96) \pm 1.96\sqrt{(0.023)}$

$lnLCL, lnUCL = -0.041 \pm 1.96 \times 0.15$

$= -0.041 \pm 0.297$

$lnLCL = e^{-0.34} = 0.71$

$lnUCL = e^{0.256} = 1.30$

Lower limit 95% = 0.71

Upper limit 95% =1.30

Hence, we assigned pass (p) to check(1) since our calculated RR ,lower and upper confidence limits agree with that reported by Augustine et al 2004.

(iii) Check (2)Internal consistency of RR and CIs.

The formula is $\sqrt{LCL \times ULC} = 0.96 = RR = 0.96$

We assigned pass (p) to check(2) because of the consistency between Relative Risk and CIs

5.1.11 Peter Lee's checks on Cochrane reviews

Rabindranath [105]	meta-analysis	(Reported)			Lee's	Check	
Study	treatment	control	M-H(RR), 95%	SE	TDS	DS	(1)(2)(6)(7)
∗**Augustine 2004**	27/40	28/40	0.96[0.72 , 1.30]	0.15	57.95	44.01	(p)(p)(f)(f)
Gasparovic 2003	37/52	31/52	1.19[0.90, 1.58]	0.14	55.98	48.52	(p)(p)(p)(f)
Mehta 2001	55/84	39/82	1.38[1.05, 1.81]	0.14	71.57	51.82	(p)(p)(p)(p)
Noble 2006	43/54	34/40	0.94[0.78, 1.13]	0.09	91.13	111.83	(p)(p)(f)(f)
Uehlinge 2005	33/70	28/55	0.93[0.65, 1.33]	0.18	63.39	29.98	(p)(p)(f)(p)
Vinsonneau 2006	118/176	126/184	0.98[0.85, 1.13]	0.073	249.86	189.53	(p)(p)(f)(f)
Total	476	453	1.03[0.92, 1.16]				

Table 5.2: *Meta-analysis by Rabindranath et al(2007)*

Terplan[106]	meta-analysis	(Reported)			Lee's	Check	
Study	treat	control	MH (RR)95% CI	SE	TDS	DS	(1)(2)(6)(7)
Elk 1998	5/6	4/6	1.25[0.64 - 2.44]	0.34	5.72	8.58	(p)(p)(p)(f)
Haug 2004	26/30	28/33	1.02[0.84 - 1.25]	0.10	51.87	97.25	(p)(p)(p)(f)
Jones 2001	44/47	36/38	0.99[0.89 - 1.10]	0.05	87.43	342.40	(p)(p)(f)(f)
Mulins 2004	17/35	23/36	0.76[.50 - 1.16]	0.21	44.64	21.70	(p)(p)(f)(f)
O'Neil 1996	40/47	40/45	0.96[0.82 - 1.12]	0.08	94.18	158.08	(p)(p)(f)(f)
Silveman 2001	11/20	8/20	1.38[0.71 - 2.68]	0.34	15.68	24.78	(p)(p)(p)(p)
Svikis 1997	27/40	18/36	1.35[0.91 - 2.00]	0.20	32.22	8.71	(p)(p)(p)(p)
Total (95% CI)	224	214	1.02[0.92 - 1.13]				

Table 5.3: *Psychosocial interventions for pregnant women in outpatient illicit drug treatment programs compared to other interventions by Terplan(2007)*

Alonso[107]	Meta-analysis	(Reported)			Lee's	Check	
Study	treatment	control	M-H(RR)	SE	TDS	DS	(1)(2)(6)(7)
Himmelfarb	46/72	35/81	1.48[1.09 - 2.00]	0.15	60.25	41.71	(p)(p)(p)(p)
Kurtal 1995	48/84	43/76	1.01[0.77 - 1.32]	0.14	90.68	52.89	(p)(p)(p)(f)
Micheal 1995	7/11	10/12	0.76[0.46 - 1.28]	0.26	21.34	14.67	(p)(p)(f)(f)
Romao 1999	12/20	16/24	0.90[0.57 - 1.42]	0.23	29.51	18.44	(p)(p)(f)(f)
Schifffl 1994	20/26	9/26	2.22[1.26 - 3.92]	0.29	11.01	11.93	(p)(p)(p)(p)
Valeri 1996	19/25	16/28	1.33[0.90 - 1.96]	0.20	24.35	25.37	(p)(p)(p)(f)
Woo 2002	21/23	17/20	1.07[0.86 - 1.34]	0.11	27.18	78.13	(p)(p)(p)(f)
Kurtal 1995	13/25	17/32	0.98[0.60 - 1.61]	0.25	30.28	15.77	(p)(p)(p)(p)
Subtotal	286	299	1.15[0.95 - 1.38]				

Table 5.4: *Biocompatible hemodialysis membranes for acute renal failure Alonso et al (2008)*

Macfacdyne[81]	Meta-analysis	(Reported)			Lee's	Check	
Study Group	quinolone	Control	MH-RR 95%	SE	TDS	DS	(1)(2)(6)(7)
Tutkun 1995	3/24	14/20	0.18(0.06, 0.53)	0.56	11.33	3.24	(p)(p)(p)(p)
Van Hasselt 1997	3/14	7/40	1.22(0.37, 4.10)	0.61	8.35	2.66	(p)(p)(p)(p)
VH 1998 daily	9/32	12/36	0.84(0.41, 1.74)	0.37	21.15	7.35	(p)(p)(p)(p)
VH 1998 wkly	16/39	13/31	0.98(0.56, 1.71)	0.28	29.09	12.33	(p)(p)(p)(p)
Fradis 1997	10/19	8/18	1.18(0.61, 2.31)	0.34	16.42	8.67	(p)(p)(p)(p)
Kaygusuz 2002	4/20	6/20	0.67(0.22, 2.01)	0.56	10.05	3.14	(p)(p)(p)(p)
Subtotal (95% CI)	148	165	0.76 (0.55, 1.04)				

Table 5.5: *Topical antibiotics without steroids for chronically discharging ears with underlying eardrum perforations by Macfadyen et al (2005)*

Martin[108]	Meta-analysis	(Reported)			Peter Lee	Check	
Study	BMSC	no BMSC	MH(RR) 95% CI	SE	TDS	DS	(1)(2)(6)(7)
Kang 2006	0/25	2/25	0.20(0.01 - 3.97)	1.53	1.72	0.43	(p)(p)(p)(p)
Lunde 2006	11/50	11/50	1.00(0.48 - 2.09)	0.38	22.15	7.10	(p)(p)(p)(p)
Meluzin 2006a	2/22	1/20	1.82(0.18 - 18.55)	1.08	2.60	0.72	(p)(p)(p)(p)
Meluzin 2006b	4/22	1/20	3.64(0.44 - 29.87)	1.18	2.83	0.86	(p)(p)(p)(p)
Mayer 2006	10/28	9/28	1.11(0.53 - 2.31)	0.38	18.23	7.09	(p)(p)(p)(p)
Subtotal (95%)	147	143	1.10[0.63 - 1.80]				

Table 5.6: *Stem cell treatment for acute myocardial infarction Martin-Rendo et al(2008)*

McCallum [109]	Meta-analysis	(Reported)			Peter Lee	Check	
Study	Open	Closed	MH(RR), 95% CI	SE	TDS	DS	(1)(2)(6)(7)
Al-Hassan 90	5/42	8/40	0.60[0.21 - 1.67]	0.53	12.60	3.57	(p)(p)(p)(p)
Fzn 94	0/45	2/46	0.20[0.01 - 4.14]	1.55	1.66	0.41	(p)(p)(p)(p)
Gencosmaoglu	1/73	12/69	0.08[0.01 - 0.59]	1.04	3.65	0.92	(p)(p)(p)(p)
Hameed 01	1/20	2/23	0.58[0.06 - 5.88]	1.17	2.78	0.73	(p)(p)(p)(p)
Khawaja 01	0/23	0/23	0.0[0.0 - 0.0]	0	0	0	undefined
Kronborg 85	4/32	14/67	0.60[0.21 - 1.67]	0.53	12.51	3.57	(p)(p)(p)(p)
Miocinovic 99	2/25	6/25	0.33[0.07 - 1.50]	0.78	6.02	1.64	(p)(p)(p)(p)
Mohammed 05	2/55	3/28	0.34[0.06 - 1.92]	0.88	4.93	1.28	(p)(p)(p)(p)
Sndenna 92	3/60	6/60	0.50[0.13 - 1.91]	0.69	8.086	2.13	(p)(p)(p)(p)
Subtotal	375	381	0.38[0.23 - 0.63]				

Table 5.7: *Healing by primary versus secondary intention after surgical treatment for pilonidal sinus by McCallum (2007)*

In Table 5.7, all the reported CIs and RR agree with calculated estimates using Lee's simple method except khawaja et al that reported zero confidence intervals.

Villatoro[110]	MA	(Reported)			Peter Lee	Check		
Study	Treat	control	MH(OR), 95% CI	SE	TS	TE	OE	(1)(2)(3)(4)(5)
Isenmann	3/41	4/35	0.61[0.13 - 2.94]	0.80	25.28	6.32	1.58	(p)(p)(p)(p)(p)
Nordback	2/25	5/33	0.49[0.09 - 2.75]	0.87	21.03	5.26	1.31	(p)(p)(p)(p)(p)
Pederzoli 1993	3/41	4/33	0.57[0.12 - 2.76]	0.80	25.01	6.25	1.56	(p)(p)(p)(p)(p)
Sainio 1995	1/30	7/30	0.11[0.01 - 0.99]	1.17	11.64	2.91	0.73	(p)(p)(p)(p)(p)
Schwarz 1997	0/13	2/13	0.17[0.01 - 3.92]	1.52	6.90	1.72	0.43	(p)(p)(p)(p)(p)
Total	150	144	0.37[0.17 - 0.83]					

Table 5.8: A*ntibiotic therapy for prophylaxis against infection of pancreatic necrosis in acute pancreatitis by Villatoro et al (2006).SE-standard error,TS-total subjects,TE-two elements,OE-one element*

Leslie[89]	(Reported)			Lee's	Check			
Study	treat	control	M-H(OR), 95% CI	SE	TS	TE	OE	(1)(2)(3)(4)(5)
H-Smith 1999	9/49	7/49	1.35[0.46 - 3.97]	0.55	52.93	13.23	3.31	(p)(p)(p)(p)(p)
Lehmann 2002	0/20	1/20	0.32[0.01 - 8.26]	1.71	5.45	1.36	0.34	(p)(p)(p)(p)(p)
l-Goytia 2005	17/45	29/45	0.33[0.14 - 0.79]	0.44	82.11	20.53	5.13	(p)(p)(p)(p)(p)
Newson 1995	5/21	5/21	1.00[0.24 - 4.14]	0.73	30.32	7.58	1.89	(p)(p)(p)(p)(p)
Passot 2002	4/27	12/27	0.22[0.06 - 0.80]	0.66	36.64	9.16	2.29	(p)(p)(p)(p)(p)
Russell 1995	9/76	21/80	0.38[0.16 - 0.89]	0.44	83.49	20.87	5.22	(p)(p)(p)(p)(p)
Struys 1997	14/31	8/29	2.16[0.74 - 6.36]	0.55	53.13	13.28	3.32	(p)(p)(p)(p)(p)
Struys 1997	14/31	7/30	2.71[0.90 - 8.15]	0.56	50.64	12.66	3.17	(p)(p)(p)(p)(p)
Total	300	301	0.76[0.38 - 1.54]					

Table 5.9: *Target-controlled infusion versus manually-controlled infusion of propofol for general anaesthesia or sedation in adults Leslie et al (2008).SE-standard error,TS-total subjects,TE-two elements,OE-one element*

.

Ritcher[111]	(Reported)			Lee's	Check			
Study	Rosigli	control	M-H(OR), 95%	SE	TS	TE	OE	(1)(2)(3)(4)(5)
Hanefeld 07	18/200	4/207	5.02[1.67 - 15.10]	0.56	50.71	12.68	3.17	(p)(p)(p)(p(p))
Kahn 2006	205/1456	123/1441	1.76[1.39 - 2.22]	0.12	1121.6	280.4	70.1	(p)(p)(p)(p)(p)
Ko 2006	2/56	0/56	5.18[0.24 - 110.45]	1.56	6.54	1.64	0.41	(p)(p)(p)(p)(p)
Lebovitz 2001	18/169	3/158	6.16[1.78 - 21.34]	0.63	39.85	9.96	2.49	(p)(p)(p)(p)(p)
Philips 2001	13/187	3/173	4.23[1.19 - 15.12]	0.65	38.05	9.51	2.38	(p)(p)(p)(p)(p)
Raskin 2004	2/62	0/63	5.25[0.25 - 111.56]	1.56	6.61	1.65	0.41	(p)(p)(p)(p)(p)
Rosenstock 06	14/112	0/104	30.77[1.81 - 522.71]	1.44	7.66	1.92	0.48	(p)(p)(p)(p)(p)
Stocker 2007	8/45	0/47	21.53[1.20 - 385.19]	1.47	7.38	1.85	0.46	(p)(p)(p)(p)(p)
Sutton 2002	7/104	1/99	7.07[0.85 - 58.57]	1.08	13.72	3.43	0.86	(p)(p)(p)(p)(p)
Total	2391	2348	2.27[1.83 - 2.81]					

Table 5.10: *Rosiglitazone for type 2 diabetes mellitus by Ritcher et al (2007).SE-standard error,TS-total subjects,TE-two elements,OE-one element*

5.1.12 Peter Lee's checks on Paper-Based Journals

Only Devereaux et al [113] used 99% CIs in their analysis as they believe that it conveys better their confidence in the estimate of treatment effect given the fact that the statistical signifcance of their analysis depends on the difference of only a handful of the events due to small sample size. Reported CIs of some of the included studies in Devereaux et al [113] and Webster et al [114] are consistent with calculated CIs obtained using Peter Lee methods. Table 5.12 and Table 5.13 show respectively, outcomes of our investigations for the two studies.

