

Use of Mixed Treatment Comparisons in Systematic Reviews



Agency for Healthcare Research and Quality
Advancing Excellence in Health Care • www.ahrq.gov

Use of Mixed Treatment Comparisons in Systematic Reviews

Prepared for:

Agency for Healthcare Research and Quality
U.S. Department of Health and Human Services
540 Gaither Road
Rockville, MD 20850
www.ahrq.gov

Contract No. 290-2007-10067-I

Prepared by:

University of Connecticut/Hartford Hospital Evidence-based Practice Center
Hartford, CT

Investigators:

Craig I. Coleman, Pharm.D.
Olivia J. Phung, Pharm.D.
Joseph C. Cappelleri, Ph.D., M.P.H., M.S.
William L. Baker, Pharm.D., BCPS
Jeffrey Kluger, M.D., FACC
C. Michael White, Pharm.D., FCP, FCCP
Diana M. Sobieraj, Pharm.D.

This report is based on research conducted by the University of Connecticut/Hartford Hospital Evidence-based Practice Center (EPC) under contract to the Agency for Healthcare Research and Quality (AHRQ), Rockville, MD (Contract No. 290-2007-10067-I). The findings and conclusions in this document are those of the authors, who are responsible for its contents; the findings and conclusions do not necessarily represent the views of AHRQ. Therefore, no statement in this report should be construed as an official position of AHRQ or of the U.S. Department of Health and Human Services.

The information in this report is intended to help health care decisionmakers—patients and clinicians, health system leaders, and policymakers, among others—make well-informed decisions and thereby improve the quality of health care services. This report is not intended to be a substitute for the application of clinical judgment. Anyone who makes decisions concerning the provision of clinical care should consider this report in the same way as any medical reference and in conjunction with all other pertinent information, i.e., in the context of available resources and circumstances presented by individual patients.

This report may be used, in whole or in part, as the basis for development of clinical practice guidelines and other quality enhancement tools, or as a basis for reimbursement and coverage policies. AHRQ or U.S. Department of Health and Human Services endorsement of such derivative products may not be stated or implied.

This document is in the public domain and may be used and reprinted without permission except those copyrighted materials that are clearly noted in the document. Further reproduction of those copyrighted materials is prohibited without the specific permission of copyright holders.

Persons using assistive technology may not be able to fully access information in this report. For assistance contact EffectiveHealthCare@ahrq.gov.

None of the investigators have any affiliations or financial involvement that conflicts with the material presented in this report.

Suggested citation: Coleman CI, Phung OJ, Cappelleri JC, Baker WL, Kluger J, White CM, Sobieraj DM. Use of Mixed Treatment Comparisons in Systematic Reviews. Methods Research Report. (Prepared by the University of Connecticut/Hartford Hospital Evidence-based Practice Center under Contract No. 290-2007-10067-I.) AHRQ Publication No. 12-EHC119-EF. Rockville, MD: Agency for Healthcare Research and Quality. August 2012. www.effectivehealthcare.ahrq.gov/reports/final.cfm.

Preface

The Agency for Healthcare Research and Quality (AHRQ), through its Evidence-based Practice Centers (EPCs), sponsors the development of evidence reports and technology assessments to assist public- and private-sector organizations in their efforts to improve the quality of health care in the United States. The reports and assessments provide organizations with comprehensive, science-based information on common, costly medical conditions and new health care technologies and strategies. The EPCs systematically review the relevant scientific literature on topics assigned to them by AHRQ and conduct additional analyses when appropriate prior to developing their reports and assessments.

To improve the scientific rigor of these evidence reports, AHRQ supports empiric research by the EPCs to help understand or improve complex methodologic issues in systematic reviews. These methods research projects are intended to contribute to the research base in and be used to improve the science of systematic reviews. They are not intended to be guidance to the EPC program, although may be considered by EPCs along with other scientific research when determining EPC program methods guidance.

AHRQ expects that the EPC evidence reports and technology assessments will inform individual health plans, providers, and purchasers as well as the health care system as a whole by providing important information to help improve health care quality. The reports undergo peer review prior to their release as a final report.

We welcome comments on this Methods Research Project. They may be sent by mail to the Task Order Officer named below at: Agency for Healthcare Research and Quality, 540 Gaither Road, Rockville, MD 20850, or by email to EffectiveHealthCare@ahrq.gov.

Carolyn M. Clancy, M.D.
Director
Agency for Healthcare Research and Quality

Jean Slutsky, P.A., M.S.P.H.
Director, Center for Outcomes and Evidence
Agency for Healthcare Research and Quality

Stephanie Chang, M.D., M.P.H.
Director
Evidence-based Practice Program
Center for Outcomes and Evidence
Agency for Healthcare Research and Quality

Parivash Nourjah, Ph.D.
Task Order Officer
Center for Outcomes and Evidence
Agency for Healthcare Research and Quality

Peer Reviewers

Jeroen Jansen, Ph.D.
Mapi Consultancy
Boston, MA

Edward Mills, M.Sc., Ph.D., L.L.M.
Interdisciplinary School of Health Sciences
University of Ottawa
Ottawa, Canada

M. Hassan Murad, M.D.
Preventive, Occupational, and Aerospace
Medicine, Mayo Clinic
Rochester, MN

Christopher Schmid, Ph.D.
Biostatistics Research Center
Institute for Clinical Research and Health
Policy Studies, Tufts Medical Center
Professor of Medicine, Tufts University
Boston, MA

Nicky Welton, B.Sc., M.Sc., Ph.D.
School of Social and Community Medicine,
University of Bristol
Bristol, United Kingdom

Use of Mixed Treatment Comparisons in Systematic Reviews

Structured Abstract

Objectives: To summarize publically available guidance for, and current use of, meta-analytic methods for mixed treatment comparison (MTC) evidence synthesis; to identify analyses using these methods and summarize their characteristics; to gain insight regarding the rationale for selection, implementation, and reporting of such methods from investigators.

Methods: In part one, we identified currently available guidance documents addressing the use of MTC in evidence synthesis by searching governmental agencies' and participating members' of the International Network of Agencies for Health Technology Assessment Web sites. Commonalities and disagreements among guidance documents were summarized qualitatively. Next, in part two, a systematic literature search for MTCs was undertaken. Characteristics of included analysis were summarized qualitatively. Last, in part three, we invited a random selection of nine investigators from the systematic literature search to participate in a focus group. Using a Web-based series of questions, we queried respondents regarding their opinion of network meta-analysis and how elements of MTC methodology were chosen in their identified analysis. Responses were summarized qualitatively.

Results: Guidance documents were typically written in a fashion to be applicable to network meta-analysis in general and not to a specific methodology. Guidance documents stressed Bayesian and Frequentist MTC approaches have strengths and limitations, while only one guidance document attempted to comprehensively address how to conduct a network meta-analysis and how to interpret and report results.

Our systematic review identified 42 MTCs of which the majority used Bayesian methods (80.9 percent). Bayesian analyses either used noninformative priors or did not report detail about priors used. Data regarding the evaluation of convergence, heterogeneity, and inconsistency were not consistently reported, and from those providing detail, it appears a broad range of methods were used.

Due to the infrequent use of Frequentist methods for MTC and poor response rate to our focus group invitation, all respondents had conducted a MTC using Bayesian methods. Consequently, we were unable to compare/contrast the viewpoints of investigators who used these two different methods.

Conclusion: Additional guidance on how and when to conduct a MTC, as well as how to interpret and report results is needed. Published meta-analyses using these methods varied in how they conducted and reported results.

Contents

Introduction	1
Background.....	1
Objectives	1
Project-Specific Terminology.....	2
Methods	3
Part One: Review of Existing Guidance Documents	3
Searching the Literature	3
Data Synthesis.....	3
Part Two: Systematic Review of Existing Bayesian or Frequentist MTCs	3
Searching the Literature	3
Inclusion and Exclusion Criteria.....	4
Data Extraction	4
Data Synthesis.....	5
Part Three: MTC Focus Group	5
Composition of the Focus Group.....	5
Data Synthesis.....	5
Results	6
Part One: Review of Existing Guidance Documents	6
Key Points.....	6
Detailed Analysis.....	7
Part Two: Systematic Review of Existing MTCs	18
Results of the Literature Search	18
Key Points.....	20
Detailed Analysis.....	21
Part Three: MTC Focus Group	29
Key Points.....	29
Detailed Analysis.....	30
Discussion	36
References	39
Abbreviations	45

Tables

Table 1. Checklist of Good Research Practices for Conducting and Reporting Network Meta-Analyses.....	9
Table 2. International Society for Pharmacoeconomics and Outcomes Guidance for the Reporting of a Network Meta-Analysis	15
Table 3. Journal-Level Characteristics	21
Table 4. General Characteristics of Bayesian Mixed Treatment Comparisons	22
Table 5. Methods Characteristics in Bayesian Mixed Treatment Comparisons.....	24

Table 6. Outcomes and Results Reporting in Bayesian Mixed Treatment Comparisons	26
Table 7. General Characteristics of Frequentist Mixed Treatment Comparisons.....	27
Table 8. Methods Characteristics in Frequentist Mixed Treatment Comparisons.....	28
Table 9. Outcomes and Results Reporting in Frequentist Mixed Treatment Comparisons.....	29
Table 10. Training of Respondents.....	30
Table 11. Role of Respondents in Their Meta-Analysis.....	31
Table 12. General Mixed Treatment Comparison Questions	31
Table 13. Question Specific to the Investigator’s Published Mixed Treatment Comparisons With Bayesian Methods.....	33
Table 14. Question Specific to the Investigator’s Published Mixed Treatment Comparison and How Much the Specific Parameters Influenced Their Decision To Use Bayesian Methods	34
Table 15. Information About How the Code Was Derived	35
Table 16. Information About How and Why the Prior Distributions Were Chosen.....	35

Figures

Figure 1. Network Pattern Types	11
Figure 2. Inclusion of Identified Citations for Part Two of This Project.....	19

Appendixes

Appendix A. Verbatim Quotes From Guidance Documents
Appendix B. Literature Search for Part Two
Appendix C. Data Extraction Tool for Part Two
Appendix D. Focus Group Questions
Appendix E. Excluded Studies
Appendix F. Evidence Tables
Appendix G. Glossary

Introduction

Background

Clinicians and decisionmakers often have to select from multiple available interventions when determining the optimal treatment for disease. Ideally, high-quality randomized controlled trials (RCTs) that estimate the effectiveness of all possible interventions directly against one another would be available to guide decisionmaking.^{1,2} However, interventions are commonly compared with placebo or non-active control in RCTs rather than another active intervention and when direct comparative trials exist they are between two of a larger group of possible treatments. As such decisionmakers are faced with a lack of adequate direct comparative data to make their judgments.

In the absence of direct comparative data, indirect comparisons may provide valuable information. For example, if two different interventions have been evaluated against a common comparator, the relative effects of the two interventions compared with each other can be estimated indirectly.^{1,2} Even in the presence of direct comparative data, indirect comparisons may add value to the interpretation of comparative effectiveness, as usually more than two treatments for a given disease are considered in clinical practice (even if all treatments are not directly compared).

According to the Conducting Quantitative Synthesis When Comparing Medical Interventions chapter of the Evidence-based Practice Center (EPC) methods guide, investigators may choose to implement an indirect or mixed treatment comparison (MTC) in order to make statistical comparisons between interventions.³ Several methodologies exist to indirectly compare interventions, as do modes to implement such methodologies.^{1,4-7} These include anchored indirect comparisons as described by Bucher et al.,⁴ Frequentist MTC⁵ and Bayesian MTC.^{6,7} In the simplest form, interventions that are evaluated against a common comparator in separate trials can be compared to each other indirectly using an anchored indirect treatment comparison approach.⁴ This and related approaches have been previously addressed in a health technology assessment report by Glenny et al. and consequently are not the focus of this report.⁸ As a generalization of indirect comparisons, when more than two treatments are being compared indirectly, and at least one pair of treatments is being compared both directly and indirectly (a closed loop is present), both direct and indirect types of data can be used to estimate effects in a network meta-analysis using a Bayesian or Frequentist framework.^{1,2,4-7} Although these latter methodologies for synthesizing networks of studies with at least one closed loop are frequently employed, best practices for their use are unclear.^{1,2,9}

Objectives

This report is divided into three parts, each with its own objective.

- Part one: Summarize publicly available guidance discussing when and how to conduct a MTC as well as how to interpret and report the results of such analysis. We will highlight guidance on methods to synthesize MTCs. However, we will also summarize guidance applicable to network meta-analysis in general when such guidance also applies to MTC.
- Part two: Identify either Bayesian or Frequentist MTCs that were published since 2006, and summarize their characteristics.
- Part three: Gather insight from investigators who have conducted either Bayesian or Frequentist MTCs, as identified in part two of this project. More specifically,

investigators will be queried about how elements of such methodology should be chosen and reported.

Project-Specific Terminology

Throughout this report we will use the following specific definitions:

- **Network meta-analysis:** Meant generically to define the simultaneous synthesis of evidence of all pairwise comparisons across more than two interventions^{8,9}
- **Closed loop:** Each comparison has both direct evidence and indirect evidence. For example, consider AB trials, AC trials, and BC trials. The BC comparison has direct evidence from the BC trials and indirect evidence from the AB and AC trials (and similarly for the AB comparison and the AC comparison)
- **Mixed treatment comparison (MTC):** A statistical approach used to analyze a network of evidence with more than two interventions which are being compared indirectly, and at least one pair of interventions compared both directly and indirectly⁹
- **Bayesian framework:** An approach that can be used to conduct MTCs (as well as simpler indirect treatment comparisons) involving a formal combination of a prior probability distribution, which reflects a prior belief of the possible values of the model parameter of interest, with a likelihood distribution of these parameters based on the observed data, to obtain a corresponding posterior probability distribution.⁹
- **Lumley's network meta-analysis approach:** A Frequentist approach to conduct a MTC originally described by Lumley et al. whereby both direct and indirect evidence are combined when there is at least one closed loop of evidence connecting two interventions of interest using a mixed model.^{7,10,12}

Methods

Part One: Review of Existing Guidance Documents

Searching the Literature

We searched for publicly available guidance reports and manuals prepared by regulatory bodies or organizations engaged in evidence synthesis for guidance related to network meta-analyses or MTCs. More specifically, we searched the following Web sites: (1) Agency for Healthcare Research and Quality (AHRQ, www.ahrq.gov); (2) Centre for Reviews and Dissemination (CRD, www.crd.york.ac.uk/crdweb/); (3) Cochrane Collaboration (www.cochrane.org/); (4) National Institute for Health and Clinical Excellence (NICE, www.nice.org.uk/); (5) International Society of Pharmacoeconomics and Outcomes Researchers (ISPOR, www.ispor.org/); (6) Drug Effectiveness Review Program (DERP) of the Oregon Health & Science University (OHSU) Center for Evidence-based Policy (www.ohsu.edu/xd/research/centers-institutes/evidence-based-policy-center/derp/?WT_rank=1/); (7) the Institute of Medicine (IOM, www.iom.edu/) and (8) all other current members of the International Network of Agencies for Health Technology Assessment (INAHTA) (as listed on the INAHTA Web site www.inahta.org/ on December 26, 2011). Finally, we supplemented the above with a Google search (www.google.com).

Data Synthesis

Each identified relevant document was read by a single researcher in detail, and key statements were summarized into the following three categories:

- Guidance on “when to conduct” a network meta-analysis/MTC, including definitions of network meta-analysis and MTC, justification for conducting such analyses and assumptions that should be met.
- Guidance on “how to conduct” a network meta-analysis/MTC, including planning and design, analysis framework, statistical modeling, detection and handling of potential bias/inconsistency, assessment of model fit and sensitivity analysis.
- Guidance on “how to report and interpret” a network meta-analysis/MTC, including requirements or suggestions for reporting and displaying results, types of permissible conclusions, limitations of analysis.

Provided synthesis was not intended to be an exhaustive representation of the content of the source documents, but rather a summary. A selection of verbatim quotes from the source documents can be found in Appendix A.

Part Two: Systematic Review of Existing Bayesian or Frequentist MTCs

Searching the Literature

A systematic literature search was conducted in Medline (2006 to July 31, 2011), the Centre for Reviews and Dissemination Databases (July 31, 2011) (including the Database of Abstracts

and Reviews of Effects, Health Technology Assessment and the National Institute for Health Research Economic Evaluation Database), The Cochrane Library (July 31, 2011), and the American College of Physicians Journal Club (2006 to July 31, 2011). The search strategy in Appendix B was used.

Inclusion and Exclusion Criteria

Two independent investigators assessed reviews for inclusion in a parallel manner based on a priori defined criteria. Systematic reviews that met the following criteria were included: (1) compared the clinical effectiveness or safety of three or more interventions (any treatment, dose, treatment regimen or clinical procedure) based on RCTs; (2) utilized either Bayesian or Frequentist methods to conduct MTC; (3) published in full text; (4) published in the English language; and (5) published between January 1, 2006 and July 31, 2011. Of note, methodological publications that presented MTCs for illustrative purposes and cost-effectiveness analyses were not considered, nor were individual patient data meta-analyses. We included all interventions regardless if pharmacologic, behavioral, or procedural.

Data Extraction

Two reviewers used a standardized tool (Appendix C) to independently extract data; disagreements were resolved through discussion. For each included MTC, all published material including the manuscript, supplements, appendices, or external Web sites which the reader was referred to for additional data were used during data extraction. Therefore, the extraction of data in this project is predicated on the reporting of the information by the authors within these sources. When extracting data, we recorded what the authors reported without ourselves judging whether the methods were appropriate or not. If there was insufficient data from all available sources, we indicated “not reported” for that criterion on data extraction.

First, general characteristics were collected on the journals in which included MTCs were published. Characteristics included journal name, impact factor, allowance of supplements or appendices, and limitations on word, table, and figure counts. Secondly, general characteristics of each MTC were extracted including: (1) the number of authors and if any authors were considered to be methodologists, (2) the number and type of intervention comparisons made; (3) country and year in which the review was conducted; (4) funding source and affiliations; (5) number of printed pages and use of supplement or appendix; (6) the number of trials and patients in the analyses; (7) clinical area (e.g., cardiology, endocrinology, etc.); and (8) the network pattern. For the purposes of this project, we defined a methodologist as an individual with affiliation to a department of statistics, biostatistics, epidemiology, clinical epidemiology, or public health services, as determined by author information and affiliations listed in the publication.¹¹ The country in which a review was conducted was determined by the corresponding author’s affiliation. The network pattern was determined by figures presented within the review. If a figure was not available investigators determined the pattern based on text descriptions of included trials.

We also extracted information regarding the methods used to conduct the MTC including (1) methods/models applied (e.g., Bayesian or Frequentist); (2) whether a fixed-effect or random-effects model was applied; (3) description of model parameters (e.g., choices of prior distributions in Bayesian analysis and assumptions in Frequentist analysis); (4) method for assessment of model fit; (5) methods for handling of potential bias, inconsistency and heterogeneity (e.g., qualitative or quantitative); (6) use of covariate adjustment in models; (7)

whether the model accommodated multi-arm trials; (8) software utilized (WinBUGS, OpenBUGS, wrappers, R, commercial software such as SAS/STATA/SPSS); and (9) availability of code. Finally we extracted data concerning the reporting of results. This included (1) type of endpoint (e.g., continuous versus binary); (2) effect size (e.g., odds ratio, relative risk, risk difference, weighted mean difference) assessed; (3) measure of variance (e.g., confidence or credible intervals); (4) use of other methods to report results (e.g., probability of treatment being best, claims of equivalence or non-inferiority); and (5) format/presentation of results (e.g., text, tables, figures, such as figure of network of studies, raw data tables).

Data Synthesis

The general characteristics of journals and MTCs were summarized qualitatively. Categorical data is presented using frequencies and continuous data as means with standard deviations (SDs).

Part Three: MTC Focus Group

Composition of the Focus Group

Up to nine individuals were allowed to participate in this focus group. We randomly identified MTCs identified in part two of this project to invite, via email, either the first or corresponding author to participate in this group. If no response was obtained we sent a reminder email. If we still did not receive a response, we attempted to contact another author on the publication. After contacting two authors unsuccessfully, we selected another publication at random. Upon investigator-expressed interest, a link was sent to the investigator via email which redirected them to the Web-based tool SurveyMonkey©. The investigators were asked to complete questions in regard to the unique MTC which we identified in Part Two of this project (Appendix D).

We intended on participation in this group to be numerically similar between investigators who used Bayesian and Frequentist methods. However, the number of Frequentist MTCs identified in part two of this report was much fewer and author response was suboptimal. In an effort to identify additional models using Frequentist MTCs, we re-ran the original literature search from part two without the publication year limit. Although an additional model was identified the author declined participation in our group and therefore we chose to continue to invite investigators who used Bayesian methods until we met the target number of nine respondents.

Data Synthesis

Responses from members of the focus group were tallied, summarized, and reported in a de-identified format. Categorical data was summarized using frequencies and continuous data as means with SDs.

Results

Part One: Review of Existing Guidance Documents

Key Points

- Publicly available guidance discussing when and how to conduct a MTC as well as how to interpret and report the results of such analysis is summarized.
 - The majority of guidance is applicable to network meta-analysis in general, and not specific to MTC.
 - Guidance is provided from many organizations including: Health Information and Quality Authority, ISPOR, AHRQ Effective Health Care (EHC) Program, DERP, CRD, Canadian Agency for Drugs and Technologies in Health (CADTH), Australian Department of Health and Ageing, NICE, Health Care Knowledge Centre in Belgium, German Institute for Quality and Efficiency in Health Care, Haute Autorite de Sante, National Department of Health – Republic of South Africa and the Cochrane Collaboration.
 - Guidance from these organizations is not comprehensive and many aspects are not fully commented on. This reflects the lack of definitive evidence in the literature on these approaches and the need for future research.
- Either a Bayesian or Frequentist framework can be used to conduct a MTC.
- Limitations of the Lumley Frequentist method include: it is restricted to studies with at least one closed loop, it does not account for correlations that may exist between effect estimates when they are obtained from a single multi-arm trial, and there are weaknesses in situations where zero cells are common.
 - These limitations can be addressed through special preparations such as using a small increment to address zero cells and adding steps to adjust for correlations between effect estimates.
- Limitations of the Bayesian method include: it requires specification of noninformative priors, it is more complex to understand, and more difficult to use the software.
- Regardless of the method used to conduct the MTC, homogeneity and consistency of factors and event rates between direct and indirect comparisons is paramount if network meta-analysis is to be conducted.
 - Homogeneity and consistency should always be assessed for as an a priori component of the review process.
 - What is regarded as homogeneous and consistent enough is not well defined and is a subjective determination.
 - Some organizations recommend presenting direct and indirect evidence separately and if deemed consistent, performing network meta-analysis/MTC.
- Sensitivity analyses should include testing alternative specifications of the prior distribution to assess robustness of model results.
- For the Bayesian method, assessment and reporting of model fit is recommended.
- ISPOR provides a comprehensive checklist for conducting and synthesizing network meta-analysis including MTC.

- Reporting the study selection process, providing a description of included individual studies, and use of a graphical representation of the network results can help to improve transparency.

Detailed Analysis

Although our objective for part one was to focus on guidance for conducting a MTC, the majority of guidance available is applicable to network meta-analysis in general. When available, we also present guidance specific to MTCs, using either Bayesian or Frequentist methods.

General Description of Guidance Documents

Searches identified 25 relevant documents from which we extracted information. These included documents from regulatory/government-affiliated groups and nongovernmental organizations and collaborations involved in comparative effectiveness review and health technology assessment. Appendix A provides noteworthy verbatim statements from the 25 documents organized according to the categories listed in the Methods section. Most guidance is for network meta-analysis in general, regardless of the specific methodology used to conduct the analysis. The documents identified include:

- A guidance document from Health Information and Quality Authority (2011)⁷
- A two part guidance document from the ISPOR Task Force on Indirect Treatment Comparisons Good Research Practices (2011)^{9,10}
- A guidance document from AHRQ's EHC Program (2010)³
- A guidance document from DERP (2011)¹⁴
- A guidance document from CRD (2009)¹⁵
- Two guidance documents from CADTH (2006 and 2009)^{11,12}
- A published proceedings paper from the Ad Hoc Network Meta-analysis Methods Meeting Working Group (Li et al, 2011)¹⁶
- Two guidance documents from the Australian Government's Department of Health and Ageing (2008, 2008)^{17,18}
- A guidance document from the NICE (2008)¹⁹
- Seven guidance documents from NICE's Decision Support Unit (DSU) (each updated in 2012)²⁰
- A guidance document from the Health Care Knowledge Centre in Belgium (2008)²¹
- A guidance document from the German Institute for Quality and Efficiency in Health Care (2011)²²
- A guidance document from Haute Autorite de Sante (2009)²³
- A guidance document from the National Department of Health – Republic of South Africa (2010)²⁴
- Two guidance documents from the Cochrane Collaboration (2011)^{25,26}

When To Conduct a Network Meta-Analysis/MTC

The definition or meaning of the term network meta-analysis varies across the identified guidance documents.^{9-11, 23} Often these documents use terms such as “indirect treatment comparison,” “multiple treatment comparison,” “multiple treatment meta-analysis,” “mixed treatment meta-analysis,” (MTM) and “mixed treatment comparison” as synonyms for network meta-analysis.^{9,10,16,18,20,25} When used in this way, these terms are meant to represent the

simultaneous synthesis of evidence of all pairwise comparisons across three or more interventions.^{9,10,20,25} However, other documents use the above terms more definitely in order to differentiate the statistical analysis framework to be applied. Many guidance documents used the term MTC as we do in this report, specifically to describe a statistical approach used to analyze a network of evidence with more than two interventions which are being compared indirectly, and at least one pair of interventions compared both directly and indirectly.^{9,10,23} However, in some cases, MTC is referred to as “an extension”²⁵ or “special case”^{9,10} of only a Bayesian framework.^{7,10} Of note, a Bayesian framework can be used for, but is not restricted to, synthesizing networks with at least one closed loop.^{9,10,23} Lumley’s mixed model approach is used to describe one common Frequentist mixed model method for “analyzing a combination of direct and indirect evidence where there is at least one closed loop of evidence connecting the two technologies of interest.”^{7,11,12,23} Other similar mixed model methods exist.^{27,28}

A key component of nearly all documents is a discussion around when conducting a network meta-analysis is justified. Here the documents are almost entirely in agreement that synthesizing direct evidence only (from sufficient head-to-head or randomized controlled trials) “should take precedence”²⁵ or is “preferred”^{14,17} over analysis containing indirect evidence.^{19,21,22} However, in the absence of sufficient direct evidence, network meta-analysis “can be considered as an additional analytic tool”,^{3,19,21-23} although one document specifically states “pursuit of qualitative or quantitative indirect comparison is never required...”.¹⁴ In cases where analysis of both direct and indirect comparisons is undertaken, two guidance documents suggest the approaches should be considered and reported separately.^{19,25} Of note, a few documents^{9,10,19,23} appear to advocate for conducting MTC even in the presence of reasonable direct evidence, suggesting the combination of indirect and direct evidence may “add information that is not available from head-to-head comparison”,¹⁹ “strengthen the assessment between treatments directly evaluated,”^{9,10} and “yield a more refined and precise estimate of the interventions directly compared and broaden inference to the population sampled because it links and maximizes existing information within the network of treatment comparisons.”^{9,10}

An additional key discussion theme of identified guidance documents revolves around the assumptions of “homogeneity” and “consistency” (also referred to as “exchangeability” in some documents) that must be met in order to undertake network meta-analysis. Documents agreed that the validity of a network meta-analysis relies on the included studies or trials being similar in all factors (other than the intervention) that may affect outcomes, an assumption also important in standard pair-wise meta-analysis and that direct and indirect estimates are similar.