Poole[112]	(Reported)			Lee's	Check			
Study	treat	contr	M-H(OR), 95%	SE	TS	TE	OE	(1)(2)(3)(4)(5)
Allegra	111/171	89/181	1.91[1.25, 2.94]	0.22	336.12	84.03	21.01	(p)(p)(p)(p)(p)
Baboli	134/254	58/241	3.52[2.40, 5.18]	0.20	415.40	103.85	25.96	(p)(p)(p)(p)(p)
Boman	46/98	29/105	2.32[1.29, 4.15]	0.29	180.08	45.02	11.25	(p)(p)(p)(p)(p)
Borgia	7/10	4/7	2.92[0.44,19.23]	0.96	17.23	4.31	1.08	(p)(p)(p)(p)(p)
Castig	240/311	179/302	2.32[1.64, 3.30]	0.18	502.87	125.72	31.43	(p)(p)(p)(p)(p)
Cremon	8/21	0/20	25.9[1.4, 485.3]	1.49	7.14	1.78	0.45	(P)(P)(p)(p)(p)
Grass 76	18/35	11/34	2.21[0.83, 5.88]	0.49	64.14	16.04	4.01	(p)(p)(p)(p)(p)
Grass 94	25/42	14/41	2.84[1.16, 6.92]	0.46	77.08	19.27	4.82	(p)(p)(p)(p)(p)
Grillage	35/54	29/55	1.65[0.77, 3.57]	0.39	104.49	26.12	6.53	(p)(p)(p)(p)(p)
Hansen	36/59	34/70	1.66[0.82, 3.35]	0.36	124.12	31.03	7.76	(p)(p)(p)(p)(p)
Jackson	41/61	36/60	1.37[0.65, 2.87]	0.38	111.48	27.87	6.97	(p)(p)(p)(p)(p)
Malerba	28/44	24/47	1.68[0.72, 3.88]	0.43	86.66	21.67	5.42	(p)(p)(p)(p)(p)
McGavin	11/72	8/76	1.53[0.58, 4.06]	0.49	64.93	16.23	4.05	(p)(p)(p)(p)(p)
Meister 1986	37/90	34/91	1.17[0.64, 2.13]	0.31	170.05	42.51	10.63	(p)(p)(p)(p)(p)
Meister 1999	79/122	56/124	2.23[1.34, 3.73]	0.26	234.6	58.65	14.66	(p)(p)(p)(p)(p)
Moretti	26/63	13/61	2.59[1.18, 5.73]	0.40	98.46	24.6	6.15	(p)(p)(p)(p)(p)
Norwak	114/147	101/148	1.61[0.96, 2.70]	0.26	229.93	57.48	14.37	(p)(p)(p)(p)(p)
Olivieri	56/110	21/104	4.10[2.23, 7.52]	0.31	166.39	41.60	10.40	(p)(p)(p)(p)(p)
Pela	37/83	17/80	2.98[1.50, 5.94]	0.35	129.81	32.45	8.11	(p)(p)(p)(p)(p)
Rasmussen	28/44	24/47	1.68[0.72, 3.88]	0.43	86.67	21.67	5.42	(p)(p)(p)(p)(p)
Total(95%CI)	1891	1896	2.23[1.95, 2.56]					

Table 5.11: *Mucolytic agents for chronic bronchitis or chronic obstructive pulmonary disease by Poole et al (2006).SE-standard error,TS-total subjects,TE-two elements,OE-one element*

Devereaux[113]	Meta-analysis	(Reported)			Lee's	Checks	
Study	B blocker	cont	MH(RR) 99%CI	SE	TDS	DS	(1)(2)(6)(7)
Cucchiara	0/37	1/37	0.33[0.01, 21.46]	1.49	1.79	0.45	(p)(p)(p)(p)
Liu	0/16	1/14	0.29[0.00, 17.86]	0.39	0.9	0.22	(p)(p)(p)(p)
Magnusson	4/15	0/15	9.00[0.22, 375.21]	1.44	1.38	0.48	(p)(p)(p)(p)
Stone	10/89	0/39	9.33[0.23,374.09]	1.43	1.72	0.48	(p)(p)(p)(p)
Mackenzie	1/50	0/50	3.00[0.05, 195.17]	1.61	1.51	0.39	(p)(p)(p)(p)
Jackobson	5/49	1/49	5.00[0.31, 80.06]	1.08	3.10	0.86	(p)(p)(p)(p)
Davies	12/20	8/20	1.50[0.64, 3.50]	0.33	14.76	9.22	(p)(p)(p)(p)
Wallace	2/99	1/101	2.04[0.09, 46.84]	1.21	2.67	0.68	(p)(p)(p)(p)
Yang	53/246	19/250	2.83[1.48, 5.42]	0.25	49.59	15.80	(p)(p)(p)(p)
Total	621	575	2.27[1.36, 3.80]				

Table 5.12: *The effect of perioperative blocker treatment in patients having non-cardiac surgery by Devereaux et al 2005*

TDS=Minimum number of total diseased subjects, DS=Minimum number of diseased subjects in either exposure groups

Webster[114]	Meta-analysis	(Reported)			Lee's	Check	
Study group	treatment	control	MH(RR) 95%	SE	TDS	DS	(1)(2)(6)(7)
Raofi 99	0/14	0/21					undefined
Yang 99	0/30	1/31	0.33(0.01, 7.87)	1.70	1.38	0.35	(p)(p)(p)(p)
Van Duji 02	1/11	0/12	3.25(0.15, 72.36)	1.58	1.46	0.40	(p)(p)(p)(p)
Wang 2000	0/25	2/32	0.25(0.01, 5.06)	1.59	6.90	1.73	(p)(p)(p)(p)
White 2000	0/52	5/50	0.09(0.005, 1.54)	2.13	1.87	0.47	(p)(p)(p)(p)
Miller 02	1/100	3/50	0.17(0.02. 1.56)	1.11	3.21	0.81	(p)(p)(p)(p)
Laskow 95	4/92	1/28	1.22(0.14-10.45)	1.10	3.16	0.83	(p)(p)(p)(p)
Shapiro 91	2/28	3/29	0.69(0.12-3.83)	0.88	4.76	1.28	(p)(p)(p)(p)
Johnson	8/148	3/75	1.35(0.37-4.95)	0.66	8.64	2.28	(p)(p)(p)(p)
Trompeter 02	6/103	13/93	0.42(0.17 - 1.05)	0.46	17.46	4.64	(p)(p)(p)(p)
Campos 02	12/84	9/80	1.27(0.57 - 2.85)	0.41	20.34	5.93	(p)(p)(p)(p)
Pirsch 1997	10/205	19/207	0.53(0.25 - 1.11)	0.38	26.31	6.92	(p)(p)(p)(p)
Margreiter 02	17/286	22/271	0.73(0.40 - 1.35)	0.31	39.078	10.39	(p)(p)(p)(p)
Mayer 1997	38/303	18/145	1.01(0.60 - 1.71)	0.27	49.01	14.01	(p)(p)(p)(p)
Subtotal (95% CI)	1481	1123	0.77(0.58 - 1.02)				

Table 5.13: *Individual datasets in Journal reviews by Webster et al(2005)*

TDS=Minimum number of total diseased subjects, DS=Minimum number of diseased subjects in either exposure groups

Arroll[115]	Meta-analysis	(Reported)			Lee's	check	
Study	treat	cont	MH(RR) 95%CI	SE	TDS	DS	(1)(2)(6)(7)
Cederlof	20/26	14/25	1.37[0.92, 2.06]	0.21	21.83	23.65	(p)(p)(p)(f)
Dieppe	10/12	1/12	10.00[1.51, 66.43]	0.97	0.72	1.07	(p)(p)(p)(p)
Friedman	15/17	12/17	1.25[0.88, 1.78]	0.17	14.57	30.97	(p)(p)(p)(f)
Gaffney	33/42	21/42	1.57[1.12, 2.21]	0.17	28.51	33.26	(p)(p)(p)(p)
Ravaud	16/25	7/28	2.56[1.26, 5.18]	0.36	11.07	7.69	(p)(p)(p)(p)
Smith	25/38	15/33	1.45[0.93, 2.24]	0.22	27.21	19.89	(p)(p)(p)(p)
Total(95% CI)	160	157	1.66[1.37, 2.01]				

Table 5.14: *Corticosteroid injections for osteoarthritis of the knee: meta-analysis, by Arroll 2004*

Arroll[91]	Meta-analysis	(Reported)			Lee's	Check	
Study	Antibiotic	placebeo	MH(RR), 95%CI	SE	TDS	DS	(1)(2)(6)(7)
Taylor 1977a	72/75	25/30	1.15[0.98, 1.36]	0.084	22.89	143.10	(p)(p)(p)(f)
Taylor 1977b	51/54	25/30	1.13[0.95, 1.35]	0.09	27.66	124.44	(p)(p)(p)(f)
Todd 1984	6/26	9/24	0.62[0.26, 1.47]	0.44	15.76	5.12	(p)(p)(p)(p)
De Sutter 02	125/180	95/179	1.31[1.11, 1.55]	0.09	168.47	137.83	(p)(p)(p)(f)
Total (95% CI);	335	263	1.08[1.05, 1.33]	1.1873			

Table 5.15: *Meta-analysis of studies of outcomes of purulent rhinitis at five to eight days by Arroll 2006*

5.1.13 The Estimation of Minimum number of subjects sizes

Lee [2] advocates that validity of reported Odds ratios/Relative risks and CIs can also be checked where the original data are not available using some simple methods presented in Check (3) - Check (7) above and the assessments were based on inequility approach. These assumed the minimum total number of subjects in the study, the minimum total number of diseased subjects in the study and the minimum number of subjects in any disease group must be greater or equal to a calculated value using Lee's methods. All studies that met the minimum number requirement were assigned (P)-Pass and (F)-fail if unable to met Lee's criteria.

Checks (3),(4) and (5) are for assessment of point estimates(Odd ratios) of case control studies corresponding to (i) Minimum number of total subjects given the CI - Check 3 (ii) Minimum number in any two groups of subjects given CI -Check 4 and (iii) Minimum number

Goldberg[116]	Meta-analysis	(Reported)			Lee's	Checks	
Study	treat	cont	MH(RR)95%CI	SE	TDS	DS	(1)(2)(6)(7)
Pizzo	5/18	3/16	1.48[0.42, 5.24]	0.64	6.97	2.41	(p)(p)(p)(p)
EORTIC	11/68	14/64	0.74[0.36, 1.51]	0.37	25.06	7.48	(p)(p)(p)(p)
Schiel	2/101	0/54	2.70[0.13, 55.17]	1.54	1.65	0.42	(p)(p)(p)(p)
Wingard	4/97	5/111	0.92[0.25, 3.31]	0.66	8.83	2.30	(p)(p)(p)(p)
Goldstone	1/64	1/69	1.08[0.07, 16.88]	1.40	2.01	0.51	(p)(p)(p)(p)
Cordonnier	3/150	7/143	0.41[0.11, 1.55]	0.67	8.61	2.20	(p)(p)(p)(p)
Total	498	457	0.82[0.50, 1.34]				

Table 5.16: *Empirical antifungal therapy for patients with neutropenia and persistent fever:Meta-analysis by Goldberg et al 2008*

Tan Fao[117]	Meta-analysis	(Reported)		Lee's	Check			
Study	Steroids	placebo	MH(OR) 95%	se	TS	TE	OE	(1)(2)(3)(4)(5)
Cheng	8/85	13/43	0.24[0.09, 0.64]	0.50	63.89	15.97	3.99	(p)(p)(p)(p)(p)
Darmon	11/327	17/337	0.66[0.03, 1.42]	0.39	101.72	25.43	6.36	(p)(p)(p)(p)(p)
Francois	11/355	76/343	0.11[0.06, 0.22]	0.33	145.64	36.41	9.10	(p)(p)(p)(p)(p)
Gaussorgues	4/138	2/138	2.03[0.37, 11.27]	0.87	21.07	5.27	1.32	(p)(p)(p)(p)(p)
Ho	7/39	10/38	0.61[0.21, 1.82]	0.55	52.72	13.18	3.29	(p)(p)(p)(p)(p)
Lee	4/40	11/40	0.29[0.08, 1.02]	0.65	37.94	9.49	2.37	(p)(p)(p)(p)(p)
Total (95%) ‘	45/984	129/939	0.30[0.15, 0.58]					

Table 5.17: *Prophylactic administration of parenteral steroids for preventing airway complications after extubation in adults by Tan Fao et al(2008).SE-standard error,TS-total subjects,TE-two elements,OE-one element*

of subjects in any group combined given the CI - Check 5.