How To Conduct a Network Meta-Analysis/MTC

A number of the identified guidance documents reaffirmed that the same “good research practices” or “principles of good practice” used when conducting a traditional systematic review and meta-analyses should be carried over to conducting a network meta-analysis.^{9,10,19,23} These documents often reminded readers, “to minimize error and ensure validity of findings from meta-analyses, the systematic review, whether it involves a standard, pairwise meta-analysis or a network meta-analysis, must be designed rigorously and conducted carefully.”¹⁶ This includes an a priori declaration of the intent to conduct a network meta-analysis and clearly stating in the protocol the methods and implementation methods to be utilized.

A particularly variable area of focus of these documents includes strategies for systematically searching for studies. While many documents suggest following “conventional guidance” when conducting systematic literature searches for a network meta-analysis, some documents also

acknowledge the additional time and resources necessary to conduct a network meta-analysis search due to larger number of interventions to assess. While one document suggests an investigator might consider restricting a search to the minimum number of interventions of interest,⁷ another document emphasizes that “different specification of eligibility criteria may result in differences in the structure or extent of a network, leading to discrepant findings for network meta-analyses on the same topic.”²⁶ Moreover, many documents acknowledged that as more interventions are included in a network meta-analysis, the greater that uncertainty is reduced,²⁰ precision is increased²⁶ and “the ability to establish whether various sources of evidence ‘agree’ with each other” is enhanced.²⁶ In doing so, the documents suggest that network meta-analyses may need to include comparisons not of direct interest (e.g., placebo controls and therapies no longer used in current practice) as they may provide valuable information for the primary comparison(s) through indirect means.^{20,26} To this end, documents propose various strategies to balance validity and efficiency, and with the understanding that in some cases inclusion of therapies no longer used in clinical practice may at times be inappropriate as erroneous conclusions may be drawn on the efficacy and/or safety of these outdated treatments versus standards of care. Some guidance suggests these strategies include restricting to direct evidence only and broadening the search only after demonstrating that no direct data exists,¹⁷ using “iterative search methods” such as those proposed by Hawkins et al.,²⁹ and using previously published, good quality and up-to-date systematic reviews to augment a search.¹⁶ While not uniformly done, some guidelines state¹⁶ or imply²⁰ that evidence should be derived from RCTs only.^{16,17}

Perhaps the most comprehensive guidance on the planning and design of a network meta-analysis is available in the ISPOR document, which provided “a checklist of good research practices”.¹⁰ Below is the checklist, which includes guidance in the areas of search strategies, data collection, statistical analysis planning, data analysis and reporting (Table 1). Of note, the checklist often refers researchers to conventional guidelines on conducting meta-analysis.

Table 1. Checklist of good research practices for conducting and reporting network meta-analyses

Item	Recommendation
Search strategies	<ul style="list-style-type: none"> • Follow conventional guidelines for systematic literature searches; be explicit about search terms, literature, and time frames, and avoid use of ad hoc data • Consider iterative search methods to identify higher-order indirect comparisons that do not come up in the initial search focusing on lower-order indirect comparisons
Data collection	<ul style="list-style-type: none"> • Set forth evidence network demonstrating direct and indirect linkages between treatments, based on identified study reports • Follow conventional guidelines for data collection; use a prespecified protocol and data extraction form • Include sufficient study detail in data extraction to permit assessment of comparability and homogeneity (e.g., patient and study characteristics, comparators, and outcome measures)
Statistical analysis plan	<ul style="list-style-type: none"> • Prepare statistical analysis plan prior to data analysis, but permit modifications during data analysis, if necessary • Provide step-by-step descriptions of all analyses, including explicit statements of all assumptions and procedures for checking them • Describe analytic features specific to network meta-analysis, including comparability and homogeneity, synthesis, sensitivity analysis, subgroup analysis and meta-regression, and special types of outcomes

Table 1. Checklist of good research practices for conducting and reporting network meta-analyses (continued)

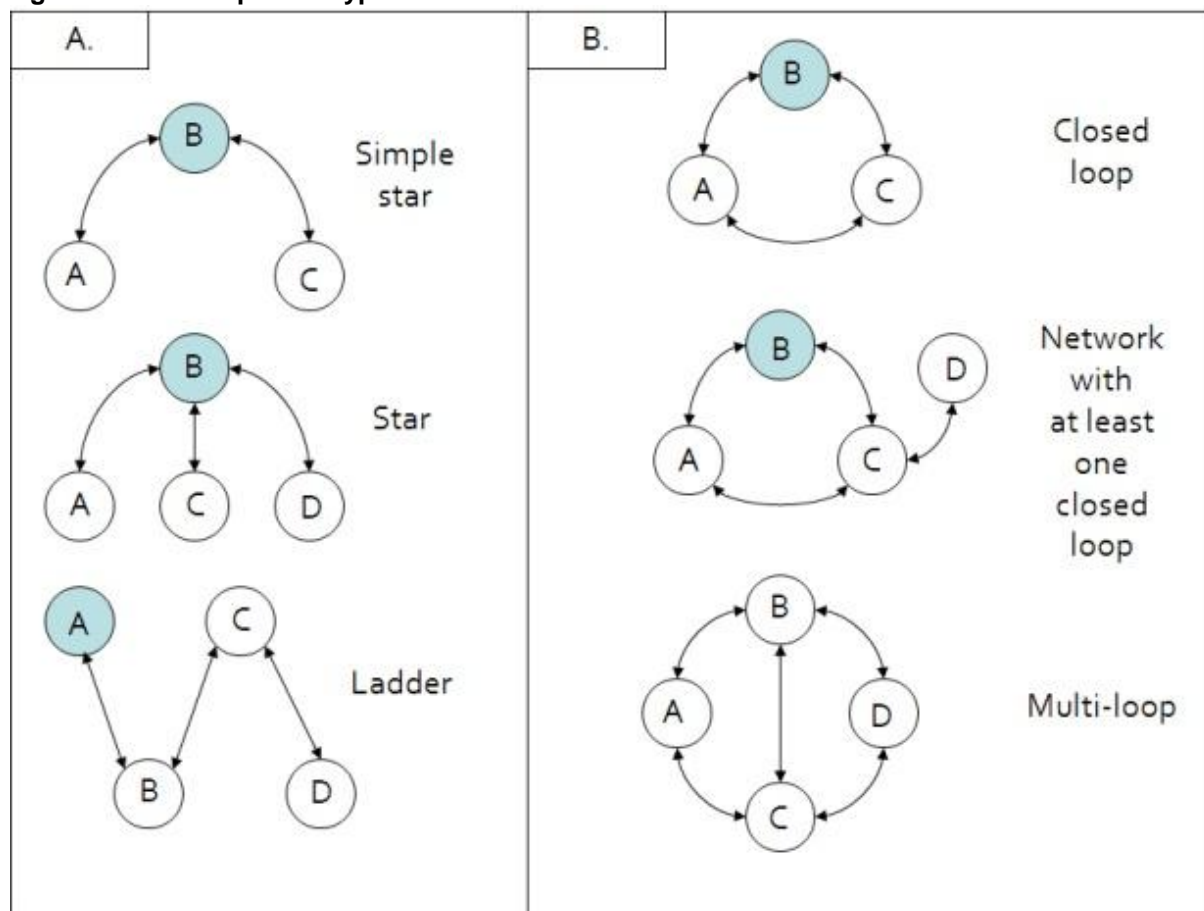
Item	Recommendation
Data analysis	<ul style="list-style-type: none"> • Follow conventional guidelines for statistical model diagnostics • Evaluate violations of similarity or consistency assumption in evidence network • If similarity or consistency is a problem, consider use of meta-regression models with treatment x covariate interactions to reduce bias
Reporting	<ul style="list-style-type: none"> • Follow PRISMA statement for reporting of meta-analysis • Explicitly state the study research questions (e.g., in Introduction or Objectives section of report) • Provide graphical depiction of evidence network • Indicate software package used in the analysis and provide code (at least in an online appendix)

Adapted with permission from: Hoaglin DC, Hawkins N, Jansen JP et al. Conducting indirect-treatment comparisons and network meta-analysis studies: report of the ISPOR task force on indirect treatment comparisons good research practices – part 2. *Value Health* 2011;14:429-437.

Abbreviations: PRISMA= preferred reporting in systematic review and meta-analysis

Many of the identified guidance documents provided advantages and disadvantages for the use of the different analysis frameworks (i.e., Frequentist and Bayesian methods) to network meta-analysis. Documents highlight that the “pattern” of the network of included studies may often dictate the framework used.^{7,9-12,23} Networks of studies that do not contain a “closed loop” such as a simple star, star or ladder pattern (Figure 1) cannot be analyzed using the Frequentist method described by Lumley, since a closed loop design is needed for calculating the estimate of incoherence, which is then used to construct 95% confidence intervals for the indirect estimate(s). However, those networks containing a closed loop (Figure 1) can be analyzed using the two of the more complex approaches, either Bayesian or Frequentist methods. The Bayesian method of conducting a MTC can be used to analyze any network pattern.

Figure 1. Network pattern types



Adapted from: Wells GA, Sultan SA, Chen L, Khan M, Coyle D. Indirect Evidence: Indirect Treatment Comparisons in Meta-Analysis. Ottawa: Canadian Agency for Drugs and Technologies in Health; 2009. Available at: <http://www.cadth.ca> (Last accessed on December 28, 2011).

Documents list a number of additional considerations when choosing between a Frequentist and Bayesian framework for analyzing these more complex closed loop networks of studies. Perhaps the most frequent consideration noted is the potential advantage of Bayesian methods in that “the method naturally leads to a decision framework that supports decisionmaking”^{9-11,23} by facilitating ranking of compared interventions.

With respect to statistical modeling, most guidance documents refer reviewers to the paper by Lumley (2002) for the statistical guidance in implementing Frequentist MTC when multi-arm trials are not present, including the necessary code. For MTCs with a Bayesian framework, the DSU of NICE has built a set of “core models” based upon the framework of generalized linear modeling.²⁰ The guidance document provides for Normal, Binomial, Poisson and Multinomial likelihoods, with identity, logit, log, complementary log-log, and probit link functions. Moreover, these “core models” can accommodate the assumptions of fixed-effect and random-effects settings, as well as multi-arm trials and multi-/shared parameter models for trials reporting results in different formats (trial versus group level data).

Identified guidance documents also comment on additional statistical modeling issues related to MTC conducted with either Frequentist or Bayesian methods. The merits of using a fixed- or random-effects model are discussed in a number of documents. While fundamentally, either a fixed or random-effects model can be used,^{9,10} at least one document¹⁸ states a preference for

using the random-effects approach “because the standard error obtained from a fixed effect analysis will be too small if there is heterogeneity between trials (beyond random variation)...”, and due to the fact “that there may be additional heterogeneity in an indirect comparison compared to a direct comparison.”¹⁸ A few documents acknowledge the potential benefit of incorporating study-level covariates into the model (extending the network to include treatment-by-covariate interactions or meta-regression); however, they also note concerns in the implementation as too few studies are often included in such meta-analyses which increases the potential for ecological bias. Guidance from NICE¹⁹ highlights that when a comparison of the results from single treatment arms from different RCTs is undertaken, the data must be treated as observational and appropriate steps taken to adjust for possible bias and increased uncertainty (including extending network to include treatment-by-covariate interactions or meta-regression). To this end, guidance typically suggest such naïve analyses “are completely untrustworthy” and should never be undertaken.²³

The implementation of Bayesian methods in a MTC was discussed in detail in many documents. Of note, the guidance from the Haute Autorite de Sante provides a detailed description of Markov chain Monte Carlo methods (simulation-based methods which can be used for the analysis of complex statistical models and to obtain estimates from distributions) and their use in MTC. While acknowledging the potential “arbitrary” nature of selection of priors (or priors whose form is not defended) in a MTC using Bayesian methods (particularly for between-study variance in a random-effects model), many of these documents suggested “vague” or perhaps more accurately described “noninformative priors” for such analyses, provided specific values (Appendix A) for different model parameters, and proposed alternative strategies for eliciting/determining priors (i.e., use of larger meta-analyses or expert clinicians in a field) when applicable. Documents also highlighted the need for checking convergence (i.e., running at least three chains, starting from widely different but sensible initial values, and examining posterior distributions visually for spikes and unwanted peculiarities) and running a “conservatively” large number of iterations for both the initial “burn-in” and the posterior sampling.

Additional statistical modeling discussion from the guidance documents included (1) the selection of the referent in MTC with Bayesian methods (as this can affect the posterior estimates), (2) the inappropriateness of treating multi-arm trials as if they were separate trials in a network meta-analysis (the correlation among the effect estimates of pair-wise comparisons must be taken into account), (3) the potential need for multi-/shared parameter models to address situations where trials report results in different formats (i.e., binomial data versus summary log odds and variance), and (4) the summary effect measure to be chosen with documents often recommending relative versus absolute measures due to concerns regarding varying baseline risk; and odds ratios as the preferred relative measure as they are symmetrical around the line of unity.

Nearly all guidance documents addressed identification and handling of potential bias and inconsistency in network meta-analyses. Inconsistency was commonly defined by documents as a conflict between “direct” evidence and “indirect” evidence of a comparison. As noted by one of the NICE guidance documents, “like heterogeneity, inconsistency is caused by effect-modifiers, and specifically by an imbalance in the distribution of effect modifiers in the direct and indirect evidence.”²⁰ Many documents reminded readers that network meta-analyses, like traditional meta-analysis, are akin to observational studies because the value of randomization does not hold across trials (albeit, they allow one to compare to or ore treatments that have not previously been directly compared, while maintaining the benefit of within trial

randomization).^{9,10,17,25} Consequently, they are prone to similar biases, particularly confounding bias. Other noted factors that might potentially influence effect estimates include the number of trials with two or more comparison arms and heterogeneity (as with traditional pair-wise meta-analysis).¹⁶

Documents unanimously agree that the “consistency” or “exchangeability” assumption must be assessed and should be an a priori component of the review protocol.^{7,9-26} Both the CADTH and Australian Government’s Department of Health and Ageing documents provide guidance for determining whether the “consistency” or “exchangeability” assumption is met based upon a detailed review of included studies (Appendix A). Both frameworks include an assessment of comparability of the common or “linking” treatment and comparability of patients in trials for presence of clinical or methodological heterogeneity. The Australian Government’s Department of Health and Ageing document more specifically suggests for the direct trials and indirect comparison, to assess whether the measure of comparative treatment effect is appropriate and assess the event rates of linking interventions. Another document further warned “with increased complexity and greater numbers of treatments, the prospect of inconsistency increases.”⁷

Documents also suggest more quantitative methods for detecting inconsistency between direct and indirect evidence. As noted in the ISPOR document, many regulatory agencies require the direct estimates and indirect estimates be calculated separately and shown to be consistent before they are combined. Within a Bayesian framework, a consistency model can be compared to an inconsistency model, with the residual deviance used as a test of “global inconsistency”. The same NICE DSU document that provided the core Bayesian code also provides these models to assess inconsistency.²⁰ Other, less favored, statistical methods noted by documents for detecting inconsistency include node splitting and use of measures of inconsistency variance.

Guidance documents are clear in their cautions about conducting network meta-analysis if the “consistency” assumption is not met. Unfortunately, as pointed out by one document, even if inconsistency is detected, “no commonly accepted standard [defines] which studies are similar enough”^{9,10} and that the determination is a “subjective” one. Moreover, some guidance documents stress that the validity of indirect comparisons may often be “unverifiable” because of limited detail in publications³ and the underpowered nature of detecting heterogeneity,^{18,20} and yet, another cautioned that inconsistency may affect different regions of a network of trials differently.¹⁶ Therefore, many documents provide more unwavering recommendations against network meta-analysis in the presence of inconsistency, while others make more flexible statements such as: “large inconsistencies rule out meta-analysis, small inconsistencies should add uncertainty to the results”^{11,12} and “...researchers must evaluate departures from consistency and determine how to interpret them.”⁹

A number of documents discussed the importance of assessing model fit when conducting a MTC using a Frequentist or Bayesian framework, both to aid in fixed- versus random-effects model (or other competing model, i.e., with or without covariate interaction) selection, and to demonstrate that the overall model fit is adequate. Examination of residual deviance (the lower the residual deviance the better the fit) and deviance information criteria (DIC) statistics were most commonly recommended when using a Bayesian approach.

Some of the guidance documents emphasized researchers should test alternate specifications of the prior distribution to assess robustness of model results. Noted assumptions to be tested in sensitivity analysis included different priors, inclusion or exclusion of covariate/modifiers in the model, and use of a fixed- or random-effects model.

How To Report and Interpret a Network Meta-Analysis/MTC

The proper interpretation of network meta-analyses is of paramount importance given their propensity to inform both clinical decisionmaking as well as coverage for third-party payers. A few guidance documents discussing the proper interpretation and reporting of network meta-analyses were identified in our literature search and are discussed here.^{7,9,10,16,19,23}

When interpreting the results of a network meta-analysis, it is important to consider the internal validity of the analyses as this “maximizes transparency and avoid(s) errors in interpretation.”¹⁶ This can be achieved by assessing the appropriateness of inclusion criteria of the evidence network, the quality of the included studies, and the existence of confounding bias.^{9,10,23} As mentioned previously, “good research practices” are necessary when conducting network meta-analyses, similar to traditional systematic reviews, and this includes use of “rigorous and extensive literature search methods”^{9,10} to minimize the potential for publication bias. Moreover, the validity of the network meta-analysis also hinges on the internal validity of the studies included in the review. It is recommended that “each study included in the network meta-analysis should be critically evaluated for bias.”^{9,10} One of these determinants should be the similarity between the included trials. This involves evaluating the clinical and methodological characteristics of the included studies in order to identify potential sources of bias and includes (but is not limited to) assessing differences in patient populations, methods for outcomes measurement, protocol requirements, duration of follow-up, and the time-frame the study was conducted.^{9,10,16} Differences in these characteristics could affect the integrity of the network and potentially impact interpretation of its results if a treatment-by-covariate interaction exists. An example would be significant differences in “baseline risks and placebo responses across trials” which “can reflect additional important differences in study or patient characteristics across studies.”^{9,10}

In addition to assessing the internal validity of both the included studies as well as the network itself, decisionmakers should assess the external validity of the meta-analysis’ findings and whether they apply to the population of interest.^{9,10} This is important since many clinical trials are conducted using selected and homogenous populations, which can compromise external validity. However, decisionmakers should embrace a certain level of dissimilarity between studies in a network meta-analysis, as this often times more closely reflects real-world clinical practice. It has been said that “some heterogeneity across trials in the network may arguably increase external validity.”^{9,10} This view should be interpreted with caution, as a high degree of heterogeneity within the direct comparisons may also significantly weaken the network and adversely affect its outputs.

As discussed above, probability statements regarding which intervention in a MTC is “best” are commonplace. It has been recommended that these “probability statements should be interpreted carefully since the difference between treatments might be small and not clinically meaningful.”¹⁶ Moreover, posterior probabilities resulting from a MTC using a Bayesian framework—which themselves are estimates and contain inherent random variability—may (in certain situations) lead to misinterpretation (of the relative efficacy of an evaluated intervention that can limit, rather than enhance, decision-making. For example, two interventions could demonstrate quite comparable safety and efficacy profiles (that is, be similar clinically), but may appear different based on their posterior probabilities. Additionally, this determination “cannot be made on the basis of efficacy endpoints alone.”^{9,10} This assessment should include evaluations of other available safety and effectiveness data not included in the network meta-analysis, including observational evidence. This will provide a more detailed picture of the totality of

information for the intervention and allow the decisionmaker to more properly assess its place in medical practice.

Guidance from the Haute Autorite de Sante provides a brief “critical review guide” with suggests users of network meta-analyses/MTC consider the following to evaluate its validity/usefulness: (1) acceptability of the approach used; (2) search strategy and selection process for data contributing to the indirect comparison calculations; (3) clinical homogeneity of trials and stability of effects; (4) consistency of estimates; (5) degree of concordance of the result with that of existing direct comparisons; and (6) correct interpretation of results in the proposed conclusions.²³ Similar guidance has recently been provided by The NICE Decision Support Unit in the form of a “reviewer checklist” for evidence synthesis reports, which addresses “issues specific to network synthesis” including: (1) adequacy of information on model specification and software implementation, (2) multi-arm trials; (3) connected and disconnected networks; and (4) inconsistency”.²⁰

Guidance documents have been published providing recommendations for the proper reporting of indirect comparisons and network meta-analyses.^{9,10,13,16,30,31} A Task Force on Indirect Treatment Comparisons Good Research Practices by the ISPOR has proposed a simplified checklist to assist decisionmakers in the proper evaluation of a reported network meta-analysis.^{9,10} The items included by this task force are included in Table 2. It should be noted that this list is not all-inclusive and does not include enough information to adequately assess either the internal or external validity of an indirect comparison or network meta-analysis.

Table 2. International Society for Pharmacoeconomics and Outcomes Guidance for the reporting of a network meta-analysis

Report Section	Checklist Item	What To Look For in the Paper
Introduction	Are the rationale for the study and the study objectives stated clearly?	A clear rationale for the review
Methods	Does the methods section include the following? Description of eligibility criteria Information sources Search strategy Study selection process Data extraction (validity/quality assessment of individual studies)	A systematic review of the literature in accordance with Centre for Reviews and Dissemination guidelines and PRISMA

Table 2. International Society for Pharmacoeconomics and Outcomes Guidance for the reporting of a network meta-analysis (continued)

Report Section	Checklist Item	What To Look For in the Paper
	Are the outcome measures described?	Justification of outcome measures selected for analysis
	Is there a description of methods for analysis/synthesis of evidence? Do the methods described include the following? Description of analyses methods/models Handling of potential bias/inconsistency Analysis framework	Description and justification of statistical model(s) used: multiple meta-analysis of pairwise comparisons versus network meta-analysis models; fixed- versus random-effects models; models without or with covariate (interactions) Description of whether analyses were performed with a frequentist or Bayesian approach Description of how possible bias/inconsistency was evaluated (either qualitative or quantitative, e.g., comparison of direct evidence with the indirect evidence). If meta-regression models are used, rationale for selection of covariates in models Description of relative-effect estimates used for presentation of findings (e.g., odds ratio, relative risk, hazard ratio, difference in change from baseline) Description of whether relative-effect measures were transformed into expected (absolute) outcomes (e.g., proportion of responders)
	Are sensitivity analyses presented?	Rationale for and description of sensitivity analyses Studies included Prior distributions for model parameters in Bayesian framework
Results	Do the results include a summary of the studies included in the network of evidence? Individual study data? Network of studies?	Description of results of study identification and selection process Table/list of studies with information regarding study design and patient characteristics (that might act as effect modifiers); these are important to judge potential similarity/consistency issues Figure of network of studies Table with raw data by study and treatment as used for the analysis/model. (Optionally present relative effects of available direct comparisons of each study)
	Does the study describe an assessment of model fit? Are competing models being compared?	Justification of model results
	Are the results of the evidence synthesis (ITC/MTC) presented clearly?	Table/ figure with results for the pairwise comparisons as obtained with analyses; Point estimates and measure of uncertainty (95% CIs) In Bayesian framework, probability to reflect decision uncertainty (i.e., probability of which treatment is best if multiple treatments are being compared and probability that one treatment is better than the comparator)
	Sensitivity/scenario analyses	Description of (different) findings with sensitivity/scenario analysis

Table 2. International Society for Pharmacoeconomics and Outcomes Guidance for the reporting of a network meta-analysis (continued)

Report Section	Checklist Item	What To Look For in the Paper
Discussion	Does the discussion include the following? Description/summary of main findings Internal validity of analysis External validity Implications of results for target audience	Summary of findings Internal validity (individual trials, publication bias, differences across trials that might violate similarity and consistency assumptions) Discussion regarding generalizability of findings (given patient population within and across trials in network) Interpretation of results from a biological and clinical perspective

Adapted with permission from: Jansen JP, Fleurence R, Devine B et al. Interpreting indirect treatment comparisons and network meta-analysis for health-care decisionmaking: report of the ISPOR task force on indirect treatment comparisons good research practices: part 1. *Value Health* 2011;14:417-428.