To obtain the estimate of Relative risks(RR) of a disease in relation to given exposure there are two unique checks- checks (6) and (7) - which assess (i) Minimum number of total diseased subjects (TDS) and (ii) Minimum number of diseased subjects in either group(DS).

As time permitted, we were able to carry out further checks on these individual studies by assuming that the original data were not reported. Using Lee's simple methods stated in subsection check(3)-check(7) , we used only reported CIs and Point estimates (RR/OR) of each study. The results obtained as calculated minimum sizes were used to compare reported OR/RR and CIs. We found consistency of estimates of OR while it seemed that most studies reporting (RR) failed to meet Lee's requirements.

Legg[118]	MA	(Reported)		Lee's	Checks			
Study	treat	cont	PETO(OR)95%	SE	TS	TE	OE	(1)(2)(3)(4)(5)
Corr	33/55	32/54	1.03[0.48, 2.21]	0.39	105.45	26.36	6.59	(p)(p)(p)(p)(p)
Gilbertson	33/66	41/67	0.64[0.32, 1.26]	0.35	130.89	32.72	8.18	(p)(p)(p)(p)(p)
Drummond	2/42	3/23	0.32[0.05, 2.11]	0.95	17.55	4.39	1.09	(p)(p)(p)(p)(p)
Logan	6/52	14/58	0.42[0.16, 1.11]	0.49	65.53	16.38	4.09	(p)(p)(p)(p)(p)
Walker	18/90	27/86	0.55[0.28, 1.08]	0.34	134.92	33.73	8.43	(p)(p)(p)(p)(p)
Sackley	27/53	36/47	0.34[0.15, 0.76]	0.41	93.37	23.34	5.84	(p)(p)(p)(p)(p)
Parker	106/248	56/123	0.89[0.58, 1.38]	0.22	327.22	81.81	20.45	(p)(p)(p)(p)(p)
Total (95% CI)	607	458	0.67[0.51, 0.87]					

Table 5.18: *Effects of occupational therapy on personal activities of daily living by Legg et al 2007.SE-standard error,TS-total subjects,TE-two elements,OE-one element*

Mohammed[119]	Meta-analysis	(Reported)			Lee's	Check	
Study	treatment	cont	MH(RR), 95%CI	SE	TDS	DS	(1)(2)(6)(7)
Mangat	1/16	2/17	0.53[0.05, 5.31]	1.19	2.65	0.71	(p)(p)(p)(p)
Nannini	1/19	1/16	0.84[0.06, 12.42]	1.36	2.048	0.54	(p)(p)(p)(p)
Hughes	12/28	17/24	0.61[0.37, 1.00]	0.25	35.53	15.54	(p)(p)(f)(p)
Kokturk	1/14	2/12	0.43[0.04, 4.16]	1.18	2.65	0.71	(p)(p)(p)(p)
Aggrarwai	9/50	10/50	0.90[0.40, 2.02]	0.41	19.22	5.86	(p)(p)(p)(p)
Drobina	2/60	1/50	1.67[0.16, 17.85]	1.20	2.67	0.69	(p)(p)(p)(p)
Mahajan	2/31	1/31	2.00[0.19, 20.93]	1.19	2.60	0.69	(p)(p)(p)(p)
Total	218	200	0.70[0.47, 1.04]				

Table 5.19: *Effect of nebulised magnesium sulphate upon hospital admission: systematic review and meta-analysis by Mohammed et al(2007)*

.

In the course of our investigation, we found that studies with large outcomes failed to meet Lee's minimum subject size requiremrnets, especially where the effect size is reported as relative risk. Hence, it seem that Lee [2] failed to consider the facts that when the outcomes of events are large, their standard errors tend to be small. This small value of the variance is the denominator in Lee's inequalities which divides a constant which is 61.47. The outcome of this division is large values of estimated minimum totals. Also, large event outcomes implies that the proportion surviving in the exposed (S = a/(a+c)) will be very small. As discussed above the only logical explanation is that Lee's approach is generally inappropriate in studies reporting relatives risks when the number of events is large compared to the number at risk. Where this happens Lee's checks are often failed by studies even though the relative risk

Barr[120]	MA	(Reported)		Lee's				Check
Study	tiotr	cont	MH(OR), 95%	SE	TS	TE	OE	(1)(2)(3)(4)(5)
Beeh	180/1236	80/403	0.69[0.51, 0.92]	0.15	706.39	176.59	44.15	(p)(p)(p)(p)(p)
Brusasco	129/402	156/400	0.74[0.55, 0.99]	0.15	711.63	177.91	44.48	(p)(p)(p)(p)(p)
Casaburi	198/550	156/371	0.78[0.59, 1.02]	0.14	820.40	205.10	51.28	(p)(p)(p)(p)(p)
Dusser	250/500	308/510	0.66[0.51, 0.84]	0.13	987.43	246.86	61.71	(p)(p)(p)(p)(p)
Niewoehner	255/914	296/915	0.81[0.66, 0.99]	0.10	1495.50	373.87	93.47	(p)(p)(p)(p)(p)
Verkindre	0/46	2/54	0.23[0.01, 4.83]	1.58	6.44	1.61	0.4023	(p)(p)(p)(p)(p)
Total(95% CI)	3648	2653	0.74[0.66, 0.83]					

Table 5.20: *Tiotropium for stable chronic obstructive pulmonary disease: meta-analysis by Barr et al 2006.SE-standard error,TS-total subjects,TE-two elements,OE-one element*

Chang[121]	MA	(Reported)		Lee's	Check			
Study	Treatment	cont	MH(OR), 95%	SE	TS	TE	OE	(1)(2)(3)(4)(5)
Ours	7/8	9/9	0.26[0.01, 7.43]	1.69	5.63	1.41	0.35	(p)(p)(p)(p) (p)
Kijander	7/9	12/12	0.12[0.01, 2.86]	1.44	7.69	1.92	0.48	(p)(p)(p)(p)(p)
Ehere	2/5	4/6	0.33[0.03, 3.93]	1.24	10.34	2.59	0.65	(p)(p)(p)(p)(p)
Total (95%)	22	27	0.24[0.04, 1.27]					

Table 5.21: *Meta-analysis of Gastro-oesophageal reflux interventions for chronic cough associated with gastro-oesophageal reflux by Chang et al(2006).SE-standard error,TS-total subjects,TE-two elements,OE-one element*

is calculated correctly. This seems to be a technical error. We recommend that further investigation be carried out to establish what could be done to rectify this.

Table 5.22 presented below shows the results of the calculated minimum subjects sizes listed at the right-hand column. We first illustate how these checks were conducted using Augustine et al 2004 published in Rabindranath et al

Example:- $*$**Augustine 2004** $A = 27$ $B = 28$ $C = 13$ $D = 12$ $RR = 0.96$ $CI = [0.72, 1.30]$ $se = 0.15$ $TDS = 57.95$ $TD = 44.01$

(iv) Check (6) - Minimum number of total diseased subjects - (TDS)

$A + B > A + (BC(B + D)/D(A + C)) \geq 61.47C/Q^2(A + C)$

More generally

$A + B \geq 61.47S/Q^2$

Where $S = C/(A + C) = 13/(13 + 27) = 13/40 = 0.33$

$Q = ln(r_u/r_l) = ln(1.30/0.72) = ln(1.81) = 0.59$

$Q^2 = (0.59)^2 = 0.35$

therefore $A + B \geq 61.47S/Q^2$

$A + B \geq 61.47(0.33)/0.35 = 57.95$

$27 + 28 = 55$ which is not greater than or equal to calculated $TSD = 57.95$, hence, we assigned fail (f) to check(6) .

Where TSD is the Minimum number of total diseased subjects(Check(6)).

(iv) Check (7)-Minimum number of diseased subjects in either exposure group

$V_1 \geq 1/z$, where z is either A or B

$z \geq 15.3664/Q^2$

A or B$\geq 15.3664/0.35 = 44.0$

A or B ≥ 44.0

where A = 27 or B = 28 is less than the calculated $DS = 44.0$, we assigned fail (f) to check(7) Where DS is the Minimum number of diseased subjects in either exposure group(Check (7))

Meta-analyses by Rabindranath et al 2007 [105] in Table 5.2, Terplan [106] in Table 5.3 and Alonso et al (2008) [107] in Table 5.4 published in Cochrane library and Arroll 2004 [115] in Table 5.14 and Arroll 2006 [91] in Table 5.15 published in Paper Journal show that the total reported number of cases in both exposure groups and the number of cases in

either exposure groups given confidence intervals, were less than the results obtained for the minimum total number of diseased subjects in the groups using formula in checks (6) and (7). Studies reporting relative risks in both Paper-Base Journals and Cochrane reviews are more likely to disagree with the Lee's requirement on minimum number of diseased subjects in either exposure groups given the CI, than those repoting odd ratios. We also note that the studies that did not meet the checks have very large outcomes, those with very few outcomes scaled all the tests.

5.1.14 Mabinary.sas Checks on Reported CIs

Finally, different versions of SAS.macros developed by Weir and Senn [3] were employed to check the reliability and validity of the reported estimates for all studies recorded by the authors. Apart from using Mapeterlee.sas.macro, Mabinary.sas macro was also used to check accuracy of reported estimates. Data input into the macro consists of study name ,treatment-event, control-event, total number in the treatment and control groups, Methods e.g MH-RR, MH-OR, Peto-OR etc and their 5% or 1% level of significance. The results obtained using these macros are consistent with those obtained using Peter Lee simple methods as shown in Table 5.12 to Table 5.21 in the previous sections.

5.1.15 Summary of Result

(1)SUMMARY OF INDIVIDUAL STUDIES THAT PASS EACH OF THE CHECK

Results obtained by carrying out error checks using simple methods designed by Peter Lee [2] on reported estimates of included individual studies in twenty selected meta-analyses, ten each from Cochrane and Non-Cochrane reviews, are summarised in Table 5.23 and Table 5.24. The tables show the number (proportion) of included studies that passed each of the checks as outlined in the previous subsections of this chapter. The checks were carried out depending on whether the authors reported either odds ratios or relative risks. Table 5.23 reveals that Odds ratios reported in both Paper base journals and Cochrane reviews satisfied all Lee's requirements . Reported confidence intervals for Odd ratios and its corresponding minimum number of subjects in any group are consistent with Lee's criteria in both Cochrane reviews and Paper base journals. However, studies reporting relative risks in Paper-based Journals are more likely to be consistent with Lee's criteria than those reporting relative risks in Cochrane database when we consider the minimum number of diseased subjects sizes in a given group as shown in Table 5.24. A possible explanation of this seemingly consistency

of the Paper-based journals is that they tend to have smaller number of outcomes than that of the cochrane reviews.

For reported relative risks, both Cochrane and Non-Cochrane reviews reported (100%) confidence intervals that were consistent with Lee's checks. Nevertheless, reported minimum number of diseased subjects in either exposed groups for studies reporting relative risks, seems to be more consistent in Paper based journals with about 91% of such studies meeting Lee's criterion as against 63% of Cochrane libraries. Majority of studies published in the Cochrane reviews are less likely to satisfy the Lee's requirements on check (6) and (7) compared with the Paper-based Journals. We noted a general pattern among studies reporting relative risk that failed the minimum subject requirements of Lee. These studies tend to have very large event outcomes. Studies with smaller or rare event outcomes are more likely to meet Lee's minimum subjects requirements

(2) SUMMARY OF INDIVIDUAL STUDIES THAT PASS ALL THE CHECKS

In Cochrane reviews, there were a total of eighty-three (83) individual studies contained in the selected 10 meta-analyses and a total of Sixty-eight (68) individual studies in the selected 10 Paper base meta-analyses. Table 5.25 was constructed to illustrate the number of studies that passed all the Lee's checks in both Cochrane and Non-Cochrane reviews. There were 67 (81%) out of 83 of included studies of meta-analysis in Cochrane reviews that satisfied all the Lee's requirements compared to 60(88%) out of 68 studies of the Paper base Journals. Conversely, about 19% and 12% of all included studies in Cochrane and Paper base journals respectively, fall short of at least one of Lee's requirement. However, as noted earlier, failure to meet Lee's minimum total requirements for studies reporting relative risks does not imply that the reported estimates were wrong but that Lee's checks seems inappropriate for such studies with large outcomes.

Table 5.22: *Estimation of Minimum number of subjects sizes*

Rabindranath	calculated	Peter Lee		Check
Study(RR)	SE	TDS	DS	(6)(7)
Augustine 2004	0.15	57.95	44.01	(f)(f)
Gasparovic 2003	0.14	55.98	48.52	(p)(f)
Mehta 2001	0.14	71.57	51.82	(p)(p)
Noble 2006	0.09	91.13	111.83	(f)(f)
Uehlinge 2005	0.18	63.39	29.98	(f)(p)
Vinsonneau 2006	0.073	249.86	189.53	(f)(f)
Terplan	calculated	Peter Lee		Check
Study	SE	TDS	DS	(6)(7)
Elk 1998	0.34	5.72	8.58	(p)(f)
Haug 2004	0.10	51.87	97.25	(p)(f)
Jones 2001	0.05	87.43	342.40	(f)(f)
Mulins 2004	0.21	44.64	21.70	(f)(f)
O'Neil 1996	0.08	94.18	158.08	(f)(f)
Silveman 2001	0.34	15.68	24.78	(p)(p)
Svikis 1997	0.20	32.22	8.71	(p)(p)
Alonso's	calculated	Peter Lee		Checks
Study	SE	TDS	DS	(6)(7)
Himmelfarb	0.15	60.25	41.71	(p)(p)
Kurtal 1995	0.14	90.68	52.89	(p)(f)
Micheal 1995	0.26	21.34	14.67	(f)(f)
Romao 1999	0.23	29.51	18.44	(f)(f)
Schifffl 1994	0.29	11.01	11.93	(p)(p)
Valeri 1996	0.20	24.35	25.37	(p)(f)
Woo 2002	0.11	27.18	78.13	(p)(f)
Kurtal 1995	0.25	30.28	15.77	(p)(p)
Arroll 2006	calculated	Peter Lee		Check
Study	SE	TDS	DS	(6)(7)
Taylor 1977a	0.084	22.89	143.10	(p)(f)
Taylor 1977b	0.09	27.66	124.44	(p)(f)
Todd 1984	0.44	15.76	5.12	(p)(p)
De Sutter 02	0.09	168.47	137.83	(p)(f)
Arroll 2004	calculated	Peter Lee		Check
Study	SE	TDS	DS	(6)(7)
Cederlof	0.21	21.83	23.65	(p)(f)
Dieppe	0.97	0.72	1.07	(p)(p)
Friedman	0.17	14.57	30.97	(p)(f)
Gaffney	0.17	28.51	33.26	(p)(p)
Ravaud	0.36	11.07	7.69	(p)(p)
Smith	0.22	27.21	19.89	(p)(p)

Table 5.23: *Summary of Consistency of OR and RR and CIs with 2 ×2table*

Checks(OR)	No of Jou Rev	No of Coch Rev
Check(1)	22(100%)	42(100%)
Check(2)	22(100%)	42(100%)
Total	22	42
Checks(RR)	Journals	Cochrane
Check(1)	46(100%)	41(100%)
Check(2)	46(100%)	41(100%)
Total	46	41

Table 5.24: *Summary of Minimum number of subjects sizes without 2×2table*

Checks(OR)	No of Jou Rev	No of Coch Rev
Check(3)	22(100%)	42(100%)
Check(4)	22(100%)	42(100%)
Check(5)	22(100%)	42(100%)
Total	22	42
Checks(RR)	Journals	Cochrane
Check(6)	44(96%)	31(75%)
Check(7)	42(91%)	26(63%)
Total	46	41

Table 5.25: *No. of included studies that passed all the Lee's checks on reported estimates*

Database	Pass	Fail	Total
Journals	60(88%)	8(12%)	68
Cochrane	67(81%)	16(19%)	83
Total	126	25	151

Chapter 6

DISCUSSION AND CONCLUSION

6.1 SUMMARY

This research work focus on developing a quality instrument to assess meta-analysis in terms of actual statistical analysis. The history of meta-analysis was traced from 1976 when Glass first coined the term "meta-analysis". The use of meta-analysis can be traced to the 1960s and has since then increased rapidly and widely as it serves its purpose in the healthcare sector, government policy and decision making. The combination of individual trials to obtain the statistical results and conclusions has lessened the huge search on medical literature by clinicians who use it daily to interpret their interventions on patients' treatment. Three types of meta-analysis were summarised by Senn [52] and its necessary steps for performance.

We obtained data for review from Cochrane and Non-Cochrane database. The objectives / aims of this research were to construct a checklist that would assess the quality of meta-analysis in terms of statistical analysis and to measure its validity and reliability. We also used Lee's methods to check errors on reported estimates.

Chapter 2 summarises the trend of checklist development from the 1960's and reviewed guidelines published to aid preparation of a valid checklist. In our review, we listed eight

instruments from 1996 till date that assessed quality of reporting, scientific quality and methodological quality but could not obtain any that prepared checklist to assess statistical quality. We also reviewed the use, benefits and applications of checklists.

In Chapter 3, we developed a tool consisting of 24-items that was used to measure effectively the quality of meta-analysis in terms of actual statistical analysis. The inclusion criteria of the studies among others, include selecting meta-analysis published from 2000-2008 in each database of Cochrane and Non-Cochrane.

We selected eligible studies that were published in English, full text and with search strategy that is very comprehensive. Error bound for the estimation of sample size was obtained and we established a sample size of 100, using the knowledge that variance of proportion attains its maximum when is p=0.5 and estimated error bound to be 0.1.

The sample size (n=100) was distributed equally to Cochrane and Non-Cochrane database where samples were selected using simple random sample conducted using R statistical package. Non-Cochrane database (Web of Knowledge and Medline) shared proportionally their allocated sample size (n=50). Each selected simple random number was used to trace and retrieve studies labelled with that number from the electronic database. We typed words in the search topic as-"meta-analysis of randomised controlled trials" OR "metaanalysisrandomised controlled trials " OR "meta analysis randomised controlled trials" published in English, in full text and an article.

There were 515, 507 and 130 results in Cochrane, Web of Knowledge and Medline respectively. During the assessment of quality of trials, some trials were excluded because they did not meet the stated inclusion criteria. Therefore, eligible studies that were fit for task were 42, 31, 10 studies in Cochrane, Web of Knowledge and Medline respectively. The two paper-base journals (Web of Knowledge and Medline) were merged to obtain a single sample. Subsequently, pilot study was conducted to check the validity and practicability of the constructed checklist.

Chapter 4 emphasizes the analysis of data obtained from the quality assessment of selected studies and results. Firstly, we evaluated the number of authors and number of trials in each review. The number of trials in each meta-analysis ranges from 2 to 51, and that of authors were between 2 and 20. Stem and leaf plot was constructed to determine their distribution. The result showed positive skewness in the two groups, and the median number of trials and authors are greater in the Cochrane than the paper-based journals.

We used the new checklist to assess quality of meta-analysis in terms of actual statistical analysis in both Cochrane and Non-Cochran studies. In Table 4.2, the items which measure the statistical quality of each study are (item 10) to (item 22). We calculated Fisher exact test (two-sided) using SAS and R statistical packages to obtain the proportions and to compare the responses to each item in the checklist. Results obtained in the pilot study were consistent with that obtained when the checklist was modified. The results show that there is no statistically significant difference between Cochrane and Non-Cochrane reviews in most of the checklist items because most estimated p-values were greater than 0.05 except for item nine(9) which is statistically significant. This item suggest that paper-based journals occasionally used OQAQ and QUOROM checklists to assess scientific and reporting quality of their meta-analysis than Cochrane journals.

There were no simple double counting of studies found in the Cochrane and Paper-base journal but we identified four examples of double counting of some aspects of studies in meta-analysis published in Cochrane reviews (Sander et al [83], Rendon et al [85], Leslie et al [89] and Urquahart et al [87]). These studies counted control arms more than once. One Paper journal split control arms into equal half among the treatment groups (Arrol et al 2006 [91] and Yin et al [93]).

During the course of review of Cochrane library, we identified meta-analysis by (Marinho et al [97] and Green et al [98] that imputed standard deviation as a result of missing data and non-response from the authors when contacted. One meta-analysis did not include

the imputed data in the analysis while the other imputed using linear regression of log of standard deviation on log mean.

In chapter 5, the precisions of reported studies were assessed using mapeterlee SAS macro version 8.2 and mabinary version 1.0. These SAS macro versions were developed by Weir and Senn [3]. Mapeterlee sas macro requires input dataset of study name and standard error while mabinary sas macro requires input dataset of study name, treatment-event, total number in treatment groups, control-event and total number in control groups. Standard error for each study was calculated using Peter Lee's method. Twenty studies (ten from each database) were selected from Cochrane and Non-Cochrane databases.

Peter Lee's method was employed to check for errors in the reported estimates of relative risks, odds ratios and confidence intervals in the selected studies from Cochrane and Paper journal reviews.

Some studies included in the meta-analysis reported zero adverse events either in the treatment or control groups or both which led to a disparity between the calculated results and the estimates reported by the authors. The addition of a continuity correction factor of 0.5 to each cell of the studies with zero events took care of the disparities. Mabinary sas macro was also used to assess and check the validity of reported odd ratios, relative risks and confidence intervals on both meta-analyses. The results obtained using the macros are consistent with the original reported results in most studies.

Reported number of diseased subjects in either exposure groups for studies reporting relative risks, seems to be more consistent in Paper based journals with about 91% of such studies meeting Lee's criterion as against 75% of Cochrane libraries. Studies published in the Cochrane reviews are less likely to satisfy the Lee's requirements on expected total number of events and number of events in either exposure group (check (6) and (7)) compared with the Paper-based Journals. We note however, that studies reporting relative risks which do not comply with Lee's minimum subject number requirements are not wrong, but it does

seem that Lee's checks are inappropriate or unliable for such studies with large outcomes. Most of the studies that did not meet Lee's criteria have large events.

Also, studies reporting relative risks in both Paper-Base Journals and Cochrane reviews are more likely to disagree with the Lee's requirement on minimum number of diseased subjects in either exposure groups given the CI, than those repoting odd ratios. The reason for this is as explained in the preceding paragraph. Studies reporting relative risks and have large outcomes are more likely not to meet this requirement. this calls for caution in the use of Lee's criteria for assessment of precision of such studies.

Overall, in Cochrane, there were a total of 83 individual studies included in the selected 10 meta-analyses of randomised controlled trials but only 67(81%) out the 83 individual studies met all the Peter Lee's requirements . Also, for the Journal reviews, there were a total of 68 individual datasets included in the selected 10 meta-analyses of randomised controlled trials but only 60(88%) of them met all the requirements. All the studies that did not meet Lee's criteria reported relative risks and have relatively large outcomes.

6.1.1 FURTHER WORK

For this checklist to be adopted as a valid tool for assessing the quality of meta-analysis, it needs be validated by two or more professionals. This validation process of the checklist is left as future work. Further investigation on the inability of Peter Lee's methods to adequately cover the assessment of Relative Risks when the number of events is large compared to the number at risk is recommended.

6.2 CONCLUSION

We have successfully developed a checklist instrument that is very unique in its content. It is designed to assess not only the statistical quality of meta-analysis but also the reporting and

scientific quality of meta-analysis. The checklist is designed not only for the meta-analyst but also for study authors and review editors. Quality of statistical information provided in the individual studies has direct relation on the quality of meta-analysis produced by combining such studies. Study authors are encouraged by this instrument to include as much statistical information as possible because information rich studies may yield high quality review and meta-analysis if the information is appropriately utilized. Each item of the instrument is intended to provoke deep consideration and reflection on an aspect of study aggregation. Items 1 to 9 of the checklist address the quality of review while items 10 to 24 deals with the actual statistical quality of studies and meta-analysis. For instance, item ten reminds the reviewer to specify the statistical method used in combining studies. Subsequent two items encourages the meta-analyst to think about the reason(s) for using the stated method and to justify that such method is adequate or valid for such analysis. Validity of statistical methods comes from its suitability with the data and being able to satisfy model assumptions. Aggregation of studies requires a careful consideration of study compatibility which is only determined from information on study design, participants, interventions, methodology and outcomes. Where this information is missing, effective assessment of meta-analysis combining such studies is hampered. Also, assessment of adequacy of statistical methods used to analyse combined studies may not be possible. We advocate through the items in the checklist that studies being combined have adequate information to enable not only assessment of compatibility but also to minimize need for imputation. Double counting of studies or some aspects of studies may distort results and conclusions from meta-analysis. We have included these two items in the checklist to sensitize meta-analysts on the need to look out for repeated studies and or arms of studies which may impact on precision of estimates. Choice of weights used for combining studies is an important element in study aggregation. This can make or mar the results of meta-analysis depending on appropriateness or inappropriateness of weights. Also information on software used in

analysis exposes its suitability or otherwise for such analysis. We strongly recommend the routine use of this checklist by study authors, readers, editors and meta-analysts to identify not only scientifically sound but also statistically strong reviews. However, we accept that a validation study is needed. This should be a collaborative project involving different assessors and hence is not suitable for work leading to an MSc. We hope that our work might encourage others to take this further.