Abbreviations: CI=confidence interval; ITC=indirect treatment comparison; MTC=mixed treatment comparison; PRISMA=preferred reporting of systematic reviews and meta-analysis

This guidance document provides recommendations on items that should be included in the introduction, methods, results, and discussion sections of a network meta-analysis report as well as a detailed description of what to look for in each of these sections.^{9,10} Many of the items discussed overlap with guidance on the proper reporting of traditional meta-analyses.³² Aspects unique to conducting a network meta-analysis deserve special mention, much of which involves appropriate reporting of methods and results. If a Bayesian framework was used to perform the data analysis, it is recommended that “the choice of prior distributions for the model parameters should be defined.”^{9,10} If sensitivity analyses were conducted evaluating the prior distribution assumptions, these results should be also reported. In addition, the software package used to analyze the data as well as the written code from the program should be provided, “at least in an online appendix.”^{9,10}

When reporting the results of a network meta-analysis, the ISPOR Task Force suggests that a graphical representation of the network be provided to “improve transparency of the analyses.”^{9,10} In addition to discussing the study selection process and description of the individual studies, the report should provide results of both the pairwise comparisons as well as indirect treatment comparisons.^{9,10,19} It has also been recommended that investigators “explain the difference between direct and indirect evidence based upon study characteristics.”^{3,19} Additionally recommended items for good reporting include goodness-of-fit of the data as well as calculations of residual deviance.^{9,10}

Additional guidance documents for reporting of studies using a Bayesian framework come from the Reporting Of Bayes Used in clinical Studies (ROBUST) criteria, BayesWatch (Bayesian analysis in biomedical research), and Bayesian Standards in Science (BaSiS).^{13,30,31} Although these documents are intended for Bayesian analyses in general, they can also be applied to meta-analyses as well. The ROBUST criteria suggests that the following information should be included in any Bayesian study report: prior distributions used, including specified, justified, and sensitivity analysis, analyses run including the statistical model and analytical techniques, and results including central tendency, standard deviation or credible intervals/Bayesian confidence interval (an interval in the domain of a posterior probability distribution used for interval estimation).¹³ The BayesWatch and BaSiS include more technical and computational items such as information about the model itself, including details about the software used, if Markov Chain Monte Carlo simulation was used, and if so the number and length of runs as well as convergence diagnostics, shape of the posterior densities, and use of

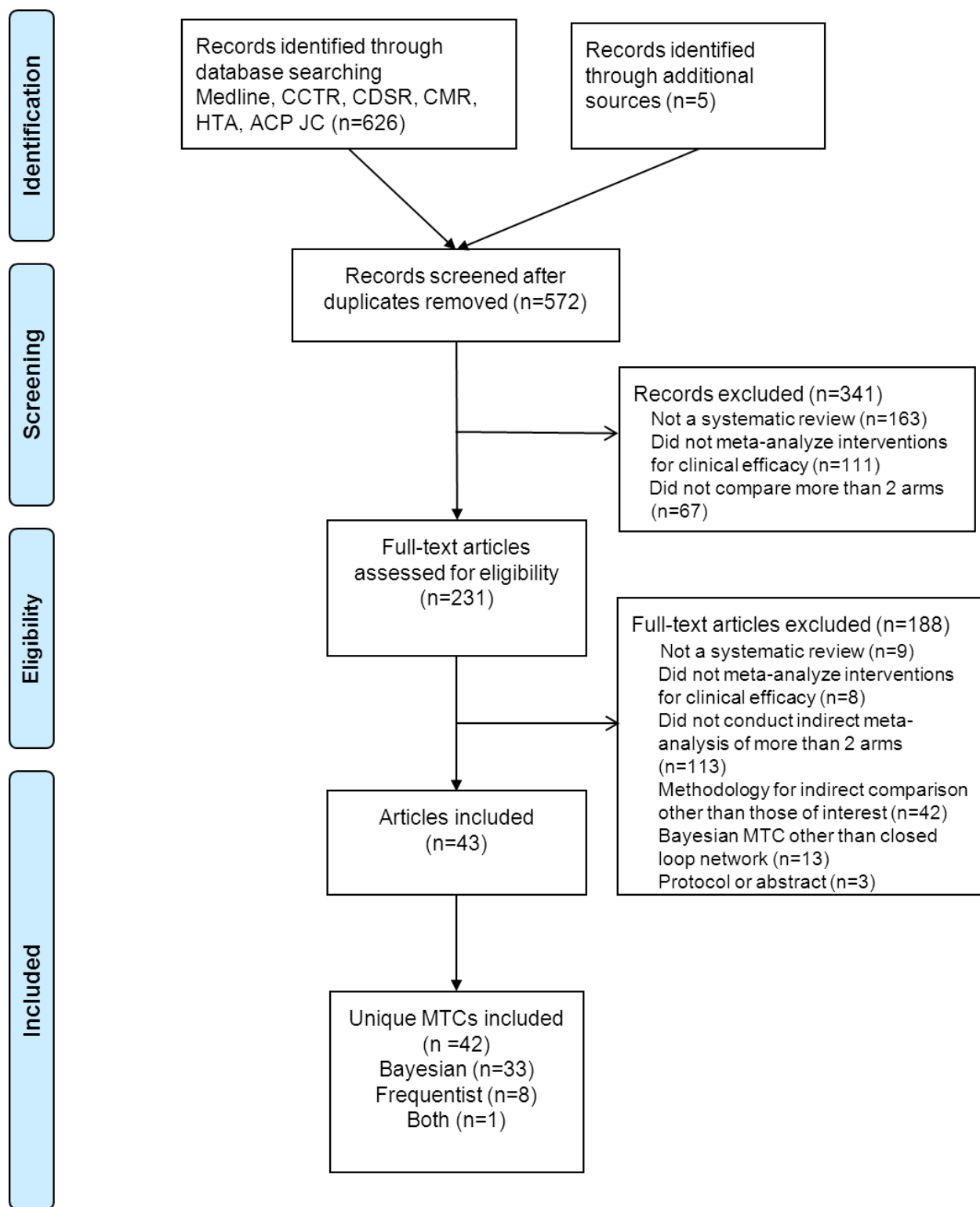
appropriate Bayes factors, amongst others.^{30,31} It has been questioned whether these more detailed requirements are important to include for a clinical journal and should be reserved for a more methodologically focused periodical.¹³

Part Two: Systematic Review of Existing MTCs

Results of the Literature Search

A total of 626 citations were identified through the database search with an additional five citations identified manually (Figure 2). After duplicates were removed, 572 citations remained and were screened at the abstract level. Of the abstracts reviewed, 341 were excluded and 231 were considered at the full-text level. After full-text review, 44 articles representing 43 unique MTCs that utilized either Bayesian or Frequentist methods to conduct a MTC were included. A list of excluded studies can be found in Appendix E.

Figure 2. Inclusion of identified citations for part two of this project



Abbreviations: ACP JC= American College of Physicians Journal Club; CCTR=Cochrane Central Register of Controlled Trials; CDSR=Cochrane Database of Systematic Reviews; CMR=Cochrane Methodology Register; HTA=Health technology Assessment; MTC=mixed treatment comparison

Key Points

- Of the included MTCs, the majority use Bayesian methods.
- Thirty-four unique MTCs that used Bayesian methods were identified and were conducted in 10 different countries. Thirteen disease categories were evaluated, with the most common being cardiovascular. Most analyses were funded government/foundation sources.
 - Pharmacologic interventions were evaluated in the majority of networks.
 - The statistical code was rarely made available to the reader, although raw data was commonly published.
 - A similar percent of MTCs either reported using vague priors or did not specify if the priors were intended to be vague or informative. Few models declared using informative priors. It was uncommon to find specific priors, and may be related to lack of code reporting. However, the majority of journals that published these MTCs allowed supplement or appendix publication and several manuscripts did utilize this option.
 - Random effects models were used in the majority of MTCs. A broad range of methods were used to evaluate convergence, heterogeneity, and inconsistency. Unfortunately, lack of reporting within manuscripts may or may not mean such evaluations were omitted.
 - It was common for authors to rank order interventions based on the probability of the intervention being best for a given outcome. Rarely did authors conclude equivalence or non-inferiority of interventions based on MTC results.
 - Most MTCs evaluated binary outcomes and reported results as odds ratios or relative risks. However, most MTCs did not specify whether these were mean or median values of the posterior distribution. All models reported 95 percent credible intervals. Of the models that reported continuous outcomes, the weighted mean difference was the effect measure used almost exclusively.
 - A mixture of tables, text, and figures was commonly used to report results of the MTCs.
- Nine MTCs used Frequentist methods.
 - These MTCs were conducted in five different countries and evaluated five disease categories including cardiology, behavioral health, pain management, rheumatology and gastro-urology.
 - Three analyses specifically referenced/used Lumley's MTC method.
 - Most analyses evaluated pharmacologic interventions with on average 7.3 interventions evaluated.
 - Eight MTCs included a traditional meta-analysis as well. It was more common for heterogeneity to be evaluated in the traditional meta-analysis than in the network meta-analysis. The majority of MTCs evaluated inconsistency.
 - None of the MTCs made claims of equivalence, non-inferiority, or defined minimally important differences. Most analyses reported binary outcomes with the majority using odds ratios as the effect estimates. All analyses reported variance using 95 percent confidence intervals.

Detailed Analysis

The results are first presented for the journals in which identified MTCs were published followed by results according to the method used to conduct the MTC, either Bayesian or Frequentist. When applicable, mean values are accompanied by SDs (mean±SD). Text and tables do not duplicate in all cases and either format may have been used to present data.

Journal-Level Characteristics

Our systematic literature search identified 42 unique MTCs that used either Bayesian or Frequentist methods to conduct MTC. The majority of MTCs used Bayesian methods (33 out of 42, 78.6 percent)³³⁻⁶⁶ and few used Frequentist methods (8 out of 42, 19.0 percent).⁶⁸⁻⁷⁵ One review (2.4 percent) used both methods.⁶⁷ Complete details of each journal in which at least one review was published and the journal's characteristics can be found in Appendix Table F-1. The 42 MTCs were published in 32 different journals, with a mean impact factor of 8.67±8.1 (Table 3). The journal which had the highest number of MTC published was the British Medical Journal (5 of the 42 reviews, 11.9 percent). The majority of journals allowed online supplements or appendices, and also imposed word count limits (Table 3). However, the majority of these journals did not impose limitations on the number of tables or figures allowed.

Table 3. Journal-level characteristics

Characteristic	Yes n/N (%) or Mean (SD)
Impact factor	8.67 (8.1)
Supplement or appendix allowed	25/32 (78.1)
Online	22/25 (88.0)
Not specified	3/25 (12.0)
Word count limit	21/32 (65.6)
Table count limit	14/32 (43.8)
Figure count limit	14/32 (43.8)

MTC Using Bayesian Methods

A summary of the results of Bayesian MTCs can be found in Table 4 to Table 6. Detailed characteristics of each analysis can be found in Appendix Tables F-2 to F-4. One analysis used both Bayesian and Frequentist methods and is considered in both sections of the results.⁶⁷ The analysis by Orme et al.⁴³ included two individual networks and whether this analysis was considered once or twice for a given characteristic is defined within table legends.

General Characteristics

The majority of identified MTCs identified in our literature search used Bayesian methods to conduct the analysis (81.0 percent). On average, 6.1±4.8 authors were listed per publication and the majority of publications (52.9 percent) did not include a methodologist as an author. The most common country from which authors published reviews was the United Kingdom (35.3 percent), followed by the United States (11.8 percent) and Greece (11.8 percent). The remaining analyses were published in a variety of countries (Table 4). The majority of analyses were funded by government/foundation sources (29.1 percent), followed by industry (26.5 percent) and analysis which did not report funding sources (23.6 percent). Only two analyses (5.9 percent) identified an affiliation, one each with the Health Technology Assessment Program and The Cochrane Collaboration. The mean number of printed pages per publication was 16.6±36.3 and over half (58.8 percent) published a supplement or appendix. Only one publication from

those that did not publish a supplement or appendix did not have the option given the journal specifications and one was an affiliated report that did not have a word or page restriction.

There were 13 different categories of disease states evaluated with a wide dispersion of categories. The most common category was cardiology (17.6 percent) (Table 4). The mean number of interventions included within the analyses was 8.5 ± 4.3 . The majority of analyses evaluated pharmacologic interventions (85.7 percent) with few evaluating devices (8.5 percent) or other interventions (2.9 percent), such as blood glucose monitoring. One analysis included both pharmacologic interventions and devices (2.9 percent). The mean number of trials included in the analyses was 35.9 ± 30.1 and the mean number of patients included was $33,460 \pm 71,233$.

Table 4. General characteristics of Bayesian mixed treatment comparisons

Characteristic	n/N (%) or Mean (SD)
Number of authors	6.1 (4.8)
Was a methodologist an author on the manuscript?*	16/34 (47.1)
Country	
U.S.A.	4/34 (11.8)
United Kingdom	12/34 (35.3)
Canada	2/34 (5.9)
Brazil	1/34 (2.9)
China	2/34 (5.9)
Switzerland	3/34 (8.8)
Netherlands	1/34 (2.9)
Italy	3/34 (8.8)
Belgium	1/34 (2.9)
Greece	4/34 (11.8)
Funding	
Industry	9/34 (26.5)
Government/Foundation	10/34 (29.4)
Unfunded	6/34 (17.6)
Other	1/34 (2.9)
Not reported	8/34 (23.6)
Declared affiliation	2/34 (5.9)
Health Technology Assessment Program	1/2 (50.0)
The Cochrane Collaboration	1/2 (50.0)
Number of printed pages	16.6 (36.3)
Supplement or appendix published	20/34 (58.8)
Disease state evaluated	
Behavioral health	4/34 (11.8)
Cardiology	6/34 (17.6)
Infectious disease	2/34 (5.9)
Endocrine	2/34 (5.9)
Pulmonary	2/34 (5.9)
Pain	3/34 (8.8)
Dermatology	2/34 (5.9)
Ophthalmology	2/34 (5.9)
Rheumatology	2/34 (5.9)
Gastroenterology	3/34 (8.8)
Dental	1/34 (2.9)
Oncology	4/34 (11.8)
Substance abuse	1/34 (2.9)
Number of interventions compared*	8.5 (4.3)

Table 4. General characteristics of Bayesian mixed treatment comparisons (continued)

Characteristic	n/N (%) or Mean (SD)
Type of intervention*	
Pharmacologic	30/35 (85.7)
Devices	3/35 (8.6)
Other	1/35 (2.9)
Device and pharmacologic	1/35 (2.9)
Number of trials included in network*	35.9 (30.1)
Number of patients included in network*	33,459 (71,233)

*The trial by Orme et al. included two individual networks and they are considered separately for this characteristic.

Methods Characteristics

The majority of analyses also included a traditional meta-analysis (76.5 percent) (Table 5). The most common model used in Bayesian MTCs was a random-effects model (58.8 percent), followed by both a random and fixed effects model (20.6 percent), unspecified (17.6 percent), or a fixed-effects model (2.9 percent). The majority of analyses did not report information about whether there was adjustment for multiple arms (82.4 percent) or adjustment for covariates (73.8 percent). Less than half of the analyses reported testing the model fit (44.1 percent), while the remaining did not comment on testing model fit. Of the 15 analyses that reported tested model fit, the most common method was use of residual deviance (40.0 percent) followed by using both residual deviance and the deviance information criterion (20.0 percent), solely the deviance information criterion (13.3 percent), unspecified methods (13.3 percent), mean sum deviation (6.7 percent), or Q-Q plots (6.7 percent).

All analyses used WinBUGS software. Two analyses also further specified additional software including BUGS XLA Wrapper and S-Plus. The majority of analyses did not make their code available to the reader (79.4 percent), although of the seven analyses that did provide the code (20.6 percent) the most common presentation was within the online supplement (five MTCs, 71.4 percent). Raw data was frequently available to the reader (61.8 percent of MTCs) and of the 21 analyses that published raw data, the most common format was within the manuscript itself (18 MTCs, 85.7 percent). Most analyses did not report evaluating convergence (64.7 percent). Of the 12 analyses (35.3 percent) that did evaluate convergence, the most common method was the Gelman Rubin statistic (58.8 percent), although several less frequent methods were used as well (Table 5). Totals of each individual method combined may not add up to the number of studies because one study may have used multiple methods.

Most analyses did not report whether the priors used were considered vague or informative (47.1 percent) while 44.1 percent of MTCs specifically described the prior distributions used as vague or non-informative. The remaining 8.8 percent of analyses used informative priors. It was uncommon for the actual prior distribution to be reported for the population treatment effect (d) and the between-study standard deviation of population treatment differences across studies (σ), as only 32.1 percent and 29.4 percent of MTCs, respectively, reported the actual priors. Most analyses did not perform sensitivity analysis based on the priors used (88.2 percent).

Evaluation of heterogeneity within traditional meta-analyses was common (16 out of 26 MTCs that included a traditional meta-analysis, 61.5 percent). Some reported multiple means to test for heterogeneity and therefore the totals of each individual method combined may not add up to the number of studies. The most common method used was the I^2 statistic (81.3 percent) followed by the Cochrane Q-statistic (43.8 percent), among many less frequent methods (Table

5). Evaluation of heterogeneity within the MTC was less common, reported in only 32.4 percent. Some analyses reported multiple means to test for heterogeneity and therefore totals of each individual method combined may not add up to the number of studies. Of these 11 analyses, the most common method used to assess heterogeneity was τ^2 (54.5 percent) followed by between study standard deviation (45.5 percent), among several other less frequent methods (Table 5).

Inconsistency was evaluated in 70.6 percent of analyses. One review reported being unable to evaluate inconsistency due to lack of direct data while the remaining MTCs (10 MTCs, 29.4 percent) did not report evaluating inconsistency. Totals of each individual method combined may not add up to the number of studies because one study may have used multiple methods. The majority of the 24 analyses that evaluated inconsistency did so through comparison of the results with either the results of their traditional meta-analysis or a previously conducted meta-analysis (50.0 percent) followed by unspecified methods (33.3 percent), among several others (Table 5).

Table 5. Methods characteristics in Bayesian mixed treatment comparisons

Characteristic	n/N (%)
Conducted traditional meta-analysis	26/34 (76.5)
Model	
Fixed effects	1/34 (2.9)
Random effects	20/34 (58.8)
Fixed and random effects	7/34 (20.6)
Not reported	6/34 (17.6)
Adjustment for covariates	9/34 (25.6)
Adjustment for multiple arms	6/34 (17.6)
Model fit tested	15/34 (44.1)
Residual deviance	6/15 (40.0)
Deviance information criterion	2/15 (13.3)
Residual deviance and deviance information criterion	3/15 (20.0)
Q-Q plots	1/15 (6.7)
Mean sum deviation	1/15 (6.7)
Method not reported	2/15 (13.3)
Code published	7/34 (20.6)
Online supplement	5/7 (71.4)
External Web site	2/7 (28.6)
Raw data published	21/34 (61.8)
Manuscript	18/21 (85.7)
Online supplement	2/21 (9.5)
External Web site	1/21 (4.8)
Evaluation of convergence*	12/34 (35.3)
Gelman Rubin statistic	7/12 (58.3)
Kernel density plot	1/12(8.3)
Visual plot inspection	1/12 (8.3)
Observation of chain mix	2/12 (16.7)
Method not reported	2/12(16.7)
Priors	
Use of noninformative	15/34 (44.1)
Use of informative priors	3/34(8.8)
Not specified	16/34 (47.1)
Prior distribution of μ reported	11/34 (32.4)
Prior distribution for sigma reported	10/34(29.4)
Sensitivity analysis based on priors	4/34 (11.8)

Table 5. Methods characteristics in Bayesian mixed treatment comparisons (continued)

Characteristic	n/N (%)
Evaluation of heterogeneity in traditional meta-analysis*	16/26(61.5)
I^2	13/16 (81.3)
Cochrane-Q statistic	7/16 (43.8)
PICO statement	1/16(6.3)
Plot visualization	2/16 (12.5)
L'Abbe plot	1/16 (6.3)
Evaluation of heterogeneity in network meta-analysis*	11/34(32.4)
Precision (τ^2)	6/11 (54.5)
Between study SD	5/11(45.5)
Heterogeneity p-values	1/11 (9.1)
Evaluation of inconsistency*	24/34 (70.6)
Comparison to traditional or prior meta-analysis	12/24 (50.0)
Inconsistency/incoherence factors	4/12 (33.3)
Posterior mean residual deviance	3/12 (25.0)
Method not reported	4/12 (33.3)
Trial sequential analysis	1/12 (8.3)
Overall inconsistency (σ^2w)	1/12 (8.3)

*Studies that used multiple methods to test heterogeneity were counted multiple times, in the respective categories.

Abbreviations: PICO=patient, intervention, comparator, outcome; SD=standard deviation

Outcome and Results Reporting

Few analyses presented graphical representation of the posterior distribution of outcomes (8.8 percent) (Table 6). The use of rank ordering of interventions based on the probability the given intervention was the best for a given outcome was reported in 61.8 percent of analyses. Only one analysis made claims of equivalence (2.9 percent) and two made claims of non-inferiority (5.9 percent). Of the three analyses that made claims of equivalence or non-inferiority, two defined a minimally important difference. Four (11.8 percent) analyses defined minimally important differences although did not make specific claims of equivalence or non-inferiority.

Most analyses reported outcomes that were binary (67.6 percent) followed by both binary and continuous outcomes (17.6), solely continuous outcomes (11.8 percent), and one reported on a categorical non-binary outcome (2.9 percent). Of the 29 analyses that reported binary outcomes, odds ratios were the most commonly reported effect measure (62.1 percent), followed by relative risks (17.2 percent) and hazard ratios (13.8 percent), among other less frequent measures. Of the 10 analyses that reported continuous outcomes, the weighted-mean difference was the most common effect measure (80.0 percent). Two network meta-analyses used multiple effect measures including standardized mean difference and a measure specific to the content (e.g., prevention fraction in a dental analysis). The one analysis that reported a categorical non-binary outcome used relative risk to measure effect. All analyses reported variance with 95 percent credible intervals and one also reported standard errors. Most analyses (85.3 percent) did not report if the posterior distribution was the mean or median value. Presentation of results data varied although most analyses used multiple media (and were therefore counted multiple times) including tables, figures, and text. Of the 34 analyses, 32 used text (94.1 percent), 24 used tables (70.6 percent), and 21 used figures (61.8 percent) to present results.

Table 6. Outcomes and results reporting in Bayesian mixed treatment comparisons

Characteristic	n/N (%) or Mean (SD)
Ranking of outcomes	21/34 (61.8)
Graphical representation of posterior distribution	3/34 (8.8)
Posterior distribution	
Mean	1/34 (2.9)
Median	4/34 (11.8)
Not reported	29/34 (85.3)
Claims of equivalence	1/34 (2.9)
Claims of non-inferiority	2/34 (5.9)
Minimally important difference	8/47 (17.0)
Type of outcome	
Binary	23/34 (67.6)
Continuous	4/34 (11.8)
Binary and continuous	6/34 (17.6)
Categorical non-binary	1/34 (2.9)
Binary effect measure	29/34 (85.3)
Relative risk	5/29 (17.2)
Odds ratio	18/29 (62.1)
Hazard ratio	4/29 (13.8)
Multiple	2/39 (6.9)
Continuous effect measure	10/34 (29.4)
Weighted mean difference	8/10 (80.0)
Multiple	2/10 (20.0)
Categorical non-binary effect measure	1/34 (2.9)
Relative risk	1/1 (100)
Presentation of Results*	
Table	24/34 (70.6)
Text	32/34 (94.1)
Figure	21/34 (61.8)

*Studies were counted multiple times when more than one method was used.

Frequentist MTCs

A summary of the results of MTCs that used Frequentist methods can be found in Table 7 to Table 9. Detailed characteristics for each analysis can be found in Appendix Tables F-5 to F-7. One analysis used both Bayesian and Frequentist methods and is considered in both sections of the results.⁶⁷ When applicable, mean values are accompanied by SDs (mean±SD).

General Characteristics

A minority of the analyses identified by our systematic review used Frequentist methods (nine MTCs, 20.9 percent). Again, one MTC used both Bayesian and Frequentist methods.⁶⁷ On average, 7.1±5.4 authors were listed per publication and a majority of publications were not considered to have a methodologist as an author (44.4 percent) (Table 7). The most common country from which authors published these MTCs were from the United States (44.4 percent), followed by the United Kingdom (22.2 percent) and France (22.2 percent). The majority of analyses were funded by government/foundation sources (44.4 percent) followed by industry (33.3 percent) among other sources. Two analyses identified an affiliation, one each with the Health Technology Assessment Program and the Cochrane Collaboration. The mean number of printed pages per publication was 16.1±16.0 and most of the publications (66.7 percent) published supplements or appendices. The two MTC with affiliations were those without a supplement.

There were five different categories of disease states evaluated in the analyses with the most in cardiology (33.3). The mean number of interventions included within the evaluated analyses

was 7.3±2.8. Eight analyses evaluated pharmacologic interventions (88.9 percent) while one evaluated multiple intervention types (11.1 percent). The mean number of trials included in the analyses was 59.0 ±51.9 and the mean number of patients included was 59615±70268.

Table 7. General characteristics of Frequentist mixed treatment comparisons

Characteristic	n/N (%) or Mean (SD)
Number of authors	7.1 (5.4)
Was a methodologist an author on the manuscript?	4/9 (44.4)
Country	
U.S.A.	4/9 (44.4)
United Kingdom	2/9 (22.2)
France	2/9 (22.2)
Thailand	1/9 (11.1)
Funding	
Industry	3/9 (33.3)
Government/Foundation	4/9 (44.4)
Other	1/9 (11.1)
Unknown	1/9 (11.1)
Declared affiliation	2/9 (22.2)
Health Technology Assessment Program	1/2 (50.0)
Cochrane	1/2 (50.0)
Number of printed pages	16.1 (16.0)
Supplement or appendix published	6/9 (66.7)
Disease state evaluated	
Behavioral Health	3/9 (33.3)
Cardiology	4/9 (44.4)
Gastro-urology	1/9 (11.1)
Rheumatology	1/9 (11.1)
Number of interventions compared	7.3 (2.8)
Type of intervention	
Pharmacologic	8/9 (88.9)
Procedure, device and pharmacologic	1/9 (11.1)
Number of trials included in network	59.0 (51.9)
Number of patients included in network	59615 (70268)

Methods Characteristics

Eight of the nine MTCs also included a traditional meta-analysis. The language used to describe the model implemented in each analysis was heterogeneous and can be found in Appendix Table F-6. Of note, three MTCs specifically referenced use of Frequentist methods described by Lumley⁷⁰⁻⁷² and the other 6 analyses used other mixed model approaches for Frequentist MTC.^{67-69,73,74} Weighting of studies was not reported in most analyses (88.9) while one (11.1 percent) weighted studies using inverse variance (Table 8). Two analyses (22.2 percent) adjusted the model for covariates while the others did not report whether adjustments were made or not. Raw data was available in most analyses (88.9 percent) and of the eight that published raw data, the format was mostly within the manuscript itself (62.5 percent) as opposed to an online supplement (37.5). Three analyses (37.5 percent) used R as the software while three (37.5 percent) used SAS, one used Stata (11.1 percent) while the last did not report software used.

Heterogeneity within traditional meta-analyses was evaluated in four of eight reviews (50.0 percent) that conducted a traditional meta-analysis. The most common method used in these four analyses was the I^2 statistic (50.0 percent) while one analysis used both the I^2 statistic and the Cochrane-Q statistic (25.0 percent) and one used the Riley Day test (25.0 percent). Evaluation of heterogeneity within network meta-analyses was less common, reported in only two of the nine

analyses (22.2 percent). One used covariance statistics and standard error and one used tau². Inconsistency was evaluated in eight of the nine analyses. The majority of analyses (62.5 percent) evaluated inconsistency by comparing results from the MTC to either the traditional meta-analysis or previously published literature. Other methods reported to evaluate inconsistency included evaluating incoherence values (25.0 percent) and t-tests based on odds ratios from the traditional and network meta-analyses (12.5 percent).

Table 8. Methods characteristics in Frequentist mixed treatment comparisons

Characteristic	n/N (%)
Weighting of studies	
Inverse variance	1/9 (11.1)
Not reported	8/9 (88.9)
Adjustment for covariates	2/9 (22.2)
Raw data published	8/9 (88.9)
Manuscript	5/8 (62.5)
Online supplement	3/8 (37.5)
Heterogeneity assessed in traditional meta-analysis	4/8 (50.0)
I ²	2/4 (50.0)
Cochrane-Q statistic and I ²	1/4 (25.0)
Riley Day test	1/4 (25.0)
Heterogeneity assessed in network meta-analysis	2/9 (22.2)
Tau ²	1/2 (50.0)
Covariance and SE	1/2 (50.0)
Inconsistency evaluation	8/9 (88.9)
Comparison to traditional or prior meta-analysis	5/8 (62.5)
Incoherence statistic	2/8 (25.0)
T-test	1/8 (12.5)

Outcome and Results Reporting

None of the analyses made claims of equivalence, noninferiority, or defined a minimally important difference (Table 9). Seven analyses reported outcomes that were binary (77.8 percent) while one analysis reported continuous outcomes and the last reported both outcome types. Of the eight analyses that reported binary outcomes, most used odds ratios as effect measures. All analyses reported variance with 95 percent confidence intervals. Presentation of results data varied although most reviews used multiple media including tables, figures, and text. Of the nine analyses, eight used text (88.9 percent), three used tables (33.3 percent), and six used figures (66.7 percent) to present results.

Table 9. Outcomes and results reporting in Frequentist mixed treatment comparisons

Characteristic	n/N (%) or Mean (SD)
Claim of equivalence	0/9 (0)
Claim of non-inferiority	0/9 (0)
Minimally important difference	0/9 (0)
Type of outcome	
Binary	7/9 (77.8)
Continuous	1/9 (11.1)
Both binary and continuous	1/9 (11.1)
Binary effect measure	
Relative risk	3/8 (37.5)
Odds ratio	4/8 (50.0)
Log odds ratio	1/8 (12.5)
Continuous effect measure	
Weighted mean difference	1/2 (50.0)
Standardized effect size	1/2 (50.0)
Presentation of Results*	
Table	3/9 (33.3)
Text	8/9 (88.9)
Figure	6/9 (66.7)

*Studies can be counted multiple times based on format used to present results

Part Three: MTC Focus Group

Key Points

- Nine individuals participated in our focus group, all of whom were authors of MTCs using Bayesian methods identified in part two of this report. Unfortunately despite all efforts, none of the limited number of investigators who conducted MTC using Frequentist methods replied to our invitation or participated in the group.
- The majority of respondents were from academic settings, have been trained in network meta-analysis methods and have conducted at least two such analyses. The respondents seemed to be involved in a variety of the steps in conducting the identified network meta-analysis.
- Respondents seem to feel the term “network meta-analysis” is used ambiguously and inconsistently in the medical literature, although they do not feel the same about the terms “mixed treatment comparison” or “Frequentist network meta-analysis.”
- Of the questions asking general opinion of network meta-analysis, most responses to questions were on average a neutral response on a 5-point scale. Of the comments which had clear majority opinions were:
 - Disagreement that investigators should consider restricting their search to the minimum number of interventions of interest when conducting a network meta-analysis
 - Agreement that the combination of indirect and direct evidence adds valuable information that is not available from head-to-head comparisons.
 - Agreement that network meta-analysis should provide a graphical depiction of the evidence network.
- When asked specifically about Bayesian methods to conduct MTC, respondents provided a variety of strengths and limitations. Although many were unique, the limitation mentioned most commonly was in regards to the software while there was no commonly mentioned strength.

- When asked specifically about their MTC, most respondents built the code from scratch or adapted the code from a previously published code. Unfortunately we did not gain insight as to how or why prior distributions were chosen but rather what the priors chosen were.
- Additionally respondents were asked to rate 11 criteria on how influential each was in their decision to use Bayesian methods for their MTC. The most influential criteria, on average, were the method’s ability to handle multi-arm studies and collaborator’s or respondent’s prior expertise and/or experience. The least influential criterion was the requirement to specify noninformative priors.

Detailed Analysis

Tables are used throughout this section to present results for each individual focus group question or to present free text responses. Not all data appear in both text and table format and some data are exclusively reported in within either format.

Composition of the Focus Group

The focus group was comprised of nine individuals (hereafter respondents), who authored a unique MTC using Bayesian methods identified in part two of this project. Despite all efforts to contact the authors of the analyses using Frequentist methods, no authors successfully replied or participated in the group. Therefore, the presented results represent the views of investigators who have used Bayesian methods to conduct their MTC. Most respondents work in academic settings (66.7 percent) and consider themselves to have the expertise needed to implement a network meta-analysis themselves (77.8 percent). Most respondents (88.9 percent) have received either formal or informal training in network meta-analysis methods (Table 10).

Table 10. Training of respondents

“I read many published analysis I initiated myself working on Anne Whitehead book I followed Bayesian courses on evidence synthesis”
“Systematic Review & Meta-Analysis of Direct, Indirect and mixed Treatment Evidence, University of Glasgow Indirect and Mixed Treatment Comparisons Course Leicester”
“3-day course on indirect comparison and MTC by Leicester & Bristol university staff”
“Took a course in Bayesian analysis, including meta-analysis”
“Dedicated course (actually after I did my first analyses)”
“2005 Bristol course. In addition involved in development of methods”
“...research fellowship, AHRQ workshop”

Three respondents are affiliated with an organization involved in conducting synthesis, systematic review, or meta-analysis, including AHRQ (n=2) and Cochrane (n=1). The referenced meta-analysis was not the first in which any of the respondents used such methods. All respondents have conducted at least two network meta-analyses and three of the nine respondents (33.3 percent) have conducted five or more of these analyses. When asked to select which activities described their involvement in the given analysis, it appears that the respondents were involved in multiple steps of the process (Table 11).

Table 11. Role of respondents in their meta-analysis

Clinical advice, clinical interpretation, policy development (n=6, 66.7%)
Protocol development (n=8, 88.9%)
Developed search strategy (n=8, 88.9%)
Data extraction (n=7, 77.8%)
Statistical advice/methodology (n=7, 77.8%)
Writing or critical revision of manuscript/report (n=8, 88.9%)
Obtaining funding (n=3, 33.3%)
Other (please specify) (n=0)

General Questions Regarding Network Meta-Analysis

Respondents were asked a series of 14 questions, using a 5-point Likert scale, regarding general principles and views of network meta-analysis. The results for each question are presented in Table 12. In summary, mixed results were obtained when asking the respondents their opinion as to the ambiguity and consistency in which certain terms were used in the literature. Respondents felt that the term “network meta-analysis” is used ambiguously and inconsistently in the medical literature, whereas the term “mixed treatment comparison” was consistently and unambiguously used. Last, most respondents were neutral to how the term “Frequentist network meta-analysis” is used in the literature. All respondents agreed that the combination of indirect and direct evidence adds valuable information that is not available from head-to-head comparisons as well as the necessity for MTCs to provide a graphical depiction of the evidence network. The majority of respondents disagreed that “when conducting a network meta-analysis, an investigator should consider restricting a search to the minimum number of interventions of interest.” All respondents agreed or were neutral with the statement “the combination of direct and indirect evidence yields a more refined and precise estimate of the interventions directly compared” and the statement “the combination of direct and indirect evidence broadens the external validity of the analysis.” The remaining questions had a mixture of responses that did not have a majority representation.

Table 12. General mixed treatment comparison questions

Question	Strongly disagree n (%)	Disagree n (%)	Neutral n (%)	Agree n (%)	Strongly agree n (%)	Mean (SD)
The term “network meta-analysis” is used unambiguously and consistently in the medical literature.	2 (22.2%)	3 (33.3%)	4 (44.4%)	0	0	2.2 (0.8)
The term “mixed treatment comparison” is used unambiguously and consistently in the medical literature.	0	1 (11.1%)	4 (44.4%)	3 (33.3%)	1 (11.1%)	3.4 (0.9)

Table 12. General mixed treatment comparison questions (continued)

Question	Strongly disagree n (%)	Disagree n (%)	Neutral n (%)	Agree n (%)	Strongly agree n (%)	Mean (SD)
The term “frequentist network meta-analysis” is used unambiguously and consistently in the medical literature.	0	1 (11.1%)	6 (66.7%)	2 (22.2%)	0	3.1 (0.6)
Synthesizing direct evidence only form sufficient head-to-head or randomized controlled trials takes precedence over analysis containing indirect evidence.	0	3 (33.3%)	3 (33.3%)	3 (33.3%)	0	3 (0.9)
The combination of indirect and direct evidence adds valuable information that is not available from head-to-head comparisons.	0	0	0	5 (55.6%)	4 (44.4%)	4.4 (0.5)
The combination of indirect and direct evidence yields a more refined and precise estimate of the interventions directly compared.	0	0	3 (33.3%)	6 (66.7%)	0	3.7 (1.1)
The combination of indirect and direct evidence broadens the external validity of the analysis.	0	0	5 (55.6%)	4 (44.4%)	0	3.4 (0.5)
When analysis of both direct and indirect comparisons is undertaken, each approach should be considered and reported separately.	0	2 (22.2%)	1 (11.1%)	4 (44.4%)	2 (22.2%)	3.7 (1.1)
When conducting a network meta-analysis, an investigator should consider restricting a search to the minimum number of interventions of interest	1 (11.1%)	5 (55.6%)	2 (22.2%)	1 (11.1%)	0	2.3 (0.9)
When conducting a network meta-analysis, an investigator should consider including comparisons not of direct interest (e.g. placebo controls and therapies no longer used in practice)	0	3 (33.3%)	2 (22.2%)	2 (22.2%)	2 (22.2%)	3.3 (1.2)
The more interventions that are included in a network meta-analysis, the greater uncertainty is reduced, precision is increased, and the ability to establish whether various sources of evidence agree with each other is enhanced.	0	2 (22.2%)	3 (33.3%)	3 (33.3%)	1 (11.1%)	3.3 (1.0)
Network meta-analysis should provide a graphical depiction of the evidence network.	0	0	0	2 (22.2%)	7 (77.8%)	4.8 (0.4)

Table 12. General mixed treatment comparison questions (continued)

Question	Strongly disagree n (%)	Disagree n (%)	Neutral n (%)	Agree n (%)	Strongly agree n (%)	Mean (SD)
The specific statistical code used should be available either as part of the manuscript, appendix/supplement material, or available on an external Web site for the reader to freely access	0	1 (11.1%)	2 (22.2%)	4 (44.4%)	2 (22.2%)	3.8 (1.0)
Current guidance on how to conduct and report a network meta-analysis is sufficient.	0	2 (22.2%)	2 (22.2%)	5 (55.6%)	0	3.3 (0.9)

Questions Specific to Bayesian Methods for MTC

The respondents were asked a series of open-ended questions. First, they were asked to list the three most significant barriers of Bayesian methods when conducting MTC, results of which are found in Table 13. All respondents listed at least one barrier, seven listed two barriers, and five listed three barriers. Respondents were also asked to list the three most significant strengths of Bayesian methods when conducting MTC, results of which are listed in Table 13. All respondents listed at least one strength, eight listed two strengths, and six listed three strengths.

Table 13. Question specific to the investigator’s published mixed treatment comparisons with Bayesian methods

Listed as the Three Most Significant Barriers to Using Bayesian Methods	Listed as the Three Most Significant Strengths to Using Bayesian Methods
Data quality-heterogeneity	Treatment efficacy ranking
Knowledge of the method	Ability to use both direct and indirect evidence
To construct the code	To analyze comparisons that were not conducted directly
Availability of researchers with the necessary expertise	High quality
Ease of use of the software	Practical factors listed on the previous page
WinBUGS	Allows for indirect comparisons when direct evidence is lacking
I cannot see any (barriers)	Because we do “informal” MTC everyday
Implementing the code	Model estimation
Ability of reader to interpret the method	Multi-arm trial adjustment
Analyze and interpret data	Handling uncertainty
Time	Ability to report ranking of interventions
Lack of user friendliness of WinBUGS	Impact
Proper understanding of the results	Intuitive interpretation
The amount of data to digest	Is growing in acceptance
Adequately reporting on methods in manuscript	Comprehensive picture of the evidence
Availability of evidence to form adequate network	Multi arm trials
Resources	No adjustment for zero cells
Acceptance of the method at time of publishing	Thorough check of the available evidence
How to present results	Novel
	Answers questions that are otherwise unanswerable
	Presentation of results (Bayesian inference in general)
	Survival endpoints

Questions Specific to Respondent’s MTC

Respondents were asked to rate 11 criteria based on how influential each criterion was in their decision to conduct a MTC using Bayesian methods. A 5-point scale was used ranging from “not at all” to “extremely.” The responses to each question can be found in Table 15. On average, the criteria with the most influence were the method’s ability to handle multi-arm trials and the collaborator’s or respondent’s prior experience and/or expertise. The next most influential criterion was the amount of methodological research supporting this method followed by the method’s ability to allow rank ordering of interventions according to the probability they are best. The remaining criteria were less influential in the respondents’ decision-making to use Bayesian methods (Table 14).

Table 14. Question specific to the investigator’s published mixed treatment comparison and how much the specific parameters influenced their decision to use Bayesian methods

Question	Not At All n (%)	A Little n (%)	Moderately n (%)	Quite a Bit n (%)	Extremely n (%)	Mean (SD)
The method allows for the ranking of interventions according to the probability they are best.	1 (11.1%)	2 (22.2%)	3 (33.3%)	2 (22.2%)	1 (11.1%)	3 (1.2)
The method allows investigators to check and compare the fit of a model	1 (11.1%)	3 (33.3%)	4 (44.4%)	0	1 (11.1%)	2.7 (1.1)
The method’s ability to handle multi-arm studies (those with more than 2 treatment groups)	0	1 (11.1%)	4 (44.4%)	3 (33.3%)	1 (11.1%)	3.4 (0.9)
Frequency of use in previously published network meta-analyses	3 (33.3%)	3 (33.3%)	1 (11.1%)	2 (22.2%)	0	2.2 (1.2)
Ease of software implementation	3 (33.3%)	2 (22.2%)	2 (22.2%)	1 (11.1%)	1 (11.1%)	2.4 (1.4)
The amount of methodological research supporting this method	2 (22.2%)	1 (11.1%)	1 (11.1%)	4 (44.4%)	1 (11.1%)	2.7 (1.4)
The methods ability to combine trials reporting results in different formats, for example binomial data and summary log odds with variance (multi- or shared parameter models)	4 (44.4%)	0	3 (33.3%)	0	2 (22.2%)	2.6 (1.7)
Access to pre-built models	3 (33.3%)	1 (11.1%)	1 (11.1%)	4 (44.4%)	0	2.7 (1.4)
Requirement to specify priors which are often arbitrary	3 (33.3%)	4 (44.4%)	2 (22.2%)	0	0	1.9 (0.8)
Collaborator(s) or your prior experiences and/or expertise	1 (11.1%)	1 (11.1%)	2 (22.2%)	3 (33.3%)	2 (22.2%)	3.4 (1.3)
The method’s ability to handle studies with “zero cells”	3 (33.3%)	1 (11.1%)	1 (11.1%)	2 (22.2%)	2 (22.2%)	2.9 (1.7)

In response to a true/false question, five of nine respondents involved a researcher /collaborator solely due to their methodological expertise in Bayesian methods. Eight of the nine respondents did not use formal guidance to guide how the MTC was conducted. One respondent replied that there was no guidance available at the time of their analysis. Respondents were asked how the code used in the analysis was derived. Three codes were adapted from a previously published code, three codes were built from scratch, one code was built from scratch with the help of WinBUGS examples, one code was adapted from a publically available code, and the last

instance the respondent was unsure how the code was derived (Table 15). The last open-ended question asked how prior distributions were chosen for the meta-analysis and why they were chosen over others. Unfortunately, the responses collected do not seem to provide insight as to how or why, but rather what the prior distributions were (Table 16).

Table 15. Information about how the code was derived

How Was the Code Used in Your Analysis Derived (e.g. built from scratch, used/adapted previously published/publically available code or wrapper [e.g. BUGSXL or other] to generate code, or other source)

Built from scratch plus help from WinBUGS examples (n=1)

Built from scratch (n=3)

Don't know (n=1)

Adapted from previously published code (n=3)

Adapted from publically available code (n=1)

Table 16. Information about how and why the prior distributions were chosen

How Were Your Prior Distributions Chosen and Why Were These Distributions Chosen Over Others? (8 replies)

Only noninformative distribution used

I don't know

Noninformative

I don't know

Used Gaussian noninformative distribution priors for treatment effects as recommended. Model didn't converge well with flat prior for between study SD, so used empirical informative half-normal prior

We chose flat, noninformative priors

We tried three different priors

Noninformative and sensitivity analyses on heterogeneity priors default in code

Discussion

This report provides the results of a three-part methods project that aimed to first review existing guidance on network meta-analysis, secondly to identify previously published MTCs and summarize their characteristics, and finally, to gather insight from investigators who have conducted network meta-analyses using these methods.

Our review of publicly available guidance documents from various governmental and evidence synthesis groups found that the majority of these documents were typically written in a fashion applicable to network meta-analysis in general, and not specific to any one methodology type. In regards to methods used to conduct meta-analyses of networks of trials containing at least one closed loop, the two approaches typically discussed by guidance included the Bayesian and the Frequentist mixed methods approach initially described by Lumley. Guidance documents stressed that both these approaches have decreased internal validity because they compromise the positive impact of individual study randomization. Common limitations of the Lumley's Frequentist approach discussed by guidance documents included the approaches' inability to synthesize networks of studies lacking at least one closed loop, the fact that the method does not account for correlations that may exist between effect estimates when they are obtained from a single multi-arm study, and a weaknesses in situations where zero cells are common. These limitations can be addressed through special preparations such as using a small increment to address zero cells and adding steps to adjust for correlations between effect estimates. The Bayesian approach was often criticized for requiring specification of noninformative priors, its complexity to understand, and the need to use non-user-friendly software to implement. Guidance noted some similarities between the methods as well. Regardless of the approach discussed, guidance documents stressed the need for consistency of factors and event rates between direct and indirect trials and the importance of assessing for consistency/inconsistency and heterogeneity. The International Society of Pharmacoeconomics and Outcomes Researchers was the only group that attempted to comprehensively address how to conduct, interpret and report a network meta-analysis. Additional guidance on how to conduct, interpret and report a network meta-analysis is needed.

Our systematic review identified 42 unique MTCs that used either Bayesian or Frequentist methods. These MTC were published in 32 different journals, most of which with accompanying supplements. Of the 42 MTCs, the vast majority used Bayesian methods. Investigators could have chosen either Bayesian or Frequentist methods as both can accommodate close loop models. Despite the option, most investigators chose a Bayesian approach. Of the analyses that utilized Bayesian approach, there was a wide distribution of disease states evaluated although cardiology was the most common area. Most analyses evaluated pharmacologic interventions and were funded by industry or a government/foundation source. There was a large variance in printed pages number of the manuscript, although two included MTCs were affiliated reports without page limitation and were likely the contributing factor. The statistical code used in the analysis was rarely made available to the reader, despite the majority of journals allowing publication of a supplement or appendix, although raw outcomes data were more commonly published. A similar number of analyses used vague priors or did not specify whether priors were intended to be vague and few analyses used informative priors. However, it was uncommon for authors to report specific priors used. Most models used a random effects model. Unfortunately, data regarding the evaluation of convergence, heterogeneity, and inconsistency were inconsistently reported and often times not mentioned throughout the publication. From the analyses that reported evaluating these three characteristics, it appears that a broad range of

methods are being utilized. We cannot say with certainty though that a lack of reporting means these characteristics were not evaluated. Perhaps with more clear guidance in the future, as to how to conduct and report these types of network meta-analyses, a more consistent approach may be taken. When investigators reported results of their findings, it was common that interventions were rank ordered based on the probability of the intervention being best for a given outcome. Rarely did authors conclude equivalence or non-inferiority of interventions based on network meta-analysis results. The most common types of outcomes evaluated were binary outcomes, measured with relative risks or odds ratios and 95% credible intervals.

As there were very few network meta-analyses identified by our systematic review that used Frequentist methods, summarizing similarities and differences across the analyses is difficult. Only nine analyses used these Frequentist type methods, despite the option of doing so amongst the majority of Bayesian MTCs. Unfortunately, we did not gain any insight as to the decisionmaking and opinions of the investigators of the Frequentist models because of a lack of response to our focus group invitation. All of the respondents had conducted a Bayesian MTC and therefore we could not compare and contrast the viewpoints between investigators who used Bayesian methods versus Frequentist methods.

The group of respondents did not appear to be new to Bayesian MTC methods as all had conducted at least two such analysis and appeared to be involved in a variety of steps in the process. However, it is unlikely the respondents were methodologists since most did not know how the code or prior distributions were chosen. Although we prefaced the questions with a list of terms and definitions for the respondents to use while answering the questions, we assume the respondents did in fact apply those definitions. Another potential limitation to this portion of the project is the one time correspondence with the investigator to obtain opinion. The process was not interactive and therefore a general consensus was not achieved in areas of discrepancy.

On average, the group felt the term “network meta-analysis” is used ambiguously and inconsistently in the medical literature, although did not feel the same about the terms “mixed treatment comparison” or “frequentist network meta-analysis.” In general, there were neutral opinions on average regarding network meta-analyses principles. However, clear majority was seen for the following: disagreement that investigators should consider restricting their search to the minimum number of interventions of interest when conducting a network meta-analysis; agreement that the combination of indirect and direct evidence adds valuable information that is not available from head-to-head comparisons; and agreement that network meta-analysis should provide a graphical depiction of the evidence network. The respondents identified several strengths and limitations of Bayesian MTC. Although most were unique statements, there was a common limitation suggested regarding the user friendliness of software used to run the analyses.

Respondents were asked specifically about their Bayesian MTC which we had identified in part two of this project. The most influential criteria in deciding to use Bayesian MTC, on average, were the method’s ability to handle multi-arm studies and collaborator’s or respondent’s prior expertise and/or experience. The least influential criterion was the requirement to specify priors which are often arbitrary. Most respondents built the code from scratch or adapted the code from a previously published code. Unfortunately we did not gain insight as to how or why prior distributions were chosen rather what the priors chosen were.

Overall, further research is needed to build on this report and develop a set of practical guidelines for conducting MTCs, developed by all relevant stakeholders, including representatives from academia and industry. Such guidelines may also lead to standardized

approaches to reporting MTCs. Future efforts should be made to continue to understand the rationale of investigators in their choice of Bayesian versus Frequentist methods to conduct MTCs.

References

1. Bucher HC, Guyatt GH, Griffith LE, et al. The results of direct and indirect treatment comparisons in meta-analysis of randomized controlled trials. *J Clin Epidemiol.* 1997;50:683-91. PMID:9250266
2. Lumley T. Network meta-analysis for indirect treatment comparisons. *Stats Med* 2002;21:2313-2324. PMID:12210616
3. Fu R, Gartlehner G, Grant M, et al. Conducting Quantitative Synthesis When Comparing Medical Interventions: AHRQ and the Effective Health Care Program. In: Agency for Healthcare Research and Quality. *Methods Guide for Comparative Effectiveness Reviews* [posted October 2010]. Rockville, MD. <http://effectivehealthcare.ahrq.gov/> (Accessed December 28, 2011).
4. Sutton A, Ades AE, Cooper N et al. Use of indirect and mixed treatment comparisons for technology assessments. *Pharmacoeconomics* 2008;26:753-767. PMID:18767896
5. Lu G, Ades AE. Combination of direct and indirect evidence in mixed treatment comparisons. *Stats Med* 2004;23:3105-3124. PMID:15449338
6. Caldwell DM, Ades AE, Higgins PT. Simultaneous comparison of multiple treatments: combining direct and indirect evidence. *BMJ* 2005;331:897-900. PMID:16223826
7. Health Information and Quality Authority. *Guidelines for Evaluating the Clinical Effectiveness of Health Technologies in Ireland.* Dublin: Health Information and Quality Authority; 2011. Available at: <http://www.hiqa.ie> (Last accessed on December 28, 2011)
8. Glenny AM, Altman DG, Song F, et al. Indirect comparisons of competing interventions. *Health Technol Assess* 2005;9(26). PMID: 16014203
9. Jansen JP, Fleurence R, Devine B, et al. Interpreting indirect treatment comparisons and network meta-analysis for health-care decision making: report of the ISPOR Task Force on Indirect Treatment Comparisons Good Research Practices: part 1. *Value Health.* 2011;14:417-28. PMID:21669366
10. Hoaglin DC, Hawkins N, Jansen JP, et al. Conducting indirect-treatment-comparison and network-meta-analysis studies: report of the ISPOR Task Force on Indirect Treatment Comparisons Good Research Practices: part 2. *Value Health.* 2011;14:429-37. PMID:21669367
11. Wells GA, Sultan SA, Chen L, et al. *Indirect Evidence: Indirect Treatment Comparisons in Meta-Analysis.* Ottawa: Canadian Agency for Drugs and Technologies in Health; 2009. www.cadth.ca (Accessed December 28, 2011).
12. Centre for Reviews and Dissemination. *CRD's Guidance for Undertaking Reviews in Health Care.* University of York, 2009. www.york.ac.uk/inst/crd/pdf/Systematic_Review_s.pdf (Accessed December 28, 2011).
13. Sung L, Hayden J, Greenberg ML, Koren G, Feldman BM, Tomlinson GA. Seven items were identified for inclusion when reporting a Bayesian analysis of a clinical study. *J Clin Epidemiol* 2005;58:261-268.
14. Drug Effectiveness Review Project. *Systematic Review Methods and Procedures.* Portland, OR, 2011. www.ohsu.edu/xd/research/centers-institutes/evidence-based-policy-center/derp/documents/upload/DERP_METHOD_S_WEB_Final_January-2011-2.pdf (Accessed December 28, 2011).
15. *Guidelines for the economic evaluation of health technologies: Canada [3rd Edition].* Ottawa: Canadian Agency for Drugs and Technologies in Health; 2006. www.cadth.ca (Last accessed December 28, 2011).