We used the developed checklist to compare statistical quality of selected publications in Cochrane and Non-Cochrane reviews. The results revealed no statistically significant difference between the reporting and methodological quality of the Cochrane and non-Cochran in most of the checklist items because most estimated p-values were greater than 0.05 (see Table 4.2). The only exception is item (9) which indicates that the non-Cochrane reviews were more likely to assess scientific and reporting quality using OQAQ and QUOROM checklists than Cochrane reviews. Both reviews performed poorly in the explicit reportage of included study designs making assessment of compatibility of combined studies difficult.

Double counting of some aspects of studies were found in Cochrane reviews. Analysis suggests that studies reporting odd ratios are likely to be consistent with Lee's checks than those reporting relative risks. We also showed that Peter Lee's checks involving large number of events cannot be relied on to assess the quality of studies reporting relative risks.

Bibliography

[1] Eric Roehm MD FACC. Meta advocacy problems meta-analysis: Beware of meta-analyses bearing false gifts. http://www.improvingmedicalstatistics.com/Meta 20Beware 20gifts.htm,Online; accessed 22 Jan 2010. [Online; accessed 22 Jan 2010].

[2] LEE PETER. Simple methods for checking for possible errors in reported odds ratios, relative risks and confidence intervals. *STATISTICS IN MEDICINE; Statist. Med.*, 18:1973–1981, 1999.

[3] Weir and Senn S. Sas macro. http://www.senns.demon.co.uk/SAS%20Macros/SASMacros.html, 2008.

[4] Kulik Jame, A Kulik, and Chen lin C. meta-analysis: Historical origins and contemporary practice. *Paper presented at the annual meeting of the American Educational Research Association,New Orleans, LA.*, 1988.

[5] Glass GV. Primary secondary and meta-analysis of research. *Educ Researcher*, 10:3–8, 1976.

[6] Pearson K. Report on certain enteric fever inoculation statistics. *BMJ*, 3:1243–6, 1904.

[7] Fisher RA. Statistical methods for research workers. *Oliver and Boyd*, pages 1–9, 1932.

[8] Cochran WG. Problems arising in the analysis of a series of similar experiments. *J Roy Stat Soc*, 4(Suppl):102–18, 1937.

[9] Rosenthal R. Meta-analytic procedures for social research. applied social science research methods series. *Beverly Hills Sage*, 1984.

[10] Hedges L V and Olkin I. Analyses, reanalyses, and meta-analysis. *Contemporary Education Review*, 1:157–165, 1982.

[11] Chalmers TC, Matta RJ, Smith H, and Kunzler AM. Evidence favoring the use of anticoagulants in the hospital phase of acute myocardial infarction. *N Engl J Med*, 297:1091–6, 1977.

[12] Moher D and Berlin J. Improving the reporting of randomized controlled trials.in a. maynard i. chalmers (eds.), non-randomr eflections on health services research. *BMJ*, pages 250–271, 1997.

[13] DARE. Database of systematic reviews. *The Cochrane Library.*, Retrieved from http://agatha.york.ac.uk/darehp.htm. Oxford:, 2001.

[14] Moher D.and Olkin I. Meta-analysis of randomized controlled trials.aconcern for standards. *Journal of the American Medical Association*, 274(128 Evaluation and the Health Professions / March 2002):1962–1964, 1995.

[15] Cochrane Collaboration. Research for development. http://www.dfid.gov.uk/r4d/news.asp?articleID=50682. [Online; accessed 18 Dec 2010].

[16] Cochrane AL. Cochrane reviews. http://www.cochrane.org/cochrane-reviews . [Online; accessed 02 Mar, 2011].

[17] Cochrane Collaboration. Overlake hospital medical center. http://medlib.overlakehospital.org/content/library/Databases.htm . [Online; accessed 10 Feb 2011].

[18] Iain K Crombie and Huw T O Davies. What is series. http://www.medicine.ox.ac.uk/bandolier/painres/download/whatis/Meta-An.pdf Iain K Crombie and Huw T O Davies, 2009.

[19] Senn Stephen. The many modes of meta. *drug information journal*, 34:537, 2000.

[20] International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use. Statistical principles for clinical trials. *Stat Med*, 18:1905–1942, 1999.

[21] Anne Whitehead. *Meta-Analysis of Controlled Clinical Trials*. 2002.

[22] DeCoster J. Meta-analysis notes. http://www.stat-help.com/notes.html, 2004. [Online; accessed 23 september 2009].

[23] Anthony Delaney, Sean .M. Bagshaw, Andre Ferland, Braden Manns, Kelvin .B. Laupland, and Christopher .J Doig. A systematic evaluation of the quality of meta-analyses in the critical care literature. *Crtical Care*, 9(5):R575–R582, 2005.

[24] Goldman L and Feinstein AR. Anticoagulants and myocardial infarctionthe problems of pooling, drowning, and floating. *Ann Intern Med*, 90:92–94, 1979.

[25] Fleiss IL and Gross AJ. Meta-analysis in epidemiology, with special reference to studies of the association between exposure to environmental tobacco smoke and lung cancer: A critique. *J Clin Epidemiol*, 44:127–139, 1991.

[26] Thomson SG and Pocock S. Can meta-analysis be trusted? *Lancet*, 338:1127–1130, 1991.

[27] Felson D. Bias in meta-analysis. *J Clin Epidemiol*, 45:885–892, 1992.

[28] Oxman AD and Guyatt GH. Validation of an index of the quality of review articles. *J Clin Epidemiol 1991*, 44:1271–1278, 1991.

[29] Streiner D L. Agreement among reviewers of review articles. *J. Clin Epidemiol 1991*, 44:91–98, 1991.

[30] Mulrow C D. The medical review article : state of the science. *Am intern Med*, 106:485–488, 1987.

[31] Oxman A D and Guyatt G H. Guidelines for reading literature reviews. *Canadian Med. Asocciation Journal*, 138:697–703, 1988.

[32] Boluty et al. the state of an evidence on acute asthma management in children. *PEDIATRICS*, 120(6):1334–1343, 2007.

[33] Basil G Bereza, Mrcio Machado, and Thomas R Einarson and. assessing the reporting and scientific quality of meta-analysis. *www.theannals.com,*, page DOI 10.1345/aph.1L204., 2008.

[34] Moher D, Cook D J., Eastwood S, Olkin I, and Rennie D.and Stroup D. F. Improving the quality of reports of meta-analyses of randomised controlled trials: The quorom statement. *Lancet*, 354:1896–1900, 1999.

[35] Junehua et al. methodology and reporting quality of systematic reviews/ meta-analysis review of traditional chinese medicine;. *the journal of alternative and complmentary medicine*, 13(18):797–806, Oct 2007.

[36] Schulz KF, Chalmers Iand Hayes RJ, and Altman DG. Empirical evidence of bias: dimensions of methodological quality associated with estimates of treatment effects in controlled trials. *JAMA*, 273:408–12, 1995.

[37] Oakes M. *Statistical inference: A commentary for the social and behavioral sciences*. New York: John Wiley, 1986.

[38] Wachter K W. Disturbed by meta-analysis? *Science*, 241((4872)):1407–8, 1988.

[39] Berk R. A. Regression analysis: A constructive critique.thousand oaks. *CA:Sage*, 2004.

[40] Berk R A and D A Freedman. Statistical assumptions as empirical commitments. in law, punishment, and social control: Essays in honor of sheldonmessinger, 2nd ed.,ed.t. g. blomberg and s. cohen. http://www.stat.berkeley.edu/census/berk2.pdf.:235–54, 2003.

[41] Hedges L V. Directions for future methodology. in the future of meta-analysis, ed. k.w.wachter and m. l. straf. *New York: Russell Sage Foundation*, pages 11–26, 1990.

[42] Bailar J C III. The practice of meta-analysis. *J Clin Epidemiol*, 48:149–157, 1995.

[43] Macarthur C, Foran P J, and Bailar J C. Qualitative assessment of studies included in a meta-analysis: Des and the risk of pregnancy loss. *J Clin Epidemiol*, 48:739–747, 1995.

[44] Shapiro S. Meta-analysis meta-analysis. *Am J Epidemiol*, 140:771–778, 1994.

[45] Shapiro S. Is meta-analysis a valid approach to the evaluation of small effects in observational studies? *J Clin Epidemiol*, 50:223–229, 1997.

[46] Kerlikowske K, Grady D, Rubin SM, Sandrock C, and Ernster VL. Efficacy of screening mammography: a meta-analysis. *JAMA*, 273:149–154, 1995.

[47] Mart CR, Hendrick RE, Rutledge JH III, and Smith RA. Benefit of mammography screening in women ages 40 to 49 years: current evidence from randomized controlled trials. *Cancer*, 75:1619–1626, 1995.

[48] Rosenfeld RM and Post JC. Meta-analysis of antibiotics for the treatment of otitis media with effusion. *Otolaryngol Head Neck Surg*, 106:378–386, 1992.

[49] Senn Stephen. overstating the evidence- double counting in meta-analysis and related problem. *BMC medical research methodology*, 9, 2009.

[50] Williams RL, Chalmers TC, Stange KC, Chalmers FT, and Bowlin SJ. Use of antibiotics in preventing recurrent acute otitis media and in treating otitis media with effusion: a meta-analytic attempt to resolve the brouhaha. *JAMA*, 270:1344–1351, 1993.

[51] Sutton A. J.and Abrams K. R., Jones D. R., Sheldon T. A, and Song F. *Methods for Meta-Analysis in Medical Research. New York:*. John Wiley and Sons, 2000.

[52] Senn Stephen J. Overstating the evidence double counting in meta-analysis and related problems. *BMC Medical Research Methodology*, 9:10, 2009.

[53] Wikipedia. Checklist — wikipedia, the free encyclopedia. 2000,.

[54] MindFire Solutions. Checklist driven testing- an overview. www.mindfiresolutions.com/mindfire/Checklists Overview, 2003. [Online; accessed 28 Nov, 2009].

[55] Mahon WA and Daniel EE. A method for the assessment of reports of drug trials. *Can Med Assoc J*, 90:565–569, 1964.

[56] Reiffenstein R, Schiltroth A, and Todd D. Current standards in reported drug trials. *Can Med Assoc J*, 99:1134–5, 1968.

[57] Freiman J, Chalmers T, Smith H, and et al. The importance of beta, the type ii error and sample size in the design and interpretation of the randomized control trial. survey of 71 "negative"trials. *N Engl J Med*, 299:690–4, 1978.

[58] DerSimonian R, Charette J, McPeek B, and et al. Reporting on methods in clinical trials. *N Engl J Med*, 306:1332–7, 1982.

[59] *Policy research Incorporated Literature Reviews validations Procedure manual*. Baltimore Policy Research inc, 1979.

[60] Mullen PD and Ramirez G. *Information Synthesis and Meta-analysis, In Advances in Health Educations and promotion*, volume 2. Greenwich Connecticut, Jai Press, 1986.

[61] Light RJ and Pillemar DB. *Summing up: the Science of Reviewing Research*. Cambridge, Massachusetis: Harvard University Press, 1984.

[62] Sacks HS, Berrier J, Reitman D, Ancona-Berk VA, and Chalmers TC. Meta-analysis of randomized controlled trials. *N Engl J Med*, 316:450–5, 1987.

[63] Alejandro R. Jadad, R. Andrew Moore, Dawn Carroll, Crispin Jenkinson, D. John, M. Reynolds, David J Gavaghan, and Henry J. McQuay D. Assessing the quality of reports of randomized clinical trials: Is blinding necessary? controlled clinical trials. *Elsevier Science Inc.*, 17:1–12, 1996.

[64] Fisher M, Friedman SE, and Strauss B. The effects of blinding on acceptance of research papers by peer review. *JAMA*, 272:143–6, 1994.

[65] Moher D, Soeken K, Sampson M, Campbell K, Ben Perot L, and Berman B. Assessing the quality of reports of systematic reviews in pediatric complementary and alternative medicine. *BMC,Pediatr 2doi: 10.1186/1471-2431-2-3. [PubMed]*, 2, 2002.

[66] Blettner M, Sauerbrei W, Schlehofer B, and et al. Traditional reviews, meta-analyses and pooled analyses in epidemiology. *Int J Epidemiol*, 28:1–9, 1999.

[67] Stroup D F, Berlin J A, Morton S C, and et al. for the MOOSE Group. Meta-analysis of observational studies in epidemiology: A proposal for reporting. *JAMA*, 283(15):2008–2012, 2000.

[68] Schlesselman JJ. Risk of endometrial cancer in relation to use of combined oral contraceptives. *Hum Reprod*, 12:1851–1863, 1997.

[69] Publication bias in clinical research. *Lancet*, 337:867–872, 1991.

[70] McShane L M, Altman D G, Sauerbrei W, Taube S E, Gion M, and Clark G M. Reporting recommendations for tumor marker prognostic studies (remark). *Journal of the National Cancer Institute*, 97(16):1180–1184, 2005.