16. Li T, Puhan MA, Vedula SS, et al. Ad Hoc Network Meta-analysis Methods Meeting Working Group. Network meta-analysis-highly attractive but more methodological research is needed. *BMC Med.* 2011;9:79. PMID:21707969
17. Pharmaceutical Benefits Advisory Committee. Guidelines for Preparing Submissions to the Pharmaceutical Benefits Advisory Committee (version 4.3); Australian Government, Department of health and Ageing. 2008. [www.health.gov.au/internet/main/publishing.nsf/content/AECB791C29482920CA25724400188EDB/\\$File/PBAC4.3.2.pdf](http://www.health.gov.au/internet/main/publishing.nsf/content/AECB791C29482920CA25724400188EDB/$File/PBAC4.3.2.pdf) (Accessed December 28, 2011).
18. Report of the Indirect Comparisons Working Group (ICWG) to the Pharmaceutical Benefits Advisory Committee: Assessing Indirect Comparisons. Australian Government, Department of health and Ageing. Available at: [www.health.gov.au/internet/main/publishing.nsf/Content/B11E8EF19B358E39CA25754B000A9C07/\\$File/ICWG%20Report%20FINAL2.pdf](http://www.health.gov.au/internet/main/publishing.nsf/Content/B11E8EF19B358E39CA25754B000A9C07/$File/ICWG%20Report%20FINAL2.pdf) (Accessed December 28, 2011).
19. National Institute for Health and Clinical Excellence. Guide to the methods of technology appraisal, 2008. Available at: www.nice.org.uk (Accessed December 28, 2011).
20. National Institute for Health and Clinical Excellence (NICE). NICE Decision Support Unit Technical Support Documents for "Evidence Synthesis". www.nicedsu.org.uk/Evidence-Synthesis-TSD-series%282391675%29.htm Accessed June 10, 2012.
21. Cleemput I, Van Wilder P, Vrijens F, Huybrechts M, Ramaekers D. Guidelines for Pharmacoeconomic Evaluations in Belgium. Health Technology Assessment (HTA). Brussels: Health Care Knowledge Centre (KCE); 2008. KCE Reports 78C (D/2008/10.273/27). www.ispor.org/peguidelines/source/Belgium_Guidelines-for-Pharmacoeconomics-Evaluation-in-Belgium_2008_English.pdf Accessed June 10, 2012.
22. German Institute for Quality and Efficiency in Health Care. Benefit Assessment of Pharmaceuticals pursuant to § 35a SGB V, 2008. Available at: www.english.g-ba.de/benefitassessment/information/ Accessed June 10, 2012.
23. Haute Autorite de Sante. Indirect comparisons Methods and validity, 2009. Available at: www.has-sante.fr/portail/upload/docs/application/pdf/2011-02/summary_report__indirect_comparisons_methods_and_validity_january_2011_2.pdf Accessed June 10, 2012
24. National Department of Health – Republic of South Africa. The Guidelines For Pharmacoeconomic Evaluations of Medicines and Scheduled Substances, 2010. Available at: www.ispor.org/PEguidelines/source/guidelines-for-pharmacoeconomic-evaluations-of-medicines-and-scheduled-substances_2010.pdf
25. Higgins JPT, Green S (eds). *Cochrane Handbook for Systematic Reviews of Interventions* Version 5.1.0 [updated March 2011]. The Cochrane Collaboration, 2011. www.cochrane-handbook.org (Accessed December 28, 2011).
26. Evolution of Cochrane Intervention Reviews and Overviews of Reviews to better accommodate comparisons among multiple interventions. Report from a meeting of the Cochrane Comparing Multiple Interventions Methods Groups, Milan, March 2011. <http://cmimg.cochrane.org/sites/cmimg.cochrane.org/files/uploads/CMIMG%20summary%20of%20meeting%20Milan%20March%202011.pdf> (Accessed December 28, 2011).
27. Salanti G, Higgins JP, Ades AE, Ioannidis JP. Evaluation of networks of randomized trials. *Stat Methods Med Res* 2008;17:279-301
28. Viechtbauer W. Conducting meta-analyses in R with the metaphor package. *J Stat Software* 2010;36:1-48.

29. Hawkins N, Scott DA, Woods B. How far do you go? Efficient searching for indirect evidence. *Med Decis Making*. 2009;29(3):273-81. PMID: 19470721
30. The BaSiS Group. Bayesian standards in science (BaSiS). <http://lib.stat.cmu.edu/bayesworkshop/2001/BaSiS.html> (Accessed September 21, 2011).
31. Spiegelhalter DJ, Myles JP, Jones DR, Abrams KR. Bayesian methods in health technology assessment: a review. *Health Technol Assess* 2000;4:1-130.
32. Preferred reporting items for systematic reviews and meta-analyses. www.prisma-statement.org/statement.htm (Accessed April 4, 2012)
33. Bangalore S, Kumar S, Kjeldsen E et al. Antihypertensive drugs and risk of cancer: network meta-analyses and trial sequential analysis of 324168 participants from randomized trials. *Lancet Oncol* 2011;12:65-82. PMID:21123111
34. Gross JL, Kramer CK, Leitaó CB, et al. Effect of antihyperglycemic agents added to metformin and a sulfonylurea on glycemic control and weight gain in type 2 diabetes: a network meta-analysis. *Ann Intern Med* 2011;154(10):672-679. PMID:21576535
35. Hartling L, Fernandes RM, Bialy L, et al Steroids and bronchodilators for acute bronchiolitis in the first two years of life: systematic review and meta-analysis. *BMJ* 2011;342:1714. PMID:21471175
36. Maund E, McDaid C, Rice S, Wright K, Jenkins B, Woolacott N. Paracetamol and selective and non-selective non-steroidal anti-inflammatory drugs for the reduction in morphine-related side-effects after major surgery: a systematic review. *Br J Anaesth* 2011;106(3):292-297. PMID:21285082
37. McDaid C, Maund E, Rice S, Wright K, Jenkins B, Woolacott N. Paracetamol and selective and non-selective non-steroidal anti-inflammatory drugs (NSAIDs) for the reduction of morphine-related side effects after major surgery: a systematic review. *Health Technol Assess* 2010;14(17):1-153. PMID:20346263
38. Sciarretta S, Palano F, Tocci G, Baldini R, Volpe M. Antihypertensive treatment and development of heart failure in hypertension: a Bayesian network meta-analysis of studies in patients with hypertension and high cardiovascular risk. *Arch Intern Med* 2011;171(5):384-394. PMID:21059964
39. Trelle S, Reichenback S, Wandel S, et al. Cardiovascular safety of non-steroidal anti-inflammatory drugs: network meta-analysis. *BMJ* 2011;342:c7086. PMID:21224324
40. van de Kerkhof P, de Peuter R, Rytto J, Jansen JP. Mixed treatment comparison of a two-compound formulation (TCF) product containing calcipotriol and betamethasone dipropionate with other topical treatments in psoriasis vulgaris. *Curr Med Res Opin* 2011;27(1):225-238. PMID:21142833
41. Van den Bruel A, Gailly J, Devriese S, Welton NJ, Shortt AJ, Vrijens F. The protective effect of ophthalmic viscoelastic devices on endothelial cell loss during cataract surgery: a meta-analysis using mixed treatment comparisons. *Br J Ophthalmol* 2011;95(1):5-10. PMID:19965836
42. Dakin H, Fidler C, Harper C. Mixed treatment comparison meta-analysis evaluating the relative efficacy of nucleos(t)ides for treatment of nucleos(t)ide-naive patients with chronic hepatitis B. *Value Health* 2010;13(8):934-945. PMID:20825624
43. Orme M, Collins S, Dakin H, Kelly S, Loftus J. Mixed treatment comparison and meta-regression of the efficacy and safety of prostaglandin analogues and comparators for primary open-angle glaucoma and ocular hypertension. *Curr Med Res Opin* 2010;26(3):511-528. PMID:20014995

44. Phung OJ, Scholle JM, Talwar M, Coleman CI. Effect of noninsulin antidiabetic drugs added to metformin therapy on glycemic control, weight gain, and hypoglycemia in type 2 diabetes. *JAMA* 2010;303(14):1410-1418. PMID:20388897
45. Uthman OA, Abdulmalik J. Comparative efficacy and acceptability of pharmacotherapeutic agents for anxiety disorders in children and adolescents: a mixed treatment comparison meta-analysis. *Curr Med Res Opin* 2010;26(1):53-59. PMID:19905879
46. Vissers D, Stam W, Nolte T, Lenre M, Jansen J. Efficacy of intranasal fentanyl spray versus other opioids for breakthrough pain in cancer. *Curr Med Res Opin* 2010;26(5):1037-1045. PMID:20199140
47. Walsh T, Worthington HV, Glenny A, Appelbe P, Marinho CCV, Shi X. Fluoride toothpastes of different concentrations for preventing dental caries in children and adolescents. *Cochrane Database of Systematic Reviews* 2011;1 Art No CD007868. PMID:20091655
48. Wandel S, Juni P, Tendal B, Nuesch E, Villiger PM, Welton NJ, et al. Effects of glucosamine, chondroitin, or placebo in patients with osteoarthritis of hip or knee: network meta-analysis. *BMJ* 2010;341:4675. PMID:20847017
49. Wang H, Huang T, Jing J, Jin J, Wang P, Yang M, et al. Effectiveness of different central venous catheters for catheter-related infections: a network meta-analysis. *J Hosp Infect* 2010;76(1):1-11. PMID:20638155
50. Woo G, Tomlinson G, Nishikawa Y, Kowgier M, Sherman M, Wong DK, et al. Tenofovir and entecavir are the most effective antiviral agents for chronic hepatitis B: a systematic review and Bayesian meta-analyses. *Gastroenterology* 2010;139(4):1218-1229. PMID:20600036
51. Baker WL, Baker EL, Coleman CI. Pharmacologic treatments for chronic obstructive pulmonary disease: a mixed-treatment comparison meta-analysis. *Pharmacotherapy* 2009;29(8):891-905. PMID:19637942
52. Bansback N, Sizto S, Sun H, Feldman S, Willian MK, Anis A. Efficacy of systemic treatments for moderate to severe plaque psoriasis: systematic review and meta-analysis. *Dermatology* 2009;219(3):209-218. PMID:19657180
53. Cipriani A, Furukawa TA, Salanti G, Geddes JR, Higgins JP, Churchill R, et al. Comparative efficacy and acceptability of 12 new-generation antidepressants: a multiple-treatments meta-analysis. *Lancet* 2009;373(9665):746-758. PMID:19185342
54. Edwards SJ, Lind T, Lundell L, Das R. Systematic review: standard- and double-dose proton pump inhibitors for the healing of severe erosive oesophagitis – a mixed treatment comparison of randomized controlled trials. *Aliment Pharmacol Ther* 2009;30(6):547-556. PMID:19558609
55. Edwards SJ, Smith CJ. Tolerability of atypical antipsychotics in the treatment of adults with schizophrenia or bipolar disorder: a mixed treatment comparison of randomized controlled trials. *Clin Ther* 2009;31(Pt 1):1345-1359. PMID:19698898
56. Golfopoulos V, Pentheroudakis G, Salanti G, Nearchou AD, Ioannidis JP, Pavlidis N. Comparative survival with diverse chemotherapy regimens for cancer of unknown primary site: multiple-treatments meta-analysis. *Cancer Treat Rev* 2009;35(7):570-573. PMID:19539430
57. Manzoli L, Georgia S, De Vito C, Boccia A, Ionnidis JPA, Villari P. Immunogenicity and adverse events of avian influenza A H5N1 vaccine in healthy adults: multiple-treatments meta-analysis. *Lancet Infect Dis* 2009;9:481-92. PMID:19628173

58. Meader N. A comparison of methadone, buprenorphine and alpha2 adrenergic agonists for opioid detoxification: a mixed treatment comparison meta-analysis. *Drug and Alcohol Dependence* 2010;108:110-114. PMID:20074867
59. Coleman CI, Baker WL, Kluger J, White CM. Antihypertensive medication and their impact on cancer incidence: a mixed treatment comparison meta-analysis of randomized controlled trials. *J Hypertens* 2008;26(4):622-629. PMID:18327066
60. Mauri D, Polyzos NP, Salanti G, Pavlidis N, Ioannidis JPA. Multiple-treatments meta-analysis of chemotherapy and targeted therapies in advanced breast cancer. *J Natl Cancer Inst* 2008;100:1780-1791. PMID:19066278
61. Stettler C, Allemann S, Wandel S, et al. Drug eluting and bare metal stents in people with and without diabetes: collaborative network meta-analysis. *BMJ* 2008;337:a1331. PMID:18757996
62. Golfinopoulos V, Salanti G, Pavlidis N, Ioannidis JP. Survival and disease-progression benefits with treatment regimens for advanced colorectal cancer: a meta-analysis. *Lancet Oncol* 2007;8(10):898-911. PMID:17888735
63. Lam SK, Owen A. Combined resynchronisation and implantable defibrillator therapy in left ventricular dysfunction: Bayesian network meta-analysis of randomised controlled trials. *BMJ* 2007;335(7626):925. PMID:17932160
64. Nixon R, Bansback N, Brennan A. The efficacy of inhibiting tumour necrosis factor alpha and interleukin 1 in patients with rheumatoid arthritis: a meta-analysis and adjusted indirect comparisons. *Rheumatology* 2007;46(7):1140-1147. PMID:17478472
65. Cooper NJ, Sutton AJ, Lu G, Khunti K. Mixed comparison of stroke prevention treatments in individuals with nonrheumatic atrial fibrillation. *Arch Intern Med* 2006;166(12):1269-1275. PMID:16801509
66. Kyrgiou M, Salanti G, Pavlidis N, Paraskeva E, Ioannidis JP. Survival benefits with diverse chemotherapy regimens for ovarian cancer: meta-analysis of multiple treatments. *J Natl Cancer Inst* 2006;98(22):1655-1663. PMID:17105988
67. Baldwin D, Woods R, Lawson R, Taylor D. Efficacy of drug treatments for generalized anxiety disorder: systematic review and meta-analysis. *BMJ* 2011;342:1199. PMID:21398351
68. Freemantle N, Lafuente-Lafuente C, Mitchell S, Eckert L, Reynolds M. Mixed treatment comparison of dronedarone, amiodarone, sotalol, flecainide, and propafenone, for the management of atrial fibrillation. *Europace* 2011;13(3):329-345. PMID:21227948
69. Anothaisintawee T, Attia J, Nickel JC, et al. Management of chronic prostatitis/chronic pelvic pain syndrome: a systematic review and network meta-analysis. *JAMA* 2011;305(1):78-86. PMID:21205969
70. Hansen RA, Gaynes BN, Gartlehner G, Moore CG, Tiwari R, Lohr KN. Efficacy and tolerability of second-generation antidepressants in social anxiety disorder. *Int Clin Psychopharmacol* 2008;23(3):170-179. PMID:18408531
71. Jalota L, Kalira V, George E, et al. Prevention of pain on injection of propofol: systematic review and meta-analysis. *BMJ* 2011;342:1110. PMID:21406529
72. Trikalinos TA, Alsheikh-Ali AA, Tatsioni A, Nallamothu BK, Kent DM. Percutaneous coronary interventions for non-acute coronary artery disease: a quantitative 20-year synopsis and a network meta-analysis. *Lancet* 2009;373(9667):911-918. PMID:19286090
73. Elliott WJ, Meyer PM. Incident diabetes in clinical trials of antihypertensive drugs: a network meta-analysis. *Lancet* 2007;369(9557):201-207. PMID:17240286

74. Singh JA, Wells GA, Christensen R et al.
Adverse effects of biologics: a network meta-analysis and Cochrane overview (review).
Cochrane database of systematic reviews 2011, Issue 2. Art No. CD008794.
75. Roskell NS, Lip GYH, Noack H, Clemens A, Plumb JM. Treatments for stroke prevention in atrial fibrillation: a network meta-analysis and indirect comparison versus dabigatran etexilate. *Thromb Haemost* 2010;104:1106-1115. PMID: 20967400
76. Eckert L, Lancon C. Duloextine compared with fluoxetine and venlafaxine: use of meta-regression analysis for indirect comparison. *BMC Psychiatry* 2006;6:30. PMID: 16867188

Abbreviations

AHRQ	Agency for Healthcare Research and Quality
BaSiS	Bayesian Standards in Science
CADTH	Canadian Agency for Drugs and Technologies in Health
CRD	Center for Reviews and Dissemination
DERP	Drug Effectiveness Review Program
DIC	deviance information criterion
DSU	Decision Support Unit
EHC	Effective Health Care
EPC	Evidence-based Practice Center
HIQA	Health Information and Quality Authority
ICWG	Indirect Comparisons Working Group
INAHTA	International Network of Agencies for Health Technology Assessment
ISPOR	International Society of Pharmacoeconomics and Health Outcomes Researchers
MTC	mixed treatment comparison
MTM	mixed treatment meta-analysis
NICE	National Institute for Health and Clinical Excellence
OHSU	Oregon Health & Science University
RCT	randomized-controlled trial
ROBUST	Reporting of Bayes Used in Clinical Studies
SD	standard deviation

Appendix A. Verbatim Quotes From Guidance Documents

This appendix contains verbatim quotations from the source documents that were reviewed. These quotations were selected for the degree of relevance to EPCs performing evidence synthesis using network meta-analysis methods. The following are not intended to be an exhaustive representation of the content of the source documents.

When to Conduct Network Meta-Analyses

Definitions/Terminology

“Terminology for indirect treatment comparisons, mixed treatment comparisons, and network meta-analysis varies in the literature.”^{2,3}

“The results of direct comparisons can be combined with those of indirect comparisons by using a mixed approach, known as a mixed treatment comparison (MTC).”¹⁸

“Methods are available for analysing, simultaneously, three or more different interventions in one meta-analysis. These are usually referred to as ‘multiple-treatments meta-analysis’ (‘MTM’), ‘network meta-analysis’, or ‘mixed treatment comparisons’ (‘MTC’) meta-analysis.”¹⁶

“Several [[proposed global estimation methods](#)] have been proposed in the literature to represent networks. Several statistical methods for estimating the parameters for these models (particularly those characterising treatment comparisons) have also been proposed. This diversity has resulted in the following classification:

- Estimation using Bayesian methods: Bayesian network meta analysis
 - Lu and Ades model
 - Model of Caldwell *et al.*
- Estimation using a mixed linear model
 - Lumley's network meta-analysis”¹⁸

“Also called mixed treatments comparison or multiple treatments comparison meta-analysis, network metaanalysis expands the scope of a conventional pair-wise meta-analysis by analyzing simultaneously both direct comparisons of interventions within randomized controlled trials (RCTs) and indirect comparisons across trials based on a common comparator (e.g., placebo or some standard treatment).”⁹

“Indirect and mixed treatment comparisons (MTC), also known as network meta-analysis, represent a recent development in evidence synthesis, particularly in decisionmaking contexts. Rather than pooling information on trials comparing treatments A and B, network meta-analysis combines data from randomised comparisons, A vs. B, A vs. C, A vs. D, B vs. D, and so on, to deliver an internally consistent set of estimates while respecting the randomisation in the evidence.”¹³

“Network analysis will be used to describe *a single synthesised analysis in which more than one common reference is used to indirectly compare the proposed drug and its main comparator.*”¹¹

“Multiple-treatments meta-analysis (MTM) is an extension to indirect comparisons that allows the combination of direct with indirect comparisons, and also the simultaneous analysis of the comparative effects of many interventions”¹⁶

“Mixed treatment comparisons, a special case of network meta-analysis, combine direct evidence and indirect evidence for particular pairwise comparisons, thereby synthesizing a greater share of the available evidence than traditional meta-analysis.”^{2,3}

“Mixed treatment comparisons (MTC), or network meta-analyses, are used to analyse studies with multiple intervention groups and to synthesise evidence across a series of studies in which different interventions were compared...They build a network of evidence that includes both direct evidence from head to head studies and indirect comparisons whereby interventions that have not been compared directly are linked through common comparators.”⁶

According to the HIQA, a multiple treatment comparison combines direct and indirect evidence to compare a technology to two or more other treatments; a network meta-analysis is appropriate for analysing a combination of direct and indirect evidence where there is at least one closed loop of evidence connecting the two technologies of interest, and a Bayesian mixed treatment comparison is appropriate for comparing multiple treatments using both direct and indirect evidence.”¹

“Often only one direct comparison trial is available. Quite often this trial has been designed with a lack of power. In other cases, the comparator may have been used in ways which are debatable. In such a situation a mixed approach, called a mixed treatment comparison 11, in which the results of direct comparisons are compared with those of indirect comparisons, is very useful as it removes or confirms any reservations that one might have about direct comparison trials.”¹⁸

Justification

“In many clinical fields, competing treatments are assessed against placebo and direct comparisons are rare. Indirect comparisons can make it possible to estimate the relative efficacy and/or safety of therapies in relation to each other before any direct comparison trials are available.”¹⁸

“In the absence of randomized, controlled trials involving a direct comparison of all treatments of interest, indirect treatment comparisons and network meta-analysis provide useful evidence for judiciously selecting the best choice(s) of treatment.”^{2,3}

“In the absence of sufficient direct head-to-head evidence and presence of sufficient indirect evidence, indirect comparisons can be considered as an additional analytic tool.”⁴

“Direct comparisons are preferred over indirect comparisons; similarly, effectiveness and long-term or serious adverse event outcomes are preferred to efficacy and short-term tolerability outcomes.”⁵

In some cases, the choice of the comparator will be difficult due to, for instance, changes in prescription behaviour and therapeutic insights over time. The comparator defined at the time of the clinical trials may no longer be the relevant comparator at the time of the pharmacoeconomic evaluation. In this case, indirect comparisons and/or modelling may be required.¹⁹

Indirect comparisons are second best solutions and are only accepted if no single trial of appropriate quality or relevance to the Belgian target population has been performed and under specific conditions regarding the analyses.¹⁹

If there are no clinical studies for a direct comparison with the pharmaceutical being assessed, or if these do not provide sufficient information about the additional benefit, indirect comparisons can be made in the dossier.²⁰

“Where relevant direct randomised trials (as defined in Part II, Subsection B.2) comparing the proposed drug directly with the main comparator are available, their analysis and presentation are preferred as the basis of the clinical evaluation (see Part II, Section B). However, in the absence of any such direct randomised trials, the second step in the hierarchy is to determine whether it is possible to present an indirect comparison based on two or more sets of randomised trials involving one or more common reference.”¹⁰

”In situations when both direct and indirect comparisons are available in a review, then unless there are design flaws in the head-to-head trials, the two approaches should be considered separately and the direct comparisons should take precedence as a basis for forming conclusions.”¹⁶

“Although it is often argued that indirect comparisons are needed when direct comparisons are not available, it is important to realize that both direct and indirect evidence contributes to the total body of evidence. The results from indirect evidence combined with the direct evidence may strengthen the assessment between treatments directly evaluated. Even when the results of the direct evidence are conclusive, combining them with the results of indirect estimates in a mixed treatment comparison (MTC) may yield a more refined and precise estimate of the interventions directly compared and broaden inference to the population sampled because it links and maximizes existing information within the network of treatment comparisons”^{2,3}

“Data from head-to-head RCTs should be presented in the reference-case analysis, if available. When head-to-head RCTs exist, evidence from mixed treatment comparison analyses may be presented if it is considered to add information that is not available from the head-to-head comparison. This mixed treatment comparison must be fully described and presented as additional to the reference-case analysis (a ‘mixed treatment comparison’ includes trials that compare the interventions head-to-head and indirectly). When multiple technologies are being appraised that have not been compared within a single RCT, data from a series of pairwise head-to-head RCTs should be presented. Consideration should also be given to presenting a combined analysis using a mixed treatment comparison framework if it is considered to add information that is not available from the head-to-head comparison. If data from head-to-head RCTs are not available, indirect treatment comparison methods should be used (an ‘indirect comparison’ is a

synthesis of data from a network of trials). The principles of good practice for standard meta-analyses should also be followed in mixed and indirect treatment comparisons.”¹²

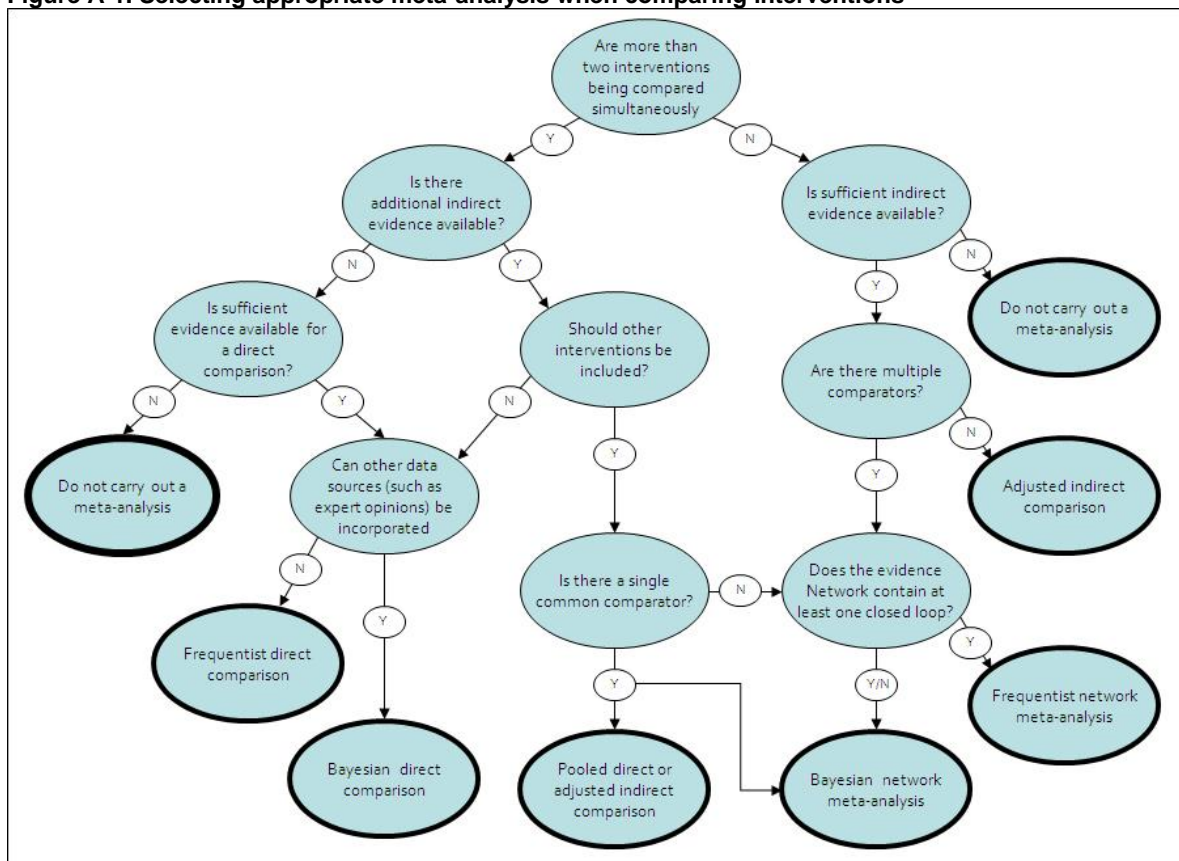
“Pursuit of qualitative or quantitative indirect comparison is never required and decisions to do so must depend on consideration of clinical, methodological, and statistical heterogeneity levels across the individual studies.”⁵

“CRGs should be encouraged to identify existing IRs that compare multiple interventions and consider the feasibility of indirect comparisons and MTM.”¹⁷

“The large majority of [intervention reviews] that involve many interventions present meta-analyses of a series of pair-wise comparisons without a specific plan to integrate the various pieces of evidence. Statistical synthesis using MTM could be performed in many cases, provided that the assumptions of this approach are fulfilled.”¹⁷

Flow chart (Figure A-1) to select proper meta-analysis when comparing interventions, from the Health Information and Quality Authority.

Figure A-1. Selecting appropriate meta-analysis when comparing interventions



Adapted from: Health Information and Quality Authority. Guidelines for Evaluating the Clinical Effectiveness of Health Technologies in Ireland. Dublin: Health Information and Quality Authority; 2011. Available at: <http://www.hiqa.ie> (Last accessed on December 28, 2011)

Assumptions

“Many assumptions behind network meta-analysis methods appear to be similar to those made in standard pair-wise meta-analysis.”⁹

“The validity of an indirect comparison relies on the different subgroups of trials being similar, on average, in all other factors that may affect outcome.”¹⁶

“Indirect comparisons are often performed on the assumption of a constant relative treatment effect across varying baseline risks (with ‘adjustment’ for the event rate in the common reference group assumed to control for differences in baseline risk). This assumption however is also usually unverifiable unless there are large numbers of trials across the indirect comparison. It also inadequately incorporates all aspects that affect the exchangeability assumption.”¹¹

“For network meta-analysis, covariates that act as relative treatment effect modifiers must be similar across trials (or adjusted for using meta-regression). And, when it combines indirect evidence with direct evidence, network meta-analysis adds the assumption of consistency: The indirect evidence must be consistent with the direct evidence.”^{2,3}

“The major assumption of indirect and mixed treatment comparisons is that the direct and the indirect evidence are consistent. That is, the treatment effect dBC estimated by the BC trials, would be the same as the treatment effect estimated by the AC and AB trials if they had included B and C arms. If this is not the case the evidence is inconsistent. Factors such as (relative) treatment effects varying with disease severity may cause inconsistency (e.g. if the BC trials are done in patient populations with higher/lower baseline risks than the AB and AC trials and the treatments interact with baseline risk, the evidence will be inconsistent).”¹⁵

“MTC analysis requires a connected network; that is, for each treatment, there is a chain of pair-wise comparisons that connects it to every other treatment.”¹⁵

How to Conduct Network Meta-analyses

Planning/Design

“Objectives of network meta-analysis may include considering all relevant evidence, answering research questions in the absence of direct evidence, improving the precision of estimates by combining direct and indirect evidence, ranking treatments, and assessing the impact of certain components of the evidence network.”^{2,3}

“When a new [intervention reviews] seeks to compare multiple interventions (i.e. to determine a preferential ordering of three or more competing interventions for a particular outcome), this should be made explicit in the protocol, and appropriate methods should be planned and implemented.”¹⁷

“The principles of good practice for systematic reviews and meta-analyses should be carefully followed when conducting mixed and indirect treatment comparisons.”¹²

“To minimize error and ensure validity of findings from meta-analyses, the systematic review, whether it involves a standard, pair-wise meta-analysis or a network meta-analysis, must be designed rigorously and conducted carefully.”⁹

“The literature search for a network meta-analysis builds the network, applying the same basic standards as for a meta-analysis leading to a direct comparison” [ispor]

“It may be difficult to identify all relevant comparators for the treatments of interest, and any search involves costs and tradeoffs. It may be efficient to proceed in stages, using one of the strategies developed by Hawkins et al.”^{2,3}

“The more interventions that are included in a MTM, the greater the potential gain in precision and the greater the ability to establish whether various sources of evidence ‘agree’ with each other. Therefore, it may sometimes be useful to include interventions that are not current candidates for clinical practice, such as placebo or no treatment, or interventions that are no longer recommended or available (‘legacy treatments’).”¹⁷

“Different specification of eligibility criteria may result in differences in the structure or extent of a network, leading to discrepant findings for network meta-analyses on the same topic. This is because different combinations of direct and indirect evidence, some independent and some overlapping, contribute to the comparisons and estimates of treatment effect. Certain interventions, for example, interventions that are no longer in use, or placebos, may not be of primary interest but may be included in the network meta-analysis if they provide information concerning the interventions of interest through indirect comparisons.”¹⁵

“To ensure that all relevant studies are identified, the network meta-analyst could search de novo for all relevant studies, but this would waste valuable resources if good systematic reviews with comprehensive searches already exist. To conserve valuable resources, one might consider using data identified through existing high quality systematic reviews of relevant pair-wise treatment comparisons provided the searches in the existing reviews are up-to-date.”¹⁵

“After demonstrating that no relevant direct randomised trials exist, broaden the literature search criteria to identify all randomised trials relevant for an indirect comparison of the proposed drug and the main comparator.” [PBAC]

“The network can be restricted to include the minimum number of comparisons required to enable an indirect comparison between the technologies of interest. Alternatively it can be expanded to include as many relevant comparators as possible.”¹

“Extending mixed treatment comparisons networks to include trial comparisons not of direct interest can reduce uncertainty in the comparisons of interest.”¹⁵

Analysis Framework

Indirect comparisons should be based on “adjusted” methods, which use the common control arm of RCTS as a way to “standardize” the comparison. Different methods of increasing complexity are available.¹⁹

“Network meta-analysis can be performed within a Frequentist or Bayesian framework.”^{2,3}

“The MTC method can be used to obtain measures of effect for each of the indicated patterns. The network meta-analysis method proposed by Lumley can compare treatments in a network geometry that contains at least one closed loop.”^{7,8}

“For syntheses where all trials are two-arm, there is no reason why frequentist methods should not be used, as long as statistically sound estimators are used and appropriate steps are taken to propagate parameter uncertainty, including correlations, through the decision model.”

“Various approaches for indirect treatment comparisons have been reviewed. The mixed treatment comparison approaches by Lu and Ades are elegant, but require information that may not be available. The challenge of Lumley’s network meta-analysis is that it needs a data-based assessment of trial consistency; therefore, it requires information from a large number of different treatment comparisons.”^{7,8}

“The common Generalised Linear Model (GLM) framework can, of course, be applied in either frequentist or Bayesian contexts. However, Bayesian Markov Chain Monte Carlo (MCMC) has for many years been the mainstay of “comprehensive decision analysis”, because simulation from a Bayesian posterior distribution supplies both statistical estimation and inference, and a platform for probabilistic decisionmaking under uncertainty”¹³

“A major advantage of the Bayesian approach is that the method naturally leads to a decision framework that supports decisionmaking”^{2,3}

“Bayesian methods based on an evidence network because of the great flexibility of the model (allowing detailed and flexible modeling of data, which can be adjusted for particular cases), estimation of inconsistency, and ability to take account of multiarm trials.”¹⁸

“A particular advantage of using a Bayesian framework is that all interventions in the analysis can be ranked, using probabilistic, rather than crude, methods.”¹⁶

“For a network meta-analysis, a specific advantage is that the posterior probability distribution allows calculating the probability of which of the competing interventions is best and other probability statements. This aspect of a Bayesian analysis is providing information that is directly relevant to health-care decisionmakers (e.g., policymakers and health-care professionals/clinicians). Other advantages of a Bayesian meta-analysis include the straightforward way to make predictions and the possibility to incorporate different sources of uncertainty.”^{2,3}

“Because Binomial and Poisson likelihoods with zero cells are allowed, special precautions do not usually need to be taken in the case of the occasional trial with a zero cell count. This is a major strength of the Bayesian MCMC approach, because some popular Frequentist approaches for log odds ratios or log relative risks have to add an arbitrary constant, usually 0.5, to cells in order to obtain non-infinite estimates of treatment effects and non-infinite variance, but in so doing they generate biased estimates of effect size.”¹³

Statistical Modeling

“Results from the naïve approach, i.e. comparing simply the treatment arm of the RCTs as if they were one single trial, are completely untrustworthy.”¹⁹

“When evidence is combined using indirect or mixed treatment comparison frameworks, trial randomisation must be preserved. A comparison of the results from single treatment arms from different randomised trials is not acceptable unless the data are treated as observational and appropriate steps taken to adjust for possible bias and increased uncertainty.”¹²

“Extending network meta-analysis models with treatment-by-covariate interactions attempts to explain heterogeneity in relative treatment effects and estimates relative treatment effects for different levels of the covariate... Unfortunately, the number of studies in a network is often limited, and in such cases, adjustment by incorporating study-level covariates with meta-regression models may sometimes be questionable. In addition, aggregate-level covariate adjustment might produce ecological bias, limiting the interpretation of estimated results for subgroups.”^{2,3}

“If confounders are present in an indirect comparison, it is only possible to adjust for them via meta-regression. However, this would be an unusual situation because at least 10 trials per adjustment variable are required in order to achieve stability in the meta-regression results”¹¹

“Network meta-analysis can be performed with fixed- or random-effects models...If there is heterogeneity, however—variation in true (or underlying) relative treatment effects for a particular pairwise comparison—random-effects models must be used. A random-effects approach typically assumes that true relative effects across studies are considered exchangeable (i.e., the prior position of expecting underlying effects to be similar but not identical) and can be described as a sample from a normal distribution whose mean is the pooled relative effect and whose SD reflects the heterogeneity.”^{2,3}

“Because the standard error obtained from a fixed effect analysis will be too small if there is heterogeneity between trials (beyond random variation), and recognising that there may be additional heterogeneity in an indirect comparison compared to a direct comparison, the Working Group supports the conclusion in the 2005 Glenny AM, et al paper that a random effects method is preferred to a fixed effect method.”¹¹

“Choices of prior distributions are, to some extent, arbitrary...”^{2,3}

“We recommend vague or flat priors, such as $N(0, 1002)$, throughout for μ and d .”¹³

“It has become standard practice to also set vague priors for the between-trial variances. For binomial with logit links models the usual practice is to place a Uniform prior on the standard deviation, for example $\sigma \sim \text{Uniform}(0,2)$...An alternative approach, which was once popular but has since fallen out of favour, is to set a vague Gamma prior on the precision, for example $1/\sigma^2 \sim \text{Gamma}(.001, .001)$.”¹³

“The parameters in the distributions of random effects have vague prior distributions: $N(0, 106)$ for the d_{Ak} (independently) and $Uniform(0, 2)$ for σ . These priors are common choices in such models.”^{2,3}

“Two further alternatives may be found useful when there is insufficient data to adequately estimate the between-trials variation. The first is the use of external data. If there is insufficient data in the metaanalysis, it may be reasonable to use an estimate for σ from a larger meta-analysis on the same trial outcome involving a similar treatment for the same condition. If there is no data on similar treatments and outcomes that can be used, an informative prior can be elicited from a clinician who knows the field.”¹³

“Particular care must be taken in checking convergence, and we suggest that at least three chains are run, starting from widely different (yet sensible) initial values...Posteriors should be examined visually for spikes and unwanted peculiarities, and both the initial “burn-in” and the posterior samples should be conservatively large and the number of iterations for both must be reported in the analysis.”¹³

“While the likelihood is not altered by a change in which treatment is taken to be “Treatment 1 [referent], the choice of the reference treatment can affect the posterior estimates because priors cannot be totally non-informative...Choice should therefore be based on ease of interpretation, with placebo or standard treatment usually taken as Treatment 1.”¹³

“It is incorrect to analyze the pairwise effects in a multiarm trial as if they came from separate studies.”^{2,3}

“If the network appropriately includes a multiarm trial, omitting it from the analysis may introduce bias. The analysis, then, must take into account the correlation among the effect estimates for the pairs of arms;”^{2,3}

“Shared parameter models allow the user to generate a single coherent synthesis when trials report results in different formats. For example some trials may report binomial data for each arm, while others report only the estimated log odds ratios and their variances; or some may report numbers of events and time at risk, while others give binomial data at given follow-up times.”¹³

“The consistency of the comparative treatment effect across trials (and sets of trials) also depends upon whether the appropriate measure of effect is used.... If an appropriate measure of comparative treatment effect is used to minimise variation in comparative treatment effect within each and all sets of included randomised trials, the exchangeability assumption is more likely to be maintained.”¹¹

“...relative measures of comparative treatment effect are often a robust way of summarising the overall result of the evidence available in order to apply it to any subgroup with a particular baseline risk.”¹¹

“Whatever the method of analysis, the pooling of individual study results and indirect comparisons should be based on relative effect measures (e.g., OR, difference in change from baseline, hazards ratio) to preserve randomization.”^{2,3}

One advantage of the OR is that, because it is symmetrical around 1.0 (unlike the RR), the OR for harm is equal to the inverse of OR for benefit, and hence is consistently estimated regardless of how the research question is framed (eg in a study that is to measure survival, the researchers could use a null hypothesis of no difference in survival, or a null hypothesis of no difference in mortality).¹¹

“If the underlying baseline risk is the same across the two sets of trials and the PBS population, then there it may be considered appropriate to use the directly synthesised RD as an absolute measure of comparative treatment effect... If the baseline risk is different, then the primary issue for the indirect comparison is whether the trials are similar in terms of potential confounders... If it is decided to proceed with the indirect comparison, then a ratio measure (OR or RR) is usually preferred to the RD, because as outlined above, it is considered that relative measures of comparative treatment effect have more often been observed to be constant across different baseline risks than absolute measures of comparative treatment effect.”¹¹

Assessment for and Handling of Potential Bias/Inconsistency

“Before comparing the proposed medicine with the main comparator, the comparability of the two sets of trials must be established.”²¹

“When direct evidence and indirect evidence are combined for a particular pairwise comparison, it is important that the indirect estimate is not biased and there is no discrepancy between the direct and indirect comparisons. Therefore, consistency between these direct and indirect comparisons should be accounted for.”^{2,3}

“Heterogeneity, inconsistency, and bias may propagate through a network of trials, and may affect the estimates differentially across regions of the network.”⁹

“The indirect comparison across trials does not have a randomisation step to allow the characteristics of the patients to differ only due to the play of chance.”¹⁰

“The indirect comparisons involved are not randomized comparisons, and may suffer the biases of observational studies, for example due to confounding”.¹⁶

“...it is important to remember that in a network meta-analysis of RCTs, the value of randomization does not hold across trials.... Hence, an ITC or network meta-analysis of RCTs is a form of observational evidence, but arguably less prone to confounding bias than is a cohort study(or any other observational design).”^{2,3}

“the mechanisms that potentially could create “bias” in indirect comparisons appear be to identical to those that cause heterogeneity in pair-wise metaanalysis.”¹⁴

“Inconsistency can be thought of as a conflict between “direct” evidence on a comparison between treatments B and C, and “indirect” evidence gained from AC and AB trials. Like heterogeneity, inconsistency is caused by effect-modifiers, and specifically by an imbalance in the distribution of effect modifiers in the direct and indirect evidence.”¹⁴

“Factors such as the total number of trials in a network, number of trials with more than two comparison arms, heterogeneity (i.e., clinical, methodological, and statistical variability within direct and indirect comparisons), inconsistency (i.e., discrepancy between direct and indirect comparisons), and bias may influence effect estimates obtained from network meta-analyses.”⁹

“In principle, the validity of indirect comparison relies on the invariance of treatment effects across study populations. However, in practice, trials can vary in numerous ways including population characteristics, interventions and cointerventions, length of followup, loss to followup, study quality, etc. Given the limited information in many publications and the inclusion of multiple treatments, the validity of indirect comparisons is often unverifiable. Moreover, indirect comparisons, like all other meta-analyses, essentially constitute an observational study, and residual confounding can always be present. Systematic differences in characteristics among trials in a network can bias indirect comparison results. In addition, all other considerations for meta-analyses, such as choice of effect measures or heterogeneity, also apply to indirect comparisons.”⁴

The ICWG report quotes Glenny et al.’s definition of inconsistency (they call it exchangeability), “... the two sets of trials should be exchangeable, in the sense that there is no reason to suppose that the results as a whole would be different had the various trialists kept the same protocol and patients, but chosen to study a different treatment comparison.”¹¹

“Most agencies to which the results of a network meta-analysis could be submitted currently require that direct estimates and indirect estimates be calculated separately and shown to be consistent before direct evidence and indirect evidence are combined.”^{2,3}

“...network meta-analysis relies on the randomization in the RCTs that compared the treatments directly. It also involves a similarity assumption: “Combining studies should only be considered if they are clinically and methodologically similar”. Nevertheless, “no commonly accepted standard [defines] which studies are ‘similar enough.’”^{2,3}

“In a multiple treatment comparison involving both direct and indirect evidence, the evidence network can become very complex with many comparisons based on only one or two studies. With increasing complexity and greater numbers of treatments, the prospect of inconsistency increases. There is also a power trade-off between the number of pair-wise comparisons and the number of studies included in the analysis – too many comparisons with too few studies and the analysis may be underpowered to detect true differences.”¹

The ICWG report¹¹ provides an example framework for assessing the exchangeability assumption of a network meta-analysis. Assuming an indirect comparisons of treatments A and B through a common comparator C is being considered, ICWG first recommends for the AvC and BvC direct randomized trials:

- Assessment of the available trials for factors that may cause heterogeneity of the AvC and BvC comparative treatment effect
- Assessment of the event rates in the drug C populations
- Assessment of whether the measure of the comparative treatment effect for AvC and BvC is appropriate
- Assessment of evidence of the statistical homogeneity of the AvC and BvC comparative treatment effect across the available trials

Then for the BvA indirect comparison:

- Assessment across the sets of trials (i.e. the AvC and the BvC trials) for factors that may cause heterogeneity of the BvA comparative treatment effect
- Assessment of the event rates in the drug C populations across the sets of trials
- Assessment of whether the measure of the comparative treatment effect for BvA is appropriate
- Assessment of evidence of statistical homogeneity of the synthesized comparative treatment effect BvA across the sets of trials (only possible if BvA has been compared via multiple common references)

According to the CADTH, “Whether an indirect treatment comparison provides a valid estimate of the relative efficacy for an intervention of interest significantly depends on the fulfillment of this primary assumption. To determine whether or not this assumption is met, trials included in the indirect comparison can be assessed according to three criteria”:

- comparability of the linking treatment;
- comparability of patients/heterogeneity;
- methodological comparability of included trials

“...whichever [indirect comparison/network meta-analysis] method the investigators choose, they should assess the invariance of treatment effects across studies and appropriateness of the chosen method on a case-by-case basis, paying special attention to comparability across different sets of trials.”⁴

“Where direct and indirect evidence are combined, inconsistencies between the direct and indirect evidence must be assessed and reported.”¹

“Decisionmakers making use of results of network meta-analyses will need to assess whether the differences between treatments are most likely true or whether they can be explained by bias in the analysis. The internal validity of the analyses is contingent on three factors: 1) the appropriate identification of the studies that make up the evidence network, 2) the quality of the individual RCTs, and 3) the extent of confounding bias due to similarity and consistency violations.”^{2,3}

“Factors such as the total number of trials in a network, number of trials with more than two comparison arms, heterogeneity (i.e., clinical, methodological, and statistical variability within direct and indirect comparisons), inconsistency (i.e., discrepancy between direct and indirect comparisons), and bias may influence effect estimates obtained from network meta-analyses.”⁹

“Evaluation of homogeneity and consistency (if the network supports both direct and indirect comparisons) should be specified as components of the analysis and should reflect the risks and benefits of combining data for the particular research question”^{2,3}

“While it is essential to carry out tests for inconsistency, the issue should not be considered in an overly mechanical way... We emphasise that while tests for inconsistency must be carried out, they are inherently underpowered, and will often fail to detect it. Investigators must therefore also ask whether, if inconsistency is not detected, conclusions from combining direct and indirect evidence can be relied upon.”¹⁴

“...tests for statistical heterogeneity have low power, and therefore, even if statistical heterogeneity is not demonstrated, doubts will remain about its presence, particularly in the presence of obvious clinical differences across the AvC and BvC trials by a factor that is known to influence drugs B and/or A.”¹¹

“When analyzing a network of comparisons, the inconsistency of the network needs to be considered, as well as between-trial heterogeneity and sampling error. Large inconsistencies rule out a meta-analysis, small inconsistencies should add uncertainty to the results.”⁸

“A departure from consistency arises when the direct and indirect estimates of an effect differ...Researchers must evaluate departures from consistency and determine how to interpret them.”²

“The assumption of constant efficacy requires all trials included in the analysis to be equivalent and attempting to measure the same treatment effect – that is, the results of one set of trials (A vs. B) should be generalisable to the other set of trials (A vs. C). Determining whether the assumption of generalisability holds is a subjective assessment based on a detailed review of the included studies in both comparisons.”¹

“Disagreement between direct and indirect evidence must be fully investigated and it may preclude pooling data if the disagreement cannot be adequately explained.”¹

“When information on heterogeneity within the direct comparisons is available, consideration of it can form a preliminary step in a network meta-analysis, but one should first examine potential effect modifiers, because disparities among studies may preclude analysis of the network.”^{2,3}

“Consistency or coherence describes the situation that direct and indirect evidence agrees with each other, and when the evidence of a network of interventions is consistent, investigators could combine direct and indirect evidence using MTM models. Conversely, they should refrain from combining multiple sources of evidence from an incoherent network where there are substantial differences between direct and indirect evidence.”⁴

“Decisions should be based on coherent models that fit the data. Careful examination of different sources of evidence may reveal that some estimates are “corroborated” and others not. If inconsistency is detected, the entire network of evidence should be reconsidered from a clinical epidemiology viewpoint with respect to the presence of potential effect modifiers.”¹⁴

“Any adjustment in response to inconsistency is *post hoc*, which emphasizes the importance of identifying potential causes of heterogeneity of effect at the scoping stage, and potential internal biases in advance of synthesis”¹⁴

“Measures of inconsistency variance or incoherence variance are not recommended as indicators of inconsistency.”¹⁴

“Within a Bayesian framework a consistency model can be compared to an “inconsistency” model. Analyses of residual deviance can provide an “omnibus” test of global inconsistency, and can also help locate it.”¹⁴

“Node splitting is another effective method for comparing direct evidence to indirect evidence in complex networks.”¹⁴

Assessment of Model Fit

“In frequentist analyses, measures of model fit are similar to those for direct evidence and depend on the particular outcome measure. Bayesian analyses customarily use deviance (a likelihood-based measure)—the lower the residual deviance, the better the fit. For comparing models, the deviance information criterion (DIC) adds a penalty term, equal to the effective number of parameters in the model. If a model fits poorly, graphical techniques can aid more-detailed examination.”^{2,3}

“The goodness-of-fit can be estimated by calculating the difference between the deviance for the fitted model and the deviance for the saturated model (which fits the data perfectly). For example, the Akaike information criterion, which uses the likelihood function, the Bayesian information criterion, or deviance information criterion can all be used for model selection”^{2,3}

“...competing models should be compared in terms of their goodness-of-fit to the data, and residual deviance calculations may be provided to justify the study’s choice of the base case model.”^{2,3}

“In this document we suggest that global DIC statistics and res D are consulted both to compare fixed and random effect models, and to ensure that overall fit is adequate.”¹³

“The choice of a fixed- or random-effects meta-analysis model, with or without covariate interactions, can be made by comparing different competing models regarding their goodness-of-fit to the data.”^{2,3}

Use of Sensitivity Analysis

“Investigators should conduct sensitivity analysis to check the assumptions of the indirect comparison. If the results are not robust to the assumptions, findings from indirect comparisons should be considered as inconclusive.”⁴

“Sensitivity analyses should focus on the areas of greatest uncertainty. Potential effect modifiers can be explored by stratifying on variations in study design or population. Comparisons between random-effects and fixed-effects analyses may be appropriate. Bayesian analyses should also explore the influence of choosing different prior distributions.”^{2,3}

“Choices of prior distributions are, to some extent, arbitrary, so they are often subjected to sensitivity analysis, which may be especially important for priors on heterogeneity in random-effects models.”^{2,3}

How to Interpret and Report Network Meta-Analyses

Interpretation

“Probability statements could be made about the effectiveness of each treatment [24]. For example, for each treatment, one can calculate the probability that the treatment is the best, second best, or third best among all treatments. Such probability statements should be interpreted carefully since the difference between treatments might be small and not clinically meaningful.”⁹

“Investigators should explicitly state assumptions underlying indirect comparisons and conduct sensitivity analysis to check those assumptions. If the results are not robust, findings from indirect comparisons should be considered inconclusive. Interpretation of findings should explicitly address these limitations.”⁴

In respect to Bayesian network meta-analysis, “Probability statements could be made about the effectiveness of each treatment.”⁹

“The external validity of the network meta-analysis will naturally be limited by the external validity of the RCTs included in the evidence network, and health-care decisionmakers will need to review whether results can be extrapolated to the population of interest.”^{2,3}

“Furthermore identification of the “best” or most appropriate treatment cannot be made on the basis of efficacy end points alone. To inform health-care decisionmaking for clinical treatment guidelines and reimbursement policies, the efficacy findings of a network meta-analysis must be interpreted in light of other available (observational) evidence and other characteristics of the competing interventions, such as safety and convenience”.^{2,3}

“The network of available evidence should be described and used to guide the selection of the method of meta-analysis. The selection of direct and indirect evidence must be clearly defined. The exclusion of relevant evidence, either direct or indirect, should be highlighted and justified. Where direct and indirect evidence are combined, inconsistencies between the direct and indirect evidence must be assessed and reported.”¹

“An approach based on a network of trials can incorporate both non-inferiority and superiority trials and so unify interpretation of the results of these different types of trials, without taking into account the non-inferiority margins used (which very frequently cannot be justified).”¹⁸

“There are two types of potential errors when interpreting the results of indirect comparisons, mainly those derived from networks of comparisons:

1. drawing conclusions of equivalent efficacy for two treatments when there is no statistically significant difference
2. and within an indirect comparison, establishing an incorrect hierarchy by naive comparison of point estimates.”¹⁸

Guidance from the Haute Autorite de Sante provides a brief “critical review guide” with six main sections:¹⁸

1. acceptability of the approach used;
2. search strategy and selection process for data contributing to the indirect comparison calculations;
3. clinical homogeneity of trials and stability of effects;
4. consistency of estimates;
5. degree of concordance of the result with that of existing direct comparisons;
6. correct interpretation of results in the proposed conclusions.

The NICE Decision Support Unit technical support document #7 provides a “reviewer checklist” for evidence synthesis reports, which addresses “issues specific to network synthesis” including:¹³⁻¹⁵

*“C1. Adequacy of information on model specification and software implementation
C2. Multi-arm trials
C3. Connected and disconnected networks
C4. Inconsistency”*

Reporting

“In addition to the estimates of treatment effects, uncertainty, clinical and methodological characteristics, and potential biases within included trials must be conveyed.”⁹

“If the analyses were performed within a Bayesian framework, the choice of prior distributions for the model parameters should be defined.”^{2,3}

“Indicate software package used in the analysis and provide code (at least in an online appendix)”

“Evidence from a mixed treatment comparison may be presented in a variety of ways. The network of evidence may be presented in tabular form. It may also be presented diagrammatically as long as the direct and indirect treatment comparisons are clearly identified and the number of trials in each comparison is stated.”¹²

“In order to appreciate the value of a network meta-analysis, it is recommended that results of all (relevant) pairwise comparisons (as a reflection of the functional parameters) are presented as well.”^{2,3}

“It is critical to report all pairwise effect estimates together with the associated confidence or credible intervals, depending on the statistical model used (i.e., frequentist or Bayesian model).”⁹

“Investigators should make efforts to explain the differences between direct and indirect evidence based upon study characteristics.”⁴

“The heterogeneity between results of pairwise comparisons and inconsistencies between the direct and indirect evidence on the technologies should be reported.”¹²

“... the choice of an indirect instead of a direct head-to-head comparison between the study treatment and the comparator should be explained, together with the limitations of the indirect comparison.”¹⁹

References

1. Caldwell DM, Ades AE, Higgins PT. Simultaneous comparison of multiple treatments: combining direct and indirect evidence. *BMJ* 2005;331:897-900. PMID:16223826Health Information and Quality Authority. Guidelines for Evaluating the Clinical Effectiveness of Health Technologies in Ireland. Dublin: Health Information and Quality Authority; 2011. Available at: <http://www.hiqa.ie> (Last accessed on December 28, 2011)
2. Jansen JP, Fleurence R, Devine B, Itzler R, Barrett A, Hawkins N, Lee K, Boersma C, Annemans L, Cappelleri JC. Interpreting indirect treatment comparisons and network meta-analysis for health-care decision making: report of the ISPOR Task Force on Indirect Treatment Comparisons Good Research Practices: part 1. *Value Health*. 2011;14:417-28. PMID:21669366
3. Hoaglin DC, Hawkins N, Jansen JP, et al. Conducting indirect-treatment-comparison and network-meta-analysis studies: report of the ISPOR Task Force on Indirect Treatment Comparisons Good Research Practices: part 2. *Value Health*. 2011;14:429-37. PMID:21669367
4. Fu R, Gartlehner G, Grant M, et al. Conducting Quantitative Synthesis When Comparing Medical Interventions: AHRQ and the Effective Health Care Program. In: Agency for Healthcare Research and Quality. *Methods Guide for Comparative Effectiveness Reviews* [posted October 2010]. Rockville, MD. Available at: <http://effectivehealthcare.ahrq.gov/> (Last accessed on December 28, 2011).
5. Drug Effectiveness Review Project. *Systematic Review Methods and Procedures*. Portland, OR, 2011. Available at: http://www.ohsu.edu/xd/research/centers-institutes/evidence-based-policy-center/derp/documents/upload/DERP_METHODS_WEB_Final_January-2011-2.pdf (Last accessed on December 28, 2011).
6. Centre for Reviews and Dissemination. *CRD's Guidance for Undertaking Reviews in Health Care*. University of York, 2009. Available at: http://www.york.ac.uk/inst/crd/pdf/Systematic_Reviews.pdf (Last accessed on December 28, 2011).
7. Wells GA, Sultan SA, Chen L, et al. *Indirect Evidence: Indirect Treatment Comparisons in Meta-Analysis*. Ottawa: Canadian Agency for Drugs and Technologies in Health; 2009. Available at: <http://www.cadth.ca> (Last accessed on December 28, 2011).