[71] Altman DG, De Stavola BL, Love SB, and Stepniewska KA. Review of survival analyses published in cancer journals. *Br J Cancer*, 72:511–8, 1995.

[72] GOLDER S, MATHER L, WRIGHT K, and GLANVILLE J. What's on the web: recommendations for reporting research. *J Health Serv Res Policy*, 10(2):124–5, 2005.

[73] Booth A. "brimful of starlite": Toward standards for reporting literature searches. *Journal of the Medical Library Association*, 94(4):421–9, e205, 2006.

[74] Oxman AD. Checklists for review articles. *BMJ*, 309:648–651, 1994.

[75] Shea B.J, Grimshaw J.M, Wells G.A., Boers M., Andersson N., Hamel C., Porter A.C.and Tugwell P., Moher D., and Bouter L.M. Development of amstar: A measurement tool to assess the methodological quality of systematic reviews. *BMC Medical Research Methodology*, 7(10), 2007.

[76] David Moher, Alessandro Liberati, Jennifer Tetzlaff, and Douglas G. Altman. The prisma group preferred reporting items for systematic reviews and meta-analyses: The prisma statement. *BMJ*, page 339:b2535, 2009.

[77] Web of Knowledge. Citation and journal database. http://www.wok.mimas.ac.uk/. [Online; accessed 4 Jan, 2011].

[78] Weisstein Eric W. Fisher's exact test. *MathWorld–A Wolfram Web Resource*, http://mathworld.wolfram.com/FishersExactTest.html.

[79] Weisstein Eric W. Stem-and-leaf diagram. *From MathWorld–A Wolfram Web Resource*, http://mathworld.wolfram.com/Stem-and-LeafDiagram.html.

[80] Oxman et al. Guidelines for reading literature reviews. *CMAJ*, 138, 1988.

[81] Macfadyen CA, Acuin JM, and Gamble CL. Topical antibiotics without steroids for chronically discharging ears with underlying eardrum perforations. *Cochrane Database of Systematic Reviews*, Issue 4(Art. No.: CD004618. DOI:10.1002/14651858.CD004618.pub2), 2005.

[82] Perry R Hinton, Charlotte Brownlow, Isabella McMurray, and Bob Cozens. *SPSS Explained*. 2004.

[83] Saunders DH, Greig CA, Young A, and Mead GE. Physical fitness training for stroke patients. *Cochrane Database of Systematic Reviews*, Issue 1, Art. No.: CD003316. DOI: 10.1002/14651858.CD003316.pub2, 2004.

[84] Pohl M, Mehrholz J, Ritschel C, and Ruckriem S. Speed-dependent treadmill training in ambulatory hemiparetic stroke patients: a randomized controlled trial. *Stroke*, 33(2):5538., 2002.

[85] Martin-Rendon E, Brunskill S, Dore C, Hyde C, Mathur A, Stanworth S, and Watt S. Stem cell treatment for acute myocardial infarction. *Cochrane Database of Systematic Reviews*, issue 4(Art. No.: CD006536. DOI:10.1002/14651858.CD006536.pub2), 2008.

[86] Meluzin J, Mayer J, Groch L, Kanousek, Hornacek I, and Hinomaz O et al. Autologous transplantation of mononuclear bone marrow cells in patients with acute myocardial infarction: the effect of the dose of transplanted cells on myocardial function. *American Heart Journal*, 152(5):975.e915., 2006.

[87] UrquhartD, Hoving JL, Assendelft WJJ, RolandM, and van TulderMW. Antidepressants for non-specific low back pain. *CochraneDatabase of SystematicReviews*, Issue 1(Art.No.:CD001703.DOI: 10.1002/14651858.CD001703.pub3.), 2008,.

[88] Atkinson JH, Slater MA, Wahlgren DR, Williams RA, Zisook S, Pruitt SD, and et al. Effects of noradrenergic and serotonergic antidepressants on chronic low back pain intensity. *Pain*, 83:137–45, 1999.

[89] Leslie K, Clavisi O, and Hargrove J. Target-controlled infusion versus manually-controlled infusion of propofol for general anaesthesia or sedation in adults. *Cochrane Database of Systematic Reviews*, Issue 3(Art. No.: CD006059.DOI: 10.1002/14651858.CD006059.pub2.), 2008.

[90] Struys MM, Versichelen LF, O Thas, L Herregods, and G Rolly. Comparison of computer-controlled administration of propofol with two manually-controlled infusion techniques. *Anaesthesia*, 52:4150, 1997.

[91] Arroll B and T Kenealy. Are antibiotics effective for acute purulent rhinitis? systematic review and meta-analysis of placebo controlled randomised trials. *BMJ*, 333:279, 2006.

[92] Taylor B, Abbott GD, Kerr MM, and Fergusson DM. Amoxycillin and co-trimoxazole in presumed viral respiratory infections in childhood: a placebo controlled trial. *BMJ*, ii:552–4, 1977.

[93] YIN Z G, ZHANG J B., KAN S. L., and WANG P. A comparison of traditional digital blocks and single subcutaneous palmar injection blocks at the base of the finger

and a meta-analysis of the digital block trials. *Journal of Hand Surgery (British and European Volume)*, 31:547, 2006.

[94] Hung VS, Bodavula VKR, and Dubin NH. Digital anaesthesia: comparison of the efficacy and pain associated with three digital nerve block techniques. *Journal of Hand Surgery*, 30B::581–584., 2005.

[95] Sachdev HPS, Tarun Gera, and Penelope Nestel. Effect of iron supplementation on mental and motor development in children: systematic review of randomised controlled trials. *Public Health Nutrition:*, 8(2):117–132, 2004.

[96] Singh H, Poluha W, Cheung M, Choptain N, Baron KI, and Taback SP. Propofol for sedation during colonoscopy. *CochraneDatabase of Systematic Reviews*, Issue 4(Art.No.: CD006268. DOI: 10.1002/14651858.CD006268.pub2), 2008.

[97] Marinho VCC, Higgins JPT, Logan S, and Sheiham A. Fluoride mouthrinses for preventing dental caries in children and adolescents. *Cochrane Database of Systematic Reviews*, Issue 3(Art. No.: CD002284. DOI:10.1002/14651858.CD002284.), 2003.

[98] Green S, Buchbinder R, Barnsley L, Hall S, White M, Smidt, and Assendelft WJJ. Non-steroidal antiinflammatory drugs (nsaids) for treating lateral elbow pain in adults. *Cochrane Database of Systematic Reviews*, Issue 4(Art. No.:CD003686. DOI: 10.1002/14651858.CD003686), 2001.

[99] Tania Winzenberg, Kelly Shaw, Jayne Fryer, and Graeme Jones. Effects of calcium supplementation on bone density in healthy children: meta-analysis of randomised controlled trials. *BMJ*, 333:775, 2006.

[100] Furukawa TA, Barbui C, Cipriani A, Brambilla P, and Watanabe N. Imputing missing standard deviations in meta-analyses can provide accurate results. *Journal of Clinical Epidemiology*, 59:7–10, 2006.

[101] Wu-Williams A H, Dai X D, Blot W, Xu Z Y, Sun X W, Xiao H P, Stone B J, Yu S F, Feng Y P, Ershow A G, Sun J, Fraumeni J F, and Henderson B E. Lung cancer among women in north-east china. *British Journal of Cancer*, 62:982–987, 1990.

[102] Geng G Y, Liang Z H, Zhang A Y, and Wu G L. On the relationship between cigarette smoking and female lung cancer. *Elsevier Science Publishers B. V. (Biomedical Division)*, pages 483–486, 1988.

[103] Hackshaw A K, Law M R, and Wald N J. The accumulated evidence on lung cancer and environmental tobacco smoke. *British Medical Journal*, 315:980–988, 1997.

[104] Cohen J J. A tutorial on sas macro language,. *AsrtZeneca LP*, pages 1–6, 2009.

[105] Rabindranath KS, Adams J, MacLeod AM, and Muirhead N. Intermittent versus continuous renal replacement therapy for acute renal failure in adults. *Cochrane Database of Systematic Reviews*, Issue 3(Art.No.: CD003773. DOI:10.1002/14651858.CD003773.pub3), 2007.

[106] Terplan M and Lui S. Psychosocial interventions for pregnant women in outpatient illicit drug treatment programs compared to other interventions. *Cochrane Database of Systematic Reviews*, Issue 4(Art. No.: CD006037. DOI:10.1002/14651858.CD006037.pub2.), 2007.

[107] Alonso A, Lau J, and Jaber BL. Biocompatible hemodialysis membranes for acute renal failure. *Cochrane Database of Systematic Reviews*, Issue 1(Art. No.: CD005283. DOI: 10.1002/14651858.CD005283.pub2.), 2008.

[108] Martin-Rendon E, Brunskill S, Dore C, Hyde C, Mathur A, Stanworth S, and Watt S. Stem cell treatment for acute myocardial infarction. *Cochrane Database of Systematic Reviews*, Issue 4(Art. No.: CD006536. DOI:10.1002/14651858.CD006536.pub2.), 2008.

[109] McCallum I, King PM, and Bruce J. Healing by primary versus secondary intention after surgical treatment for pilonidal sinus. *Cochrane Database of Systematic Reviews*, Issue 4(Art. No.: CD006213. DOI:10.1002/14651858.CD006213.pub2.), 2007.

[110] Villatoro E, Bassi C, and Larvin M. Antibiotic therapy for prophylaxis against infection of pancreatic necrosis in acute pancreatitis. *Cochrane Database of Systematic Reviews*, Issue 4(Art. No.: CD002941. DOI:10.1002/14651858.CD002941.pub2.), 2006.

[111] Richter B, Bandeira-Echtler E, Bergerhoff K, Clar C, and Ebrahim SH. Rosiglitazone for type 2 diabetes mellitus. *Cochrane Database of Systematic Reviews*, Issue 3(Art. No.: CD006063. DOI: 10.1002/14651858.CD006063.pub2), 2007.

[112] Poole P and Black PN. Mucolytic agents for chronic bronchitis or chronic obstructive pulmonary disease. *Cochrane Database of Systematic Reviews*, Issue 3(Art. No.: CD001287. DOI: 10.1002/14651858.CD001287.pub2), 2006.

[113] Devereaux P J, W Scott Beattie, Peter T-L Choi, Neal H Badner, Gordon H Guyatt, Juan C Villar, Claudio S Cin, Kate Leslie, Michael J Jacka, Victor M Montori, Mohit Bhandari, Alvaro Avezum, Alexandre B Cavalcanti, Julian W Giles, Thomas Schricker, Homer Yang, Carl-Johan Jakobsen, and Salim Yusuf. How strong is the evidence for the use of perioperative blockers in non-cardiac surgery? systematic review and meta-analysis of randomised controlled trials. *BMJ*, (doi:10.1136/bmj.38503.623646.8F), 2005.

[114] Angela C Webster, Rebecca C Woodroffe, Rod S Taylor, Jeremy R Chapman, and Jonathan C Craig. Tacrolimus versus ciclosporin as primary immunosuppression for kidney transplant recipients: meta-analysis and meta-regression of randomised trial data. *BMJ*, 331(doi:10.1136/bmj.38569.471007.AE):810, 2005.

[115] Arroll Bruce and Felicity Goodyear-Smith. Corticosteroid injections for osteoarthritis of the knee: meta-analysis. *BMJ*, 328(doi:10.1136/bmj.38039.573970.7C):869, 2004.

[116] Elad Goldberga, Anat Gafter-Gvilib, Eyal Robenshtoka, Leonard Leibovici, and Mical Paulb. Empirical antifungal therapy for patients with neutropenia and persistent fever: Systematic review and meta-analysis. *EUROPEAN JOURNAL OF CANCER*, 44:2192–2203, 2008.

[117] Tao Fan, Gang Wang, Bing Mao, Zeyu Yu, Xuemei Liu, Lei Wang, and Sai Yang. Prophylactic administration of parenteral steroids for preventing airway complications after extubation in adults: meta-analysis of randomised placebo controlled trials. *BMJ*, 337:a1841, 2008.

[118] Lynn Legg, Avril Drummond, Jo Leonardi-Bee, J R F Gladman, Susan Corr, Pip Logan, Catherine Sackley, Marion Walker, Peter Langhorne Mireille Donkervoort, Judi Edmans, Louise Gilbertson, and Lyn Jongbloed. Occupational therapy for patients with problems in personal activities of daily living after stroke: systematic review of randomised trials. *BMJ*, 335(doi:10.1136/bmj.39343.466863.55):922, 2007.

[119] Mohammed S and S Goodacre. Intravenous and nebulised magnesium sulphate for acute asthma: systematic review and meta-analysis. *Emerg Med J*, 24(doi: 10.1136/emj.2007.052050):823–830, 2007.

[120] Barr R G, J Bourbeau, C A Camargo, and F S F Ram. Tiotropium for stable chronic obstructive pulmonary disease: a meta-analysis. *Thorax*, 61(doi: 10.1136/thx.2006.063271):854–62, 2006.

[121] Chang A B, Lasserson T J, Kiljander T O, F L Connor, J T Gaffney, and L A Garske. Systematic review and meta-analysis of randomised controlled trials of gastro-

oesophageal reflux interventions for chronic cough associated with gastro-oesophageal reflux. *BMJ*, 332:11–17, 2006.

INCLUDED STUDIES COCHRANE COLLABORATION DATABASE

1. Martin-Rendon E, Brunskill S, Dore C, Hyde C, Mathur A, Stanworth S, Watt S. Stem cell treatment for acute myocardial infarction. Cochrane Database of Systematic Reviews 2008, Issue 4. Art. No.: CD006536. DOI: 10.1002/14651858.CD006536.pub2.

2. Hawke F, Burns J, Radford JA, du Toit V. Custom-made foot orthoses for the treatment of foot pain. Cochrane Database of Systematic Reviews 2008, Issue 3. Art. No.: CD006801. DOI: 10.1002/14651858.CD006801.pub2.

3. Yousefi-Nooraie R, Schonstein E, Heidari K, Rashidian A, Pennick V, Akbari-Kamrani M, Irani S, Shakiba B, Mortaz Hejri S, Jonaidi AR, Mortaz-Hedjri S. Low level laser therapy for nonspecific low-back pain. Cochrane Database of Systematic Reviews 2008, Issue 2. Art. No.: CD005107. DOI: 10.1002/14651858.CD005107.pub4.

4. Webster AC, Pankhurst T, Rinaldi F, Chapman JR, Craig JC. Polyclonal and monoclonal antibodies for treating acute rejection episodes in kidney transplant recipients. Cochrane Database of Systematic Reviews 2006, Issue 2. Art. No.:CD004756. DOI: 10.1002/14651858.CD004756.pub3.

5. Nasser M, Fedorowicz Z, Khoshnevisan MH, Shahiri Tabarestani M. Acyclovir for treating primary herpetic gingivostomatitis. Cochrane Database of Systematic Reviews 2008, Issue 4. Art. No.: CD006700. DOI: 10.1002/14651858.CD006700.pub2.

6. Villatoro E, Bassi C, Larvin M. Antibiotic therapy for prophylaxis against infection of pancreatic necrosis in acute pancreatitis. Cochrane Database of Systematic Reviews 2006, Issue 4. Art. No.: CD002941. DOI: 10.1002/14651858.CD002941.pub2.

7. UrquhartD, Hoving JL, Assendelft WJJ, RolandM, van TulderMW. Antidepressants for non-specific lowback pain.CochraneDatabase of SystematicReviews 2008, Issue 1.Art.No.:CD00170
10.1002/14651858.CD001703.pub3.

8. Richter B, Bandeira-Echtler E, Bergerhoff K, Clar C, Ebrahim SH. Rosiglitazone for type 2 diabetes mellitus. Cochrane Database of Systematic Reviews 2007, Issue 3. Art. No.: CD006063. DOI: 10.1002/14651858.CD006063.pub2

9. Alonso A, Lau J, Jaber BL. Biocompatible hemodialysis membranes for acute renal failure. Cochrane Database of Systematic Reviews 2008, Issue 1. Art. No.: CD005283. DOI: 10.1002/14651858.CD005283.pub2.

10. Quartero AO, Meiniche-Schmidt V, Muris J, Rubin G, de Wit N. Bulking agents, antispasmodic and antidepressant medication for the treatment of irritable bowel syndrome. Cochrane Database of Systematic Reviews 2005, Issue 2.Art. No.: CD003460. DOI: 10.1002/14651858.CD003460.pub2.

11. Macfadyen CA, Acuin JM,GambleCL. Topical antibiotics without steroids for chronically discharging ears with underlying eardrum perforations. Cochrane Database of Systematic Reviews 2005, Issue 4. Art. No.: CD004618. DOI: 10.1002/14651858.CD004618.pub2.

12. Granger RH, Walters JAE, Poole P, Lasserson TJ, Mangtani P, Cates CJ, Wood-Baker R. Injectable vaccines for preventing pneumococcal infection in patients with chronic obstructive pulmonary disease. Cochrane Database of Systematic Reviews 2006, Issue 4. Art. No.: CD001390. DOI: 10.1002/14651858.CD001390.pub2.

13. Navaneethan SD, Shrivastava R. HMG CoA reductase inhibitors (statins) for dialysis patients. Cochrane Database of Systematic Reviews 2004, Issue 4. Art. No.: CD004289. DOI: 10.1002/14651858.CD004289.pub2.

14. McCallum I, King PM, Bruce J. Healing by primary versus secondary intention after surgical treatment for pilonidal sinus. Cochrane Database of Systematic Reviews 2007, Issue 4. Art. No.: CD006213. DOI: 10.1002/14651858.CD006213.pub2.

15. TerplanM, Lui S. Psychosocial interventions for pregnant women in outpatient illicit drug treatment programs compared to other interventions. Cochrane Database of Systematic Reviews 2007, Issue 4. Art. No.: CD006037. DOI: 10.1002/14651858.CD006037.pub2.

16. Harrison JE, OBrien KD,Worthington HV. Orthodontic treatment for prominent upper front teeth in children. Cochrane Database of Systematic Reviews 2007, Issue 3. Art. No.: CD003452. DOI: 10.1002/14651858.CD003452.pub2.

17. Saunders DH, Greig CA, Young A, Mead GE. Physical fitness training for stroke patients. Cochrane Database of Systematic Reviews 2004, Issue 1. Art. No.: CD003316. DOI: 10.1002/14651858.CD003316.pub2

18. Cepeda MS, Carr DB, Lau J, AlvarezH.Music for pain relief. Cochrane Database of Systematic Reviews 2006, Issue 2. Art. No.: CD004843. DOI: 10.1002/14651858.CD004843.pub2

19. Singh H, Poluha W, Cheung M, Choptain N, Baron KI, Taback SP. Propofol for sedation during colonoscopy. CochraneDatabase of Systematic Reviews 2008, Issue 4. Art.No.: CD006268. DOI: 10.1002/14651858.CD006268.pub2

20. Jacobson Vann JC, Szilagyi P. Patient reminder and recall systems to improve immunization rates. Cochrane Database of Systematic Reviews 2005, Issue 3. Art. No.: CD003941. DOI: 10.1002/14651858.CD003941.pub2

21. Meremikwu MM, Oyo-Ita A. Paracetamol for treating fever in children. Cochrane Database of Systematic Reviews 2002, Issue 2. Art. No.: CD003676. DOI: 10.1002/14651858.CD0036

22. Ruddy R,House A. Psychosocial interventions for conversion disorder. Cochrane-Database of Systematic Reviews 2005, Issue 4. Art. No.: CD005331. DOI: 10.1002/14651858.CD0053

23. Barlow J, Parsons J. Group-based parent-training programmes for improving emotional and behavioural adjustment in 0-3 year old children. Cochrane Database of Systematic

Reviews 2003, Issue 2. Art. No.: CD003680. DOI: 10.1002/14651858.CD003680.

24. Palmer SC, McGregor DO, Strippoli GFM. Interventions for preventing bone disease in kidney transplant recipients. Cochrane Database of Systematic Reviews 2007, Issue 3. Art. No.: CD005015. DOI: 10.1002/14651858.CD005015.pub3.

25. Magee K, Sevcik WW, Moher D, Rowe BH. Low molecular weight heparins versus unfractionated heparin for acute coronary syndromes. Cochrane Database of Systematic Reviews 2003, Issue 1. Art. No.: CD002132. DOI: 10.1002/14651858.CD002132.

26. Kerkhoffs GMMJ, Struijs PAA, Marti RK, Assendelft WJJ, Blankevoort L, van Dijk CN. Different functional treatment strategies for acute lateral ankle ligament injuries in adults. Cochrane Database of Systematic Reviews 2002, Issue 3. Art. No.: CD002938. DOI: 10.1002/14651858.CD002938

27. Macdonald G, Higgins JPT, Ramchandani P. Cognitive-behavioural interventions for children who have been sexually abused. Cochrane Database of Systematic Reviews 2006, Issue 4. Art. No.: CD001930. DOI: 10.1002/14651858.CD001930.pub2.

28. Dengfeng W, Lina H, Marjoribanks J, Haijun J, Ying S, Zhang J, Liu GJ, Wu T. Chinese herbal medicines in the treatment of ectopic pregnancy. Cochrane Database of Systematic Reviews 2007, Issue 4. Art. No.: CD006224. DOI: 10.1002/14651858.CD006224.pub2.

29. Buchbinder R, Green S, Youd JM. Corticosteroid injections for shoulder pain. Cochrane Database of Systematic Reviews 2003, Issue 1. Art. No.: CD004016. DOI: 10.1002/14651858.CD0040

30. Evans DJ, Cullinan P, Geddes DM, Walters EH, Milan SJ, Jones P. Cyclosporin as an oral corticosteroid sparing agent in stable asthma. Cochrane Database of Systematic Reviews 2000, Issue 4. Art. No.: CD002993. DOI: 10.1002/14651858.CD002993.

31. Cramp F,Daniel J. Exercise for the management of cancer-related fatigue in adults. Cochrane Database of Systematic Reviews 2008, Issue 2. Art. No.: CD006145. DOI: 10.1002/14651858.CD006145.pub2.

32. Cundiff DK, Manyemba J, Pezzullo JC. Anticoagulants versus non-steroidal anti-inflammatories or placebo for treatment of venous thromboembolism. Cochrane Database of Systematic Reviews 2006, Issue 1. Art.No.: CD003746. DOI: 10.1002/14651858.CD003746.pub2.

33. Poole P, Black PN.Mucolytic agents for chronic bronchitis or chronic obstructive pulmonary disease. Cochrane Database of Systematic Reviews 2006, Issue 3. Art. No.: CD001287. DOI: 10.1002/14651858.CD001287.pub2

34. EM,Willis NS, Craig JC. Non-corticosteroid treatment for nephrotic syndrome in children. Cochrane Database of Systematic Reviews 2008, Issue 1. Art. No.: CD002290. DOI: 10.1002/14651858.CD002290.pub3

35. Green S, Buchbinder R, Barnsley L, Hall S, WhiteM, Smidt N, Assendelft WJJ. Non-steroidal antiinflammatory drugs (NSAIDs) for treating lateral elbow pain in adults. Cochrane Database of Systematic Reviews 2001, Issue 4. Art. No.: CD003686. DOI: 10.1002/14651858.CD003686.

36. OConnor D, Marshall SC, Massy-Westropp N. Non-surgical treatment (other than steroid injection) for carpal tunnel syndrome. Cochrane Database of Systematic Reviews 2003, Issue 1. Art. No.: CD003219. DOI: 10.1002/14651858.CD003219.

37. Leslie K, Clavisi O, Hargrove J. Target-controlled infusion versus manually-controlled infusion of propofol for general anaesthesia or sedation in adults. Cochrane Database of Systematic Reviews 2008, Issue 3. Art. No.: CD006059.DOI: 10.1002/14651858.CD006059.pub2.

38. Price J, Leng GC. Steroid sex hormones for lower limb atherosclerosis. Cochrane Database of Systematic Reviews 2001, Issue 3. Art. No.: CD000188. DOI: 10.1002/14651858.CD0001

39. Mailis-Gagnon A, Furlan,MD PhD AD, Sandoval JA, Taylor RS. Spinal cord stimulation for chronic pain. Cochrane Database of Systematic Reviews 2004, Issue 3. Art. No.: CD003783. DOI: 10.1002/14651858.CD003783.pub2.

40. Rabindranath KS, Adams J, MacLeod AM, Muirhead N,Intermittent versus continuous renal replacement therapy for acute renal failure in adults. Cochrane Database of Systematic Reviews 2007, Issue 3.Art.No.: CD003773. DOI:10.1002/14651858.CD003773.pub3.

41. Marinho VCC, Higgins JPT, Logan S, Sheiham A. Fluoride mouthrinses for preventing dental caries in children and adolescents. Cochrane Database of Systematic Reviews,2003, Issue 3. Art. No.: CD002284. DOI:10.1002/14651858.CD002284

42. Strippoli GFM, Tong A, Johnson DW, Schena FP, Craig JC:Antimicrobial agents for preventing peritonitis in peritoneal dialysis patients,Cochrane Database of Systematic Reviews,2004, Issue 4;Art. No.: CD004679. DOI:10.1002/14651858.CD004679.