8. Guidelines for the economic evaluation of health technologies: Canada [3rd Edition]. Ottawa:Canadian Agency for Drugs and Technologies in Health; 2006. Available at: <http://www.cadth.ca> (Last accessed on December 28, 2011).
9. Li T, Puhan MA, Vedula SS, et al. Ad Hoc Network Meta-analysis Methods Meeting Working Group. Network meta-analysis-highly attractive but more methodological research is needed. *BMC Med.* 2011;9:79. PMID:21707969
10. Pharmaceutical Benefits Advisory Committee. Guidelines for Preparing Submissions to the Pharmaceutical Benefits Advisory Committee (version 4.3); Australian Government, Department of health and Ageing. 2008. Available at: [http://www.health.gov.au/internet/main/publishing.nsf/content/AECB791C29482920CA25724400188EDB/\\$File/PBAC4.3.2.pdf](http://www.health.gov.au/internet/main/publishing.nsf/content/AECB791C29482920CA25724400188EDB/$File/PBAC4.3.2.pdf) (Last accessed on December 28, 2011).
11. Report of the Indirect Comparisons Working Group (ICWG) to the Pharmaceutical Benefits Advisory Committee: Assessing Indirect Comparisons. Australian Government, Department of health and Ageing. Available at: [http://www.health.gov.au/internet/main/publishing.nsf/Content/B11E8EF19B358E39CA25754B000A9C07/\\$File/ICWG%20Report%20FINAL2.pdf](http://www.health.gov.au/internet/main/publishing.nsf/Content/B11E8EF19B358E39CA25754B000A9C07/$File/ICWG%20Report%20FINAL2.pdf) (Last accessed on December 28, 2011).
12. National Institute for Health and Clinical Excellence. Guide to the methods of technology appraisal, 2008. Available at: <http://www.nice.org.uk> (Last accessed on December 28, 2011).
13. Dias, S., Welton, N.J., Sutton, A.J. et al. A.E. NICE DSU Technical Support Document 2:A Generalised Linear Modeling Framework for Pairwise and Network Meta-Analysis of Randomised Controlled Trials, 2011; last updated August 2011. Available at: <http://www.nicedsu.org.uk> (Last accessed on December 28, 2011).
14. Dias, S., Welton, N.J., Sutton, A.J. et al. A.E. NICE DSU Technical Support Document 4: Inconsistency in Networks of Evidence Based on Randomised Controlled Trials. 2011. Available at: <http://www.nicedsu.org.uk> (Last accessed on December 28, 2011).
15. National Institute for Health and Clinical Excellence. Briefing paper for methods review workshop on evidence synthesis(indirect and mixed treatment comparisons), 2007. Available at: <http://www.nice.org.uk/media/4A6/2F/EvidenceSynthesisBriefingPaper.pdf> (Last accessed on December 28, 2011).
16. Higgins JPT, Green S (eds). *Cochrane Handbook for Systematic Reviews of Interventions* Version 5.1.0 [updated March 2011]. The Cochrane Collaboration, 2011. Available from www.cochrane-handbook.org (Last accessed on December 28, 2011).
17. Evolution of Cochrane Intervention Reviews and Overviews of Reviews to better accommodate comparisons among multiple interventions. Report from a meeting of the Cochrane Comparing Multiple Interventions Methods Groups, Milan, March 2011. Available at:<http://cmimg.cochrane.org/sites/cmimg.cochrane.org/files/uploads/CMIMG%20summary%20of%20meeting%20Milan%20March%202011.pdf> (Last accessed on December 28, 2011).
18. Haute Autorite de Sante. Indirect comparisons Methods and validity, 2009. Available at: http://www.has-sante.fr/portail/upload/docs/application/pdf/2011-02/summary_report__indirect_comparisons_methods_and_validity_january_2011_2.pdf Accessed June 10, 2012

19. Cleemput I, Van Wilder P, Vrijens F, Huybrechts M, Ramaekers D. Guidelines for Pharmacoeconomic Evaluations in Belgium. Health Technology Assessment (HTA). Brussels: Health Care Knowledge Centre (KCE); 2008. KCE Reports 78C (D/2008/10.273/27). Available at: http://www.ispor.org/peguidelines/source/Belgium_Guidelines-for-Pharmacoeconomics-Evaluation-in-Belgium_2008_English.pdf Accessed June 10, 2012.
20. German Institute for Quality and Efficiency in Health Care. Benefit Assessment of Pharmaceuticals pursuant to § 35a SGB V, 2008. Available at: <http://www.english.g-ba.de/benefitassessment/information/> [Accessed June 10, 2012.](#)
21. National Department of Health – Republic of South Africa. The Guidelines For Pharmacoeconomic Evaluations of Medicines and Scheduled Substances, 2010. Available at: http://www.ispor.org/PEguidelines/source/guidelines-for-pharmacoeconomic-evaluations-of-medicines-and-scheduled-substances_2010.pdf Accessed June 10, 2012.

Appendix B. Literature Search for Part Two

1. Randomized Controlled Trial/
2. Clinical Trial/
3. randomi\$ control\$ trial\$.tw.
4. controlled clinical trial.sh.
5. clinical trial\$.tw.
6. trial\$.tw.
7. 1 or 2 or 3 or 4 or 5 or 6
8. review literature/
9. meta-analysis.sh.
10. meta-analy\$.tw.
11. metaanaly\$.tw.
12. (meta adj analy\$).tw.
13. 8 or 9 or 10 or 11 or 12
14. (indirect adj2 comparison\$).tw.
15. (indirect adj2 evaluat\$).tw.
16. (indirectly adj2 compare\$).tw.
17. bayesian.tw.
18. (mixed treatment adj compar\$).tw.
19. MTC.tw.
20. 14 or 15 or 16 or 17 or 18 or 19
21. 7 and 13
22. 20 and 21
23. limit 22 to english language
24. limit 23 to yr="2006 -Current"
25. remove duplicates from 24

Appendix C. Data Extraction Tool for Part Two

Study identification

Unique ID		First author last name, year	
# authors		In there a methodologist listed as an author?	<input type="checkbox"/> Yes <input type="checkbox"/> No
Journal Name		Journal impact factor	
Is the journal classified as a methods journal? <input type="checkbox"/> Yes <input type="checkbox"/> No		Does journal allow online supplement/ appendix? <input type="checkbox"/> Yes <input type="checkbox"/> No	
Was there a published appendix or online supplement? <input type="checkbox"/> Yes <input type="checkbox"/> No		Does the journal impose a word/table/figure limit? Word: <input type="checkbox"/> Yes <input type="checkbox"/> No Table/figure: <input type="checkbox"/> Yes <input type="checkbox"/> No If yes, what is the limit:	
Geographic location of conduction?		# printed pages in main document	
Funding Source: <input type="checkbox"/> Industry <input type="checkbox"/> Government/Foundation <input type="checkbox"/> Academia <input type="checkbox"/> Other <input type="checkbox"/> Unknown			
Publication type: <input type="checkbox"/> Full text journal article <input type="checkbox"/> Report (government, etc) <input type="checkbox"/> Other			
Work affiliated with an agency? (ex. AHRQ, NICE, Cochrane, etc.) <input type="checkbox"/> Yes <input type="checkbox"/> No If yes, which agency:			
What terms were used to describe the indirect comparison? <input type="checkbox"/> Network meta-analysis <input type="checkbox"/> Mixed treatment comparison <input type="checkbox"/> Multiple treatment comparison <input type="checkbox"/> Other (i.e., simply by reference(s) used; exact terms):			

Study characteristics

Study objective:	
	Was it clear how the research question pertains to a network meta-analysis? <input type="checkbox"/> Yes <input type="checkbox"/> No
Disease state evaluated	<input type="checkbox"/> Endocrinology <input type="checkbox"/> Behavioral health <input type="checkbox"/> Cardiology <input type="checkbox"/> Oncology <input type="checkbox"/> Pain <input type="checkbox"/> Substance abuse <input type="checkbox"/> Respiratory <input type="checkbox"/> Infectious disease <input type="checkbox"/> Rheumatology <input type="checkbox"/> Gastroenterology <input type="checkbox"/> Neurology <input type="checkbox"/> Other:
Methodological inclusion criteria?	
What network pattern was present? <input type="checkbox"/> simple star <input type="checkbox"/> star <input type="checkbox"/> ladder <input type="checkbox"/> closed loop <input type="checkbox"/> network with at least one closed loop	
Was a diagram displayed to show the network? <input type="checkbox"/> Yes <input type="checkbox"/> No	
#and type of interventions compared? (e.g device, procedure, pharmacologic, behavioral, other)	

# of trials / # patients included in analysis:	
--	--

Methods Characteristics

Method/model applied: <input type="checkbox"/> Bayesian <input type="checkbox"/> Frequentist
Was traditional pair-wise meta-analysis also conducted? <input type="checkbox"/> Yes <input type="checkbox"/> No

For Bayesian networks

<p>Model (all that apply): <input type="checkbox"/> Fixed-effects <input type="checkbox"/> Random-effects <input type="checkbox"/> Adjustment of model for studies with ≥ 3 treatments? <input type="checkbox"/> Evaluation on the dependence of treatment effect on a co-variate (adjustment) performed?</p>
<p>Software used (including wrappers):</p> <p>Was the code published in the main manuscript? <input type="checkbox"/> Yes <input type="checkbox"/> No If no, was the code made available to the reader? <input type="checkbox"/> Yes <input type="checkbox"/> No If it was made available to the reader, in what format? <input type="checkbox"/> online supplement <input type="checkbox"/> referral to another website/source <input type="checkbox"/> email author <input type="checkbox"/> other:</p> <p>If email author, were we able to obtain the code for this project? <input type="checkbox"/> Yes <input type="checkbox"/> No</p> <p>Was the raw data published in the main manuscript? <input type="checkbox"/> Yes <input type="checkbox"/> No If no, was the raw data made available to the reader? <input type="checkbox"/> Yes <input type="checkbox"/> No If it was made available to the reader, in what format? <input type="checkbox"/> online supplement <input type="checkbox"/> referral to another website/source <input type="checkbox"/> email author <input type="checkbox"/> other:</p> <p>If email author, were we able to obtain the raw data for this project? <input type="checkbox"/> Yes <input type="checkbox"/> No</p>
<p>Was Markov-chain Monte Carlo modeling used? <input type="checkbox"/> Yes <input type="checkbox"/> No If no, what sampling method was used?</p> <p>Were the starting value(s) reported (this can be obtained from provided code)? <input type="checkbox"/> Yes <input type="checkbox"/> No</p> <p>Number of chains:</p> <p>Number of iterations per chain:</p> <p>Number of iterations used for final results (after excluding burn-in):</p> <p>Were convergence statistics evaluated? <input type="checkbox"/> Yes <input type="checkbox"/> No</p>
<p>Were prior distributions specified anywhere in the paper? (this can be obtained from provided code) <input type="checkbox"/> Yes <input type="checkbox"/> No If yes, what distribution was used for "D" and "σ" [often $N(0, 10^6)$ for D and Uniform(0, 2) for σ] (this can be obtained from provided code)?</p> <p>Were prior distributions justified in the paper? <input type="checkbox"/> Yes <input type="checkbox"/> No <input type="checkbox"/> NA if not specified</p> <p>Was sensitivity analysis performed based on prior distribution chosen? <input type="checkbox"/> Yes <input type="checkbox"/> No If yes, what was the distribution changed to?</p>
<p>Was a graphical representation of the posterior distribution provided?</p>

<input type="checkbox"/> Yes <input type="checkbox"/> No
Do the authors rank order the efficacy and/or safety of different interventions compared? <input type="checkbox"/> Yes <input type="checkbox"/> No
Was model fit tested (i.e., sum deviation, residual deviation, DIC)? <input type="checkbox"/> Yes <input type="checkbox"/> No If so, which was used?
Was a description of how possible heterogeneity was evaluated (either qualitative or quantitative, e.g., I^2, Cochrane Q, etc.) provided? <input type="checkbox"/> Yes <input type="checkbox"/> No If yes, how? <input type="checkbox"/> traditional meta-analysis, how: <input type="checkbox"/> network meta-analysis, how:
Was a description of how possible inconsistency was evaluated (either qualitative or quantitative, e.g., comparison of direct evidence with the indirect evidence) provided? <input type="checkbox"/> Yes <input type="checkbox"/> No
Does the analysis try to make a claim of: Equivalence <input type="checkbox"/> Yes <input type="checkbox"/> No Non-inferiority? <input type="checkbox"/> Yes <input type="checkbox"/> No
Was there an <i>a priori</i> decision rule/minimally important difference described? <input type="checkbox"/> Yes <input type="checkbox"/> No

For Frequentist networks

Model (all that apply): <input type="checkbox"/> Fixed-effects <input type="checkbox"/> Random-effects <input type="checkbox"/> Evaluation on the dependence of treatment effect on a co-variate (adjustment) performed?
Software used: Was the raw data published in the main manuscript? <input type="checkbox"/> Yes <input type="checkbox"/> No If no, was the raw data made available to the reader? <input type="checkbox"/> Yes <input type="checkbox"/> No If it was made available to the reader, in what format? <input type="checkbox"/> online supplement <input type="checkbox"/> referral to another website/source <input type="checkbox"/> email author <input type="checkbox"/> other: If email author, were we able to obtain the raw data for this project? <input type="checkbox"/> Yes <input type="checkbox"/> No
Was a Linear Mixed Model Used? <input type="checkbox"/> Yes <input type="checkbox"/> No If no, how was the model fit? How were studies weighted (inverse variance, inverse sample size etc?):
Was a description of how possible heterogeneity was evaluated (either qualitative or quantitative, e.g., I^2, Cochrane Q, etc.) provided? <input type="checkbox"/> Yes <input type="checkbox"/> No If yes, how? <input type="checkbox"/> traditional meta-analysis, how: <input type="checkbox"/> network meta-analysis, how:
Was a description of how possible inconsistency was evaluated (either qualitative or quantitative, e.g., comparison of direct evidence with the indirect evidence) provided? <input type="checkbox"/> Yes <input type="checkbox"/> No

Does the analysis try to make a claim of: Equivalence <input type="checkbox"/> Yes <input type="checkbox"/> No Non-inferiority? <input type="checkbox"/> Yes <input type="checkbox"/> No
Was there an <i>a priori</i> decision rule/minimally important difference described? <input type="checkbox"/> Yes <input type="checkbox"/> No

Posterior Distribution

Outcome 1: <input type="checkbox"/> Binary <input type="checkbox"/> Continuous <input type="checkbox"/> Categorical non binary Is this outcome effect measure reported as mean or median data? <input type="checkbox"/> Mean <input type="checkbox"/> Median <input type="checkbox"/> NR Format presented: <input type="checkbox"/> Text <input type="checkbox"/> Table <input type="checkbox"/> Figure
Effect size measured: <input type="checkbox"/> Relative risk <input type="checkbox"/> Odds ratio <input type="checkbox"/> Risk difference <input type="checkbox"/> Weighted-mean difference <input type="checkbox"/> Other:
Measure of variance: <input type="checkbox"/> Credible interval, if yes <input type="checkbox"/> 99% <input type="checkbox"/> 95% <input type="checkbox"/> SD <input type="checkbox"/> Other:
Outcome 2: <input type="checkbox"/> Binary <input type="checkbox"/> Continuous <input type="checkbox"/> Categorical non binary Is this outcome effect measure reported as mean or median data? <input type="checkbox"/> Mean <input type="checkbox"/> Median <input type="checkbox"/> NR Format presented: <input type="checkbox"/> Text <input type="checkbox"/> Table <input type="checkbox"/> Figure
Effect size measured: <input type="checkbox"/> Relative risk <input type="checkbox"/> Odds ratio <input type="checkbox"/> Risk difference <input type="checkbox"/> Weighted-mean difference <input type="checkbox"/> Other:
Measure of variance: <input type="checkbox"/> Credible interval, if yes <input type="checkbox"/> 99% <input type="checkbox"/> 95% <input type="checkbox"/> SD <input type="checkbox"/> Other:
Outcome 3: <input type="checkbox"/> Binary <input type="checkbox"/> Continuous <input type="checkbox"/> Categorical non binary Is this outcome effect measure reported as mean or median data? <input type="checkbox"/> Mean <input type="checkbox"/> Median <input type="checkbox"/> NR Format presented: <input type="checkbox"/> Text <input type="checkbox"/> Table <input type="checkbox"/> Figure
Effect size measured: <input type="checkbox"/> Relative risk <input type="checkbox"/> Odds ratio <input type="checkbox"/> Risk difference <input type="checkbox"/> Weighted-mean difference <input type="checkbox"/> Other:
Measure of variance: <input type="checkbox"/> Credible interval, if yes <input type="checkbox"/> 99% <input type="checkbox"/> 95% <input type="checkbox"/> SD <input type="checkbox"/> Other:
Outcome 4: <input type="checkbox"/> Binary <input type="checkbox"/> Continuous <input type="checkbox"/> Categorical non binary Is this outcome effect measure reported as mean or median data? <input type="checkbox"/> Mean <input type="checkbox"/> Median <input type="checkbox"/> NR Format presented: <input type="checkbox"/> Text <input type="checkbox"/> Table <input type="checkbox"/> Figure
Effect size measured: <input type="checkbox"/> Relative risk <input type="checkbox"/> Odds ratio <input type="checkbox"/> Risk difference <input type="checkbox"/> Weighted-mean difference <input type="checkbox"/> Other:
Measure of variance: <input type="checkbox"/> Credible interval, if yes <input type="checkbox"/> 99% <input type="checkbox"/> 95% <input type="checkbox"/> SD <input type="checkbox"/> Other:

Appendix D. Focus Group Questions

AHRQ Network Meta-analysis Methods Project

Insight into Meta-analyses of Networks of Studies

Background: Several methodologies exist to indirectly compare interventions, as do modes to implement such methodologies. These include anchored indirect comparisons as described by Bucher et al., Lumley's Frequentist network meta-analysis (of networks with at least one closed loop) and Bayesian network meta-analysis (commonly referred to as mixed treatment comparison (MTC)). In the simplest form, interventions that are compared in separate trials to a common comparator can be compared indirectly in the anchored indirect treatment comparison. However, as a generalization of indirect comparisons, when more than two treatments are being compared indirectly, and at least one pair of treatments is being compared both directly and indirectly (a closed loop is present), both direct and indirect types of data can be used to estimate effects in a network meta-analysis. Although Lumley's Frequentist network meta-analysis and Bayesian MTC have been used to synthesize networks of studies with at least one closed loop, best practices for their use are unclear.

Invitation to Participate: You have been chosen to participate in this focus group given your involvement as a producer of a Lumley's Frequentist network meta-analysis or a Bayesian MTC in the past few years.

This research is funded by the Agency of Healthcare Research and Quality (AHRQ) and is being undertaken by the University of Connecticut/Hartford Hospital Evidence-based Practice Center (UC/HH EPC). The Lead Investigator of this study is Dr. Craig I. Coleman, Co-Director and Methods-Chief of the UC/HH EPC, based at the University of Connecticut School of Pharmacy (UCSoP).

Instructions: As a participant, we are asking that you thoroughly and conscientiously complete the following questionnaire. We anticipated this should take you about 10-15 minutes. All participants will be acknowledged in the resulting published AHRQ report.

When asked to answer questions regarding your specific network meta-analysis, please note that we are referring to the published work defined in the email message sent to you.

If you have any questions related to this survey, please contact:

Craig I. Coleman

Co-Director and Methods-Chief

University of Connecticut/Hartford Hospital Evidence Based Practice Center

Hartford, Connecticut, USA

Email: ccolema@harthosp.org

Tel: 860-545-2096

Fax: 860-545-2277

Please review the following before starting.

Please Note: For the purposes of this questionnaire, we will use the following specific definitions:

- **Network meta-analysis** = Simultaneous synthesis of evidence of all pairwise comparisons across >2 interventions.
- **Closed loop network of evidence** = A network of evidence where >2 interventions are being compared indirectly, and at least one pair of interventions is being compared both directly and indirectly.
- **Mixed treatment comparison (MTC)** = The Bayesian approach as described by Lu and Ades whereby both direct and indirect evidence for particular pair-wise comparisons can be combined, and interventions that have not been compared directly are linked through common comparators.
- **Lumley's Frequentist network meta-analysis** = The Frequentist approach originally described by Lumley whereby both direct and indirect evidence are combined when there is at least one closed loop of evidence connecting two interventions of interest (not Bucher's method of anchored/adjusted indirect comparison).

Demographic Information

1. Work setting
 - a. Academic
 - b. Nonacademic

2. Are you affiliated with an organization involved in conducting evidence synthesis/systematic review/meta-analysis (i.e., AHRQ, Cochrane, NICE)?
 - a. Yes
 - b. No

If yes, which _____ (list all that apply)

3. Do you consider yourself to personally have the expertise needed to implement a network meta-analysis on your own?
 - a. Yes
 - b. No

If yes, which of the following methods (check all that apply)?

- i. Bayesian mixed treatment comparison
 - ii. Frequentist network meta-analysis

4. Prior to conducting your network meta-analysis identified at the beginning of this survey, how would you describe your experience with the methodology?
 - a. Knew about network meta-analysis and had used it before

- b. Knew about network meta-analysis but had not used it before
 - c. Never heard of it
5. Have you had any formal or informal training in network meta-analysis methods?
- a. Yes
 - b. No
 - c. If yes, please describe:_____
6. How many network meta-analyses have you been involved in conducting?
- a. Just this one
 - b. 2-4
 - c. 5 or more
7. What was your role on the network meta-analysis identified at the beginning of this questionnaire (check all that apply)?
- a. Clinical advice/clinical interpretation/policy development
 - b. Protocol development
 - c. Developed search strategy
 - d. Data extraction
 - e. Statistical advice/methodologist
 - f. Writing or critical revision of manuscript/report
 - g. Obtaining of funding
 - h. Other (explain):_____

Using the 5-point Likert scale, please respond to the following statements in regard to network meta-analysis in general.

1= strongly disagree; 2=disagree; 3=neutral; 4=agree; 5=strongly agree

8. The terms “network meta-analysis” is used unambiguously and consistently in the medical literature.
- a. Strongly disagree
 - b. Disagree
 - c. Neutral
 - d. Agree
 - e. Strongly agree

If strongly disagree or disagree, please explain:

9. The terms “mixed treatment comparison” is used unambiguously and consistently in the medical literature.
- a. Strongly disagree
 - b. Disagree
 - c. Neutral
 - d. Agree
 - e. Strongly agree

If strongly disagree or disagree, please explain:

10. The terms “Frequentist network meta-analysis” is used unambiguously and consistently in the medical literature.
- Strongly disagree
 - Disagree
 - Neutral
 - Agree
 - Strongly agree

If strongly disagree or disagree, please explain:

11. Synthesizing direct evidence only from sufficient head-to-head or randomized controlled trials takes precedence over analysis containing indirect evidence.
- Strongly disagree
 - Disagree
 - Neutral
 - Agree
 - Strongly agree
12. The combination of indirect and direct evidence adds valuable information that is not available from head-to-head comparisons.
- Strongly disagree
 - Disagree
 - Neutral
 - Agree
 - Strongly agree
13. The combination of indirect and direct evidence yields a more refined and precise estimate of the interventions directly compared.
- Strongly disagree
 - Disagree
 - Neutral
 - Agree
 - Strongly agree
14. The combination of indirect and direct evidence broadens the external validity of the analysis.
- Strongly disagree
 - Disagree
 - Neutral
 - Agree
 - Strongly agree

15. Where analysis of both direct and indirect comparisons is undertaken, each approach should be considered and reported separately.
- Strongly disagree
 - Disagree
 - Neutral
 - Agree
 - Strongly agree
16. When conducting a network meta-analysis, an investigator should consider restricting a search to the minimum number of interventions of interest.
- Strongly disagree
 - Disagree
 - Neutral
 - Agree
 - Strongly agree
17. When conducting a network meta-analysis, an investigator should consider including comparisons not of direct interest (e.g., placebo controls and therapies no longer used in current practice).
- Strongly disagree
 - Disagree
 - Neutral
 - Agree
 - Strongly agree
18. The more interventions that are included in a network meta-analysis, the greater uncertainty is reduced, precision is increased, and the ability to establish whether various sources of evidence agree with each other is enhanced.
- Strongly disagree
 - Disagree
 - Neutral
 - Agree
 - Strongly agree
19. Network meta-analyses should provide a graphical depiction of the evidence network.
- Strongly disagree
 - Disagree
 - Neutral
 - Agree
 - Strongly agree
20. The specific statistical code used should be available either as part of the manuscript, appendix/supplemental material, or available on an external website for the reader to freely access.
- Strongly disagree
 - Disagree

- c. Neutral
- d. Agree
- e. Strongly agree

21. Current guidance on how to conduct and report a network meta-analysis is sufficient.

- a. Strongly disagree
- b. Disagree
- c. Neutral
- d. Agree
- e. Strongly agree

If strongly disagree or disagree, please explain:

22. How much did the following play into your decision to conduct a Bayesian mixed treatment comparison meta-analysis?

The method allows for the ranking of Interventions according to the probability they are best.	Not at all	A little	Moderately	Quite a bit	Extremely
The method allows investigators to check and compare the fit of a model(s).	Not at all	A little	Moderately	Quite a bit	Extremely
The methods ability to handle multi-arm studies (those with more than 2 treatment groups).	Not at all	A little	Moderately	Quite a bit	Extremely
Frequency of use in previously published network meta-analyses.	Not at all	A little	Moderately	Quite a bit	Extremely
Ease of software implementation.	Not at all	A little	Moderately	Quite a bit	Extremely
The amount of methodological research supporting this method.	Not at all	A little	Moderately	Quite a bit	Extremely
The method's ability to combine trials reporting result in different formats, for example binomial data and summary log odds with variance (multi- or shared parameter models).	Not at all	A little	Moderately	Quite a bit	Extremely
Access to pre-built models.	Not at all	A little	Moderately	Quite a bit	Extremely
Requirement to specify priors which are often arbitrary.	Not at all	A little	Moderately	Quite a bit	Extremely
Collaborator(s) or your prior experience and/or expertise.	Not at all	A little	Moderately	Quite a bit	Extremely

23. We involved a researcher/collaborator in your project, solely due to their methodological expertise in Bayesian mixed treatment comparison meta-analysis?