WEB OF KNOWLEDGE

43. Muhammed Ashraf Memon and Shahjahan Khan and Rossita Mohamad Yunus and Richard Barr and Breda Memon.Meta-analysis of laparoscopic and open distal gastrectomy for gastric carcinoma.Surg Endosc (2008) 22:17811789 DOI 10.1007/s00464-008-9925-9

44. Z. G. YIN, J. B. ZHANG, S. L. KAN and P. WANG. A Comparison of Traditional Digital Blocks and Single Subcutaneous Palmar Injection Blocks at the Base of the Finger and a Meta-Analysis of the Digital Block Trials. Journal of Hand Surgery (British and European Volume) 2006; 31; 547, DOI: 10.1016/J.JHSB.2006.06.001

45. Tania Lourenco, Robert Pickard, Luke Vale, Adrian Grant, Cynthia Fraser,Graeme MacLennan, James NDow and and the Benign Prostatic Enlargement team. Alterna-

tive approaches to endoscopic ablation systematic review of randomised controlled for benign enlargement of the prostate:BMJ 2008;337;a449,doi:10.1136/bmj.39575.517674.BE

46. Glioma Meta-analysis Trialists (GMT) Group*. Chemotherapy in adult high-grade glioma: a systematic review and meta-analysis of individual patient data from 12 randomised trials.Lancet 2002; 359: 101118

47. Tao Fan, Gang Wang, Bing Mao, Zeyu Xiong, Yu Zhang, Xuemei Liu, Lei,Wang and Sai Yang. Prophylactic administration of parenteral steroids for preventing airway complications after extubation in adults: meta-analysis of,Randomised placebo controlled trials BMJ 2008;337:a1841

48. Henry S. Tilney and Parvinder S. Sains and Richard E. Lovegrove and George E. Reese and Alexander G. Heriot and Paris P. Tekkis:Comparison of Outcomes Following Ileostomy versus Colostomy for Defunctioning Colorectal Anastomoses.World J Surg (2007) 31: 11421151, DOI: 10.1007/s00268-006-0218-y

49. Robin Christensen, Pernelle Kruse Kristensen, Else Marie Bartels, Henning Bliddal, Arne Astrup: Efficacy and safety of the weight-loss drug rimonabant:a meta-analysis of randomised trials,Lancet 2007; 370: 170613

50. John L. Moran and Andrew D. Bersten and Patricia J. Solomon: Meta-analysis of controlled trials of ventilator therapy in acute lung injury and acute respiratory distress syndrome:an alternative perspective,Intensive Care Med (2005) 31:227235,DOI 10.1007/s00134-004-2506-z

51. B Arroll and T Kenealy: Are antibiotics effective for acute purulent rhinitis? Systematic review and meta-analysis of placebo controlled randomised trials,BMJ 2006;333;279;doi:10.1136/b

52. S Mohammed and S Goodacre: Intravenous and nebulised magnesium sulphate for acute asthma: systematic review and meta-analysis,Emerg Med J 2007;24:823830. doi: 10.1136/emj.2007.052050

53. T. Ibrahim and I. M. Tleyjeh and O. Gabbar: Surgical versus non-surgical treatment of chronic low back pain: a meta-analysis of randomised trials, International Orthopaedics (SICOT) (2008) 32:107113,DOI 10.1007/s00264-006-0269-6

54. HPS Sachdev and Tarun Gera and Penelope Nestel:Effect of iron supplementation on mental and motor development in children: systematic review of randomised controlled trials, Public Health Nutrition: 8(2), 117132 DOI: 10.1079/PHN2004677

55. Lynn Legg and Avril Drummond and Jo Leonardi-Bee and J R F Gladman and Susan Corr and Pip Logan and Catherine Sackley and Marion Walker and Peter Langhorne Mireille Donkervoort and Judi Edmans and Louise Gilbertson and Lyn Jongbloed: Occupational therapy for patients with problems in personal activities of daily living after stroke: systematic review of randomised trials, BMJ 2007;335;922; originally published online 27 Sep 2007,doi:10.1136/bmj.39343.466863.55.

56. Xiaobin Wang, Xianhui Qin, Hakan Demirtas, Jianping Li, Guangyun Mao, Yong Huo, Ningling Sun, Lisheng Liu, Xiping Xu: Effi cacy of folic acid supplementation in stroke prevention:a meta-analysis,Lancet 2007; 369: 187682.

57. J R Hughes and S Shiffman and P Callas and J Zhang:A meta-analysis of the efficacy of over-the-counter nicotine replacement Tobacco Control 2003;12:2127

58. Angela C Webster, Rebecca C Woodroffe, Rod S Taylor, Jeremy R Chapman, Jonathan C Craig:Tacrolimus versus ciclosporin as primary immunosuppression for kidney transplant recipients: meta-analysis and meta-regression of randomised trial data, BMJ 2005;331;810; originally published online 12 Sep 2005,doi:10.1136/bmj.38569.471007.AE

59. A L Takeda and J Colquitt and A J Clegg and J Jones:Pegaptanib and ranibizumab for neovascular age-related macular degeneration: a systematic review, Br J Ophthalmol 2007;91:11771182. doi: 10.1136/bjo.2007.118562

60. R G Barr and J Bourbeau and C A Camargo and F S F Ram:Tiotropium for stable chronic obstructive pulmonary disease: a meta-analysis, Thorax 2006;61:854862. doi: 10.1136/thx.2006.063271

61. W Y Zhang, A Li Wan Po, H S Dua, A Azuara-Blanco:Meta-analysis of randomised controlled trials comparing latanoprost with timolol in the treatment of patients with open angle glaucoma or ocular hypertension,Br J Ophthalmol 2001;85:983990

62. Bruce Arroll, Felicity Goodyear-Smith: Corticosteroid injections for osteoarthritis of the knee: meta-analysis,BMJ 2004;328;869; originally published online 23 Mar 2004,doi:10.1136/bmj.38039.573970.7C

63. M. D. Shelley and T. J. Wilt and J. Barber and M. D. Mason:A Meta-analysis of Randomised Trials Suggests a Survival Benefit for Combined Radiotherapy and Radical Cystectomy Compared with Radical Radiotherapy for Invasive Bladder Cancer: Are These Data Relevant to Modern Practice?,Clinical Oncology (2004) 16: 166171 doi:10.1016/j.clon.2003.10.001

64. Darren M. Ashcroft PhD1 and David Millson PhD:Naratriptan for the treatment of acute migraine: meta-analysis of randomised controlled trials,pharmacoepidemiology and drug safety 2004; 13: 7382,DOI: 10.1002/pds.890

65. Christine Rose and Louise Wallace and Rumona Dickson and Jon Ayres and Richard Lehman and Yvonne Searle and p.Sherwood Burge:The most effective psychologically-based treatments to reduce anxiety and panic in patients with chronic obstructive

pulmonary disease(COPD) :a systematic review,Patient Education and Counselling 47(2002)311-318

66. P S Sanmuganathan, P Ghahramani, P R Jackson, E J Wallis, L E Ramsay.Aspirin for primary prevention of coronary heart disease: safety and absolute benefit related to coronary risk derived from meta-analysis of randomised trials Heart 2001;85:265271

67. Pallavi Latthe, Luciano Mignini, Richard Gray, Robert Hills, Khalid Khan.Factors predisposing women to chronic pelvic pain: systematic review.BMJ 2006;332;749-755; originally published online 16 Feb 2006;doi:10.1136/bmj.38748.697465.55

68. Abigail Fraser and Mical Paul and Elad Goldberg and Camilo J. Acosta and Leonard Leibovici.Typhoid fever vaccines: Systematic review and meta-analysis of randomised controlled trials,Vaccine 25 (2007) 78487857

69. John W Eikelboom, Daniel J Quinlan, James D Douketis.Extended-duration prophylaxis against venous thromboembolism after total hip or knee replacement: a meta-analysis of the randomised trials,Lancet 2001; 358: 915

70. A Messori, S Trippoli, M Vaiani, M Gorini, A Corrado.Bleeding and pneumonia in intensive care patients given ranitidine and sucralfate for prevention of stress ulcer:metaanalysis of randomised controlled trials. BMJ 2000;321:17

71. S. R. Walsh,1 T. Tang,1 S. Bass,2 M. E. Gaunt1.Doppler-guided intra-operative fluid management during major abdominal surgery: systematic review and meta-analysis.Journal compilation 2007 Blackwell Publishing Ltd Int J Clin Pract, March 2008, 62, 3, 466470

72. P von Dadelszen, M P Ornstein, S B Bull, A G Logan, G Koren, L A Magee.Fall in mean arterial pressure and fetal growth restriction in pregnancy hypertension: a meta-analysis,Lancet 2000; 355: 8792

73. Kathleen Hempenstall1, Turo J. Nurmikko2, Robert W. Johnson3, Roger P. AHern4, Andrew S.C. Rice5. Analgesic Therapy in Postherpetic Neuralgia:A Quantitative Systematic Review,PLoS Med 2(7): e164.

MEDLINE

74. L. H. J. EBERHART1, A. M. MORIN1, U. BOTHNER2 and M. GEORGIEFF1.Droperidol and 5-HT3-receptor antagonists, alone or in combination, for prophylaxis of postoperative nausea and vomiting:A meta-analysis of randomised controlled trials,Acta Anaesthesiol Scand 2000; 44: 12521257.

75. Javier Ballesterosa and Luis F. Calladoc.Effectiveness of pindolol plus serotonin uptake inhibitors in depression: a meta-analysis of early and late outcomes from randomised controlled trials,Journal of Affective Disorders 79 (2004) 137147.

76. Alia El-Kadiki, Alexander J Sutton. Role of multivitamins and mineral supplements in preventing infections in elderly people: systematic review and meta-analysis of randomised controlled trials.BMJ 2005;330;871; originally published online 31 Mar 2005

77. P J Devereaux,W Scott Beattie, Peter T-L Choi, Neal H Badner, Gordon H Guyatt, Juan C Villar, Claudio S Cin, Kate Leslie, Michael J Jacka, Victor M Montori, Mohit Bhandari, Alvaro Avezum, Alexandre B Cavalcanti, Julian W Giles, Thomas Schricker, Homer Yang, Carl-Johan Jakobsen, Salim Yusuf. How strong is the evidence for the use of perioperative blockers in non-cardiac surgery? Systematic review and meta-analysis of randomised controlled trials.BMJ, doi:10.1136/bmj.38503.623646.8F (published 4 July 2005)

78. A B Chang, T J Lasserson, T O Kiljander, F L Connor, J T Gaffney, L A Garske. Systematic review and meta-analysis of randomised controlled trials of gastro-oesophageal

reflux interventions for chronic cough associated with gastro-oesophageal reflux BMJ 2006;332;11-17

79. Tania Winzenberg, Kelly Shaw, Jayne Fryer, Graeme Jones.Effects of calcium supplementation on bone density in healthy children: meta-analysis of randomised controlled trials,BMJ 2006;333;775; originally published online 15 Sep 2006,doi:10.1136/bmj.38950.561400.55

80. G Wright, R L Manser, G Byrnes, D Hart, D A Campbell.Surgery for non-small cell lung cancer: systematic review and meta-analysis of randomised controlled trials.Thorax 2006;61:597603. doi: 10.1136/thx.2005.051995

81. Li-Chia Chen PhDz and Darren M Ashcroft.Risk of myocardial infarction associated with selective COX-2 inhibitors: Meta-analysis of randomised controlled trials,pharmacoepidemiology and drug safety 2007; 16: 762772, DOI: 10.1002/pds.1409

82. Elad Goldberga,b,*, Anat Gafter-Gvilib,c, Eyal Robenshtoka and Leonard Leibovicia,b, Mical Paulb.Empirical antifungal therapy for patients with neutropenia and persistent fever: Systematic review and meta-analysis.EUROPEAN JOURNAL OF CANCER 44(2008) 21922203

83. Matthew MASOLI,1 Mark WEATHERALL,2 Shaun HOLT3 AND Richard BEASLEY1.Budesonide once versus twice-daily administration: meta-analysis,Respirology (2004) 9, 528534 ,kkk,

APPENDIX COCHRANE DATABASE

There were 515 meta-analyses and each was given a unique identification number from 1 to 515. The studies to be sampled were then identified using the following R code. Given $N = 515$ records of Meta-analysis of RCT's which met the inclusion criteria and $n = 50$ is

the sample size , then

$$> N < -515$$

$$> n < -50$$

$$> Identify < -c(1 : N)$$

$$> sort(sample(Identify, size = n, replace = F))$$

The list of simple random numbers for Cochrane is 27, 31, 32, 36, 40, 49, 62, 67, 76, 79, 81, 97, 103, 117, 119, 204, etc,

WEB OF KNOWLEDGE

,

$N = 507$ records of meta-analysis of RCT's that met inclusion criteria, $n = 40$ sample papers

$$> N < -507$$

$$> n < -40$$

$$> Identify < -c(1 : N)$$

$$> sort(sample(Identify, size = n, replace = F))$$

The list of simple random numbers for WOK is 12, 19, 29, 63, 78, 110, 115, 125,142, 151, 175, 181 etc

MEDLINE

,

has $N = 130$ of meta-analysis of RCT's meeting the inclusion criteria, $n = 39$ sample

papers

$$> N < -130$$

$$> n < -10$$

$$> Identify < -c(1 : N)$$

$$> sort(sample(Identify, size = n, replace = F))$$

The list of simple random numbers for Medline is 7, 8, 13, 23, 58, 79, 69, 90, 100, 123 etc