- a. True
- b. False

24. Formal guidance was used to guide the conduction of your Bayesian mixed treatment comparison meta-analysis.

- a. True
- b. False

If true, please specify the guidance used and provide a complete reference if possible:

- 25. What are the three most significant barriers to conducting a Bayesian mixed treatment comparison meta-analysis?
 - a.
 - b.
 - c.

- 26. What are the three most significant strengths of conducting a Bayesian mixed treatment comparison meta-analysis?
 - a.
 - b.
 - c.

- 27. How was the code used in your analysis derived (e.g., built from scratch; used/adapted previously published/publically available code; used a wrapper such as BugsXLA or other to generate code, other source)?

- 28. How were your prior distributions chosen and why were these distributions chosen over others?

Appendix E. Excluded Studies

Table 4. Excluded studies at the full text level

Not a systematic review published in the English language from January 1, 2006 to July 30, 2011 (n=9)
Basu A, Meltzer HY, Dukic V. Estimating transitions between symptom severity states over time in schizophrenia: a Bayesian meta-analytic approach. <i>Stat Med</i> 2006 Sep 15;25(17):2886-2910.
Biondi-Zoccai GG, Lotrionte M, Abbate A, Valgimigli M, Testa L, Burzotta F, et al. Direct and indirect comparison meta-analysis demonstrates the superiority of sirolimus- versus paclitaxel-eluting stents across 5854 patients. <i>Int J Cardiol</i> 2007 Jan 2;114(1):104-105.
Biondi-Zoccai G, Lotrionte M, Moretti C, Agostoni P, Sillano D, Laudito A, et al. Percutaneous coronary intervention with everolimus-eluting stents (Xience V): systematic review and direct-indirect comparison meta-analyses with paclitaxel-eluting stents (Taxus) and sirolimus-eluting stents (Cypher). <i>Minerva Cardioangiol</i> 2008 Feb;56(1):55-65.
Buti J, Glennly AM, Worthington HV, Nieri M, Baccini M. Network meta-analysis of randomised controlled trials: direct and indirect treatment comparisons. <i>Eur j oral implantol</i> 2011;4(1):55-62.
Moayyedi P, Shelly S, Deeks JJ, Delaney B, Innes M, Forman D. Pharmacological interventions for non-ulcer dyspepsia. <i>Cochrane Database of Systematic Reviews</i> 2011;2.
Singh JA, Christensen R, Wells GA, Suarez-Almazor ME, Buchbinder R, Lopez-Olivo MA, et al. A network meta-analysis of randomized controlled trials of biologics for rheumatoid arthritis: a Cochrane overview. <i>CMAJ</i> 2009 Nov 24;181(11):787-796.
Trkulja V, Kolundzic R. Rivaroxaban vs dabigatran for thromboprophylaxis after joint-replacement surgery: exploratory indirect comparison based on meta-analysis of pivotal clinical trials. <i>Croat Med J</i> 2010 Apr 15;51(2):113-123.
Virgili G, Novielli N, Menchini F, Murro V, Giacomelli G. Pharmacological treatments for neovascular age-related macular degeneration: can mixed treatment comparison meta-analysis be useful? <i>Curr Drug Targets</i> 2011 Feb;12(2):212-220.
Wong MC, Clarkson J, Glennly AM, Lo EC, Marinho VC, Tsang BW, et al. Cochrane reviews on the benefits/risks of fluoride toothpastes. <i>J Dent Res</i> 2011 May;90(5):573-579.
Did not conduct meta-analysis of clinical effectiveness using randomized controlled trials (n=8)
Jefferson T, Di Pietrantonj C, AlAnsary LA, Ferroni E, Thorning S, Thomas RE. Vaccines for preventing influenza in the elderly. <i>Cochrane Database of Systematic Reviews</i> 2010;6.
Jefferson T, Rivetti A, Harnden A, Di Pietrantonj C, Demicheli V. Vaccines for preventing influenza in healthy children. <i>Cochrane Database of Systematic Reviews</i> 2009;1.
Medical Advisory Secretariat. Ontario Ministry of Health and Long-Term Care (MAS). Artificial disc replacement for lumbar and cervical degenerative disc disease- update: an evidence-based analysis (Structured abstract). <i>Health Technology Assessment Database</i> 2011 Issue 3, John Wiley & Sons, Ltd. Chichester:UK. Dson: ST.
Miller, J. Chan, BKS. Nelson, H. Hormone replacement therapy and risk of venous thromboembolism (Structured abstract). <i>Health Technology Assessment Database</i> 2011 Issue 3, John Wiley & Sons, Ltd. Chichester:UK. Dson: ST.
Paravastu SCV, Mendonca D, Da Silva A. Beta blockers for peripheral arterial disease. <i>Cochrane Database of Systematic Reviews</i> 2010;3.
Richy FF, Banerjee S, Brabant Y, Helmers S. Levetiracetam extended release and levetiracetam immediate release as adjunctive treatment for partial-onset seizures: an indirect comparison of treatment-emergent adverse events using meta-analytic techniques. <i>Epilepsy Behav</i> 2009 Oct;16(2):240-245.
Takeda AL, Colquitt J, Clegg AJ, Jones J. Pegaptanib and ranibizumab for neovascular age-related macular degeneration: a systematic review. <i>Br J Ophthalmol</i> 2007 Sep;91(9):1177-1182.
van Till JO, van Ruler O, Lamme B, Weber RJ, Reitsma JB, Boermeester MA. Single-drug therapy or selective decontamination of the digestive tract as antifungal prophylaxis in critically ill patients: a systematic review. <i>Crit Care</i> 2007;11(6):R126.
Did not conduct an indirect comparison of more than two arms (n=113)
Abba K, Ramaratnam S, Ranganathan NL. Anthelmintics for people with neurocysticercosis. <i>Cochrane Database of Systematic Reviews</i> 2010;6.
Avenell A, Gillespie WJ, Gillespie LD, O'Connell D. Vitamin D and vitamin D analogues for preventing fractures associated with involutional and post-menopausal osteoporosis. <i>Cochrane Database of Systematic Reviews</i> 2009;1.

BlueCross BlueShield A. Metal-on-metal total hip resurfacing (Structured abstract). Health Technology Assessment Database 2011 Issue 3, John Wiley & Sons, Ltd. Chichester:UK. Dson: ST.

Bohlius J, Herbst C, Reiser M, Schwarzer G, Engert A. Granulopoiesis-stimulating factors to prevent adverse effects in the treatment of malignant lymphoma. Cochrane Database of Systematic Reviews 2009;1.

Bolen S, Feldman L, Vassy J, Wilson L, Yeh HC, Marinopoulos S, et al. Systematic review: comparative effectiveness and safety of oral medications for type 2 diabetes mellitus. Ann Intern Med 2007 Sep 18;147(6):386-399.

Brownfoot FC, Crowther CA, Middleton P. Different corticosteroids and regimens for accelerating fetal lung maturation for women at risk of preterm birth. Cochrane Database of Systematic Reviews 2009;1.

Brozek J, Akl EA, Jaeschke R, Terrenato I, Cilenti V, Cazzola M, et al. Long-acting beta2-agonists for chronic obstructive pulmonary disease: serious adverse events. Cochrane Database of Systematic Reviews 2009;1.

Bunn F, Trivedi D, Ashraf S. Colloid solutions for fluid resuscitation. Cochrane Database of Systematic Reviews 2011;5.

Cahill K, Lancaster T, Green N. Stage-based interventions for smoking cessation. Cochrane Database of Systematic Reviews 2011;1.

Caldwell D, Hunot V, Moore HMT, Davies P, Jones H, Lewis G, et al. Behavioural therapies versus treatment as usual for depression. Cochrane Database of Systematic Reviews 2010;6.

Castells X, RamosQuiroga AJ, Bosch R, Nogueira M, Casas M. Amphetamines for Attention Deficit Hyperactivity Disorder (ADHD) in adults. Cochrane Database of Systematic Reviews 2011;4.

Chandrasekaran B, Arumugam A, Davis F, Kumaran D S, Chandrashaarma B, Khundrakpam C, et al. Resistance exercise training for hypertension. Cochrane Database of Systematic Reviews 2010;11.

Churchill R, Caldwell D, Moore HMT, Davies P, Jones H, Lewis G, et al. Behavioural therapies versus other psychological therapies for depression. Cochrane Database of Systematic Reviews 2010;6.

Churchill R, Davies P, Caldwell D, Moore HMT, Jones H, Lewis G, et al. Interpersonal, cognitive analytic and other integrative therapies versus treatment as usual for depression. Cochrane Database of Systematic Reviews 2010;6.

Churchill R, Davies P, Caldwell D, Moore HMT, Jones H, Lewis G, et al. Humanistic therapies versus other psychological therapies for depression. Cochrane Database of Systematic Reviews 2010;6.

Churchill R, Moore HMT, Caldwell D, Davies P, Jones H, Furukawa TA, et al. Cognitive behavioural therapies versus other psychological therapies for depression. Cochrane Database of Systematic Reviews 2010;6.

Churchill R, Moore HMT, Davies P, Caldwell D, Jones H, Lewis G, et al. Psychodynamic therapies versus other psychological therapies for depression. Cochrane Database of Systematic Reviews 2010;6.

Churchill R, Moore HMT, Davies P, Caldwell D, Jones H, Lewis G, et al. Mindfulness-based 'third wave' cognitive and behavioural therapies versus treatment as usual for depression. Cochrane Database of Systematic Reviews 2010;6.

Cipriani A, La Ferla T, Furukawa TA, Signoretti A, Nakagawa A, Churchill R, et al. Sertraline versus other antidepressive agents for depression. Cochrane Database of Systematic Reviews 2011;2.

Cipriani A, Santilli C, Furukawa TA, Signoretti A, Nakagawa A, McGuire H, et al. Escitalopram versus other antidepressive agents for depression. Cochrane Database of Systematic Reviews 2009;1.

Clarkson JE, Worthington HV, Eden TOB. Interventions for preventing oral candidiasis for patients with cancer receiving treatment. Cochrane Database of Systematic Reviews 2009;1.

Coppin C, Porzolt F, Autenrieth M, Kumpf J, Coldman A, Wilt T. Immunotherapy for advanced renal cell cancer. Cochrane Database of Systematic Reviews 2009;1.

Cross NB, Webster AC, Masson P, O'Connell PJ, Craig JC. Antihypertensive treatment for kidney transplant recipients. Cochrane Database of Systematic Reviews 2010;2.

Davies P, Hunot V, Moore HMT, Caldwell D, Jones H, Lewis G, et al. Humanistic therapies versus treatment as usual for depression. Cochrane Database of Systematic Reviews 2010;6.

de Peuter OR, Lussana F, Peters RJ, Buller HR, Kamphuisen PW. A systematic review of selective and non-selective beta blockers for prevention of vascular events in patients with acute coronary syndrome or heart failure. Neth J Med 2009 Oct;67(9):284-294.

Deacon SA, Glenny A, Deery C, Robinson PG, Heanue M, Walmsley DA, et al. Different powered toothbrushes for plaque control and gingival health. Cochrane Database of Systematic Reviews 2011;4.

Delahoy P, Thompson S, Marschner IC. Pregabalin versus gabapentin in partial epilepsy: a meta-analysis of dose-response relationships. BMC Neurol 2010;10:104.

Derry S, Faura C, Edwards J, McQuay HJ, Moore AR. Single dose dipyron for acute postoperative pain. Cochrane

Database of Systematic Reviews 2011;5.

Derry S, Moore AR, McQuay HJ. Single dose oral codeine, as a single agent, for acute postoperative pain in adults. Cochrane Database of Systematic Reviews 2010;8.

Di Nisio M, Wichers IM, Middeldorp S. Treatment for superficial thrombophlebitis of the leg. Cochrane Database of Systematic Reviews 2011;2.

Eisenberg MJ, Filion KB, Yavin D, Belisle P, Mottillo S, Joseph L, et al. Pharmacotherapies for smoking cessation: a meta-analysis of randomized controlled trials. CMAJ 2008 Jul 15;179(2):135-144.

Faraone SV, Glatt SJ. A comparison of the efficacy of medications for adult attention-deficit/hyperactivity disorder using meta-analysis of effect sizes. J Clin Psychiatry 2010 Jun;71(6):754-763.

Farion KJ, Russell KF, Osmond MH, Hartling L, Klassen TP, Durec T, et al. Tissue adhesives for traumatic lacerations in children and adults. Cochrane Database of Systematic Reviews 2011;4.

Filion KB, El Khoury F, Bielinski M, Schiller I, Dendukuri N, Brophy JM. Omega-3 fatty acids in high-risk cardiovascular patients: a meta-analysis of randomized controlled trials. BMC Cardiovasc Disord 2010;10:24.

Galaal K, Godfrey K, Naik R, Kucukmetin A, Bryant A. Adjuvant radiotherapy and/or chemotherapy after surgery for uterine carcinosarcoma. Cochrane Database of Systematic Reviews 2011;4.

Gartlehner G, Gaynes BN, Hansen RA, Thieda P, DeVeauugh-Geiss A, Krebs EE, et al. Comparative benefits and harms of second-generation antidepressants: background paper for the American College of Physicians. Ann Intern Med 2008 Nov 18;149(10):734-750.

Gillespie LD, Robertson CM, Gillespie WJ, Lamb SE, Gates S, Cumming RG, et al. Interventions for preventing falls in older people living in the community. Cochrane Database of Systematic Reviews 2010;10.

Gotzsche PC, Johansen HK. House dust mite control measures for asthma. Cochrane Database of Systematic Reviews 2009;1.

Guaiana G, Barbui C, Cipriani A. Hydroxyzine for generalised anxiety disorder. Cochrane Database of Systematic Reviews 2009;1.

Gurusamy KS, Kumar Y, Davidson BR. Methods of preventing bacterial sepsis and wound complications for liver transplantation. Cochrane Database of Systematic Reviews 2009;1.

Henry DA, Carless PA, Moxey AJ, O'Connell D, Stokes BJ, Fergusson DA, et al. Anti-fibrinolytic use for minimising perioperative allogeneic blood transfusion. Cochrane Database of Systematic Reviews 2011;5.

Hodson EM, Craig JC, Strippoli FMG, Webster AC. Antiviral medications for preventing cytomegalovirus disease in solid organ transplant recipients. Cochrane Database of Systematic Reviews 2010;4.

Hooper L, Summerbell CD, Thompson R, Sills D, Roberts FG, Moore H, et al. Reduced or modified dietary fat for preventing cardiovascular disease. Cochrane Database of Systematic Reviews 2011;7.

Hunot V, Moore HMT, Caldwell D, Davies P, Jones H, Furukawa TA, et al. Cognitive behavioural therapies versus treatment as usual for depression. Cochrane Database of Systematic Reviews 2010;6.

Hunot V, Moore HMT, Caldwell D, Davies P, Jones H, Lewis G, et al. Mindfulness-based 'third wave' cognitive and behavioural therapies versus other psychological therapies for depression. Cochrane Database of Systematic Reviews 2010;6.

Hunot V, Moore HMT, Caldwell D, Davies P, Jones H, Lewis G, et al. Interpersonal, cognitive analytic and other integrative therapies versus other psychological therapies for depression. Cochrane Database of Systematic Reviews 2010;6.

Iorio A, Marchesini E, Awad T, Gluud LL. Antiviral treatment for chronic hepatitis C in patients with human immunodeficiency virus. Cochrane Database of Systematic Reviews 2010;5.

Ipser JC, Stein DJ, Hawkrigde S, Hoppe L. Pharmacotherapy for anxiety disorders in children and adolescents. Cochrane Database of Systematic Reviews 2010;6.

Itchaki G, Gaftanovili A, Lahav M, Vidal L, Raanani P, Shpilberg O, et al. Anthracyclines-containing regimens for treatment of follicular lymphoma in adults. Cochrane Database of Systematic Reviews 2010;8.

Jackson PPR, Aarabi M, Wallis E. Aspirin for primary prevention of coronary heart disease. Cochrane Database of Systematic Reviews 2009;1.

Jefferson T, Jones MA, Doshi P, Del Mar CB, Heneghan CJ, Hama R, et al. Neuraminidase inhibitors for preventing and treating influenza in healthy adults and children - a review of clinical study reports. Cochrane Database of Systematic Reviews 2011;4.

Johansson K, Sundstrom J, Neovius K, Rossner S, Neovius M. Long-term changes in blood pressure following orlistat and sibutramine treatment: a meta-analysis. Obes Rev 2010 Nov;11(11):777-791.

Kaizar EE, Greenhouse JB, Seltman H, Kelleher K. Do antidepressants cause suicidality in children? A Bayesian

meta-analysis. *Clin trials* 2006 discussion 91-8;3(2):73-90.

Kamphuisen PW, Agnelli G. What is the optimal pharmacological prophylaxis for the prevention of deep-vein thrombosis and pulmonary embolism in patients with acute ischemic stroke? *Thromb Res* 2007;119(3):265-274.

Kulisevsky J, Pagonabarraga J. Tolerability and safety of ropinirole versus other dopamine agonists and levodopa in the treatment of Parkinson's disease: meta-analysis of randomized controlled trials. *Drug Saf* 2010 Feb 1;33(2):147-161.

Lancaster T, Stead LF. Individual behavioural counselling for smoking cessation. *Cochrane Database of Systematic Reviews* 2009;1.

Law S, Derry S, Moore AR. Triptans for acute cluster headache. *Cochrane Database of Systematic Reviews* 2011;1.

Lin Y, Stanworth S, Birchall J, Doree C, Hyde C. Recombinant factor VIIa for the prevention and treatment of bleeding in patients without haemophilia. *Cochrane Database of Systematic Reviews* 2011;4.

Logman JF, Stephens J, Heeg B, Haider S, Cappelleri J, Nathwani D, et al. Comparative effectiveness of antibiotics for the treatment of MRSA complicated skin and soft tissue infections. *Curr Med Res Opin* 2010 Jul;26(7):1565-1578.

Maas T, Kaper J, Sheikh A, Knottnerus JA, Wesseling G, Dompeling E, et al. Mono and multifaceted inhalant and/or food allergen reduction interventions for preventing asthma in children at high risk of developing asthma. *Cochrane Database Syst Rev* 2009(3):006480.

Maas T, Kaper J, Sheikh A, Knottnerus AJ, Wesseling G, Dompeling E, et al. Mono and multifaceted inhalant and/or food allergen reduction interventions for preventing asthma in children at high risk of developing asthma. *Cochrane Database of Systematic Reviews* 2011;4.

Manheimer E, Cheng K, Linde K, Lao L, Yoo J, Wieland S, et al. Acupuncture for peripheral joint osteoarthritis. *Cochrane Database of Systematic Reviews* 2010;5.

Massey T, Derry S, Moore AR, McQuay HJ. Topical NSAIDs for acute pain in adults. *Cochrane Database of Systematic Reviews* 2010;6.

McQuay HJ, Moore RA. Dose-response in direct comparisons of different doses of aspirin, ibuprofen and paracetamol (acetaminophen) in analgesic studies. *Br J Clin Pharmacol* 2007 Mar;63(3):271-278.

Moore HMT, Hunot V, Davies P, Caldwell D, Jones H, Lewis G, et al. Psychodynamic therapies versus treatment as usual for depression. *Cochrane Database of Systematic Reviews* 2010;6.

Mottillo S, Filion KB, Belisle P, Joseph L, Gervais A, O'Loughlin J, et al. Behavioural interventions for smoking cessation: a meta-analysis of randomized controlled trials. *Eur Heart J* 2009 Mar;30(6):718-730.

Nakagawa A, Watanabe N, Omori IM, Barbui C, Cipriani A, McGuire H, et al. Milnacipran versus other antidepressive agents for depression. *Cochrane Database of Systematic Reviews* 2009;1.

Nguyen ND, Eisman JA, Nguyen TV. Anti-hip fracture efficacy of bisphosphonates: a Bayesian analysis of clinical trials. *J Bone Miner Res* 2006 Feb;21(2):340-349.

Omori IM, Watanabe N, Nakagawa A, Cipriani A, Barbui C, McGuire H, et al. Fluvoxamine versus other anti-depressive agents for depression. *Cochrane Database of Systematic Reviews* 2010;6.

Pavlikis N, Schmidt RL, Stockler MR. Bisphosphonates for breast cancer. *Cochrane Database of Systematic Reviews* 2009;1.

Pepas L, Kaushik S, Bryant A, Nordin A, Dickinson HO. Medical interventions for high grade vulval intraepithelial neoplasia. *Cochrane Database of Systematic Reviews* 2011;4.

Perletti G, Marras E, Wagenlehner MEF, Magri V. Antimicrobial therapy for chronic bacterial prostatitis. *Cochrane Database of Systematic Reviews* 2011;5.

Phillips RS, Gopaul S, Gibson F, Houghton E, Craig JV, Light K, et al. Antiemetic medication for prevention and treatment of chemotherapy induced nausea and vomiting in childhood. *Cochrane Database of Systematic Reviews* 2011;2.

Piscione F, Piccolo R, Cassese S, Galasso G, De Rosa R, D'Andrea C, et al. Effect of drug-eluting stents in patients with acute ST-segment elevation myocardial infarction undergoing percutaneous coronary intervention: a meta-analysis of randomised trials and an adjusted indirect comparison. *EuroIntervention* 2010 Feb;5(7):853-860.

Playford EG, Webster AC, Sorrell TC, Craig JC. Systematic review and meta-analysis of antifungal agents for preventing fungal infections in liver transplant recipients. *Eur J Clin Microbiol Infect Dis* 2006 Sep;25(9):549-561.

Rada G, Capurro D, Pantoja T, Corbalan J, Moreno G, Letelier LM, et al. Non-hormonal interventions for hot flushes in women with a history of breast cancer. *Cochrane Database of Systematic Reviews* 2010;8.

Rayner L, Price A, Evans A, Valsraj K, Higginson IJ, Hotopf M. Antidepressants for depression in physically ill people. *Cochrane Database of Systematic Reviews* 2010;6.

Razavian M, Di Micco L, Palmer SC, Craig JC, Perkovic V, Zoungas S, et al. Antiplatelet agents for chronic kidney

disease. *Cochrane Database of Systematic Reviews* 2010;11.

Reinhart M, Musini VM, Salzwedel DM, Dormuth C, Wright JM. First-line diuretics versus other classes of antihypertensive drugs for hypertension. *Cochrane Database of Systematic Reviews* 2011;4.

Revez L, Cardona FA, Ospina GE. Antibiotics for acute laryngitis in adults. *Cochrane Database of Systematic Reviews* 2011;5.

Rigotti N, Munafo' MR, Stead LF. Interventions for smoking cessation in hospitalised patients. *Cochrane Database of Systematic Reviews* 2009;1.

Rossi A, Garassino MC, Cinquini M, Sburlati P, Di Maio M, Farina G, et al. Maintenance or consolidation therapy in small-cell lung cancer: a systematic review and meta-analysis. *Lung Cancer* 2010 Nov;70(2):119-128.

Rowe BH, Spooner C, Ducharme F, Bretzlaff J, Bota G. Corticosteroids for preventing relapse following acute exacerbations of asthma. *Cochrane Database of Systematic Reviews* 2009;1.

Sandercock AGP, Counsell C, Kamal AK. Anticoagulants for acute ischaemic stroke. *Cochrane Database of Systematic Reviews* 2010;8.

Sandercock AGP, Gibson LM, Liu M. Anticoagulants for preventing recurrence following presumed non-cardioembolic ischaemic stroke or transient ischaemic attack. *Cochrane Database of Systematic Reviews* 2010;10.

Sandercock PAG, Counsell C, Gubitz GJ, Tseng M. Antiplatelet therapy for acute ischaemic stroke. *Cochrane Database of Systematic Reviews* 2009;1.

Sasse EC, Sasse AD, Brandalise SR, Clark OAC, Richards S. Colony-stimulating factors for prevention of myelosuppressive therapy-induced febrile neutropenia in children with acute lymphoblastic leukaemia. *Cochrane Database of Systematic Reviews* 2009;1.

Scholz H, Trenkwalder C, Kohnen R, Kriston L, Riemann D, Hornyak M. Dopamine agonists for the treatment of restless legs syndrome. *Cochrane Database of Systematic Reviews* 2011;5.

Shanbhag S, Cairns M, Cruickshank M, Dickinson HO, Gokul S, Parkin D. Effectiveness of different treatment modalities for the management of adult onset granulosa cell tumours of the ovary (primary and recurrent). *Cochrane Database of Systematic Reviews* 2009;1.

Siegfried N, van der Merwe L, Brocklehurst P, Sint TT. Antiretrovirals for reducing the risk of mother-to-child transmission of HIV infection. *Cochrane Database of Systematic Reviews* 2011;7.

Simpson SA, Lewis R, van der Voort J, Butler CC. Oral or topical nasal steroids for hearing loss associated with otitis media with effusion in children. *Cochrane Database of Systematic Reviews* 2011;5.

Small LN, Lau J, Snyderman DR. Preventing post-organ transplantation cytomegalovirus disease with ganciclovir: a meta-analysis comparing prophylactic and preemptive therapies. *Clin Infect Dis* 2006 Oct 1;43(7):869-880.

Solomon R, Dumouchel W. Contrast media and nephropathy: findings from systematic analysis and Food and Drug Administration reports of adverse effects. *Invest Radiol* 2006 Aug;41(8):651-660.

Soomro GM, Altman DG, Rajagopal S, Oakley Browne M. Selective serotonin re-uptake inhibitors (SSRIs) versus placebo for obsessive compulsive disorder (OCD). *Cochrane Database of Systematic Reviews* 2009;1.

Spaulding A, Rutherford GW, Siegfried N. Stavudine or zidovudine in three-drug combination therapy for initial treatment of HIV infection in antiretroviral-naïve individuals. *Cochrane Database of Systematic Reviews* 2011;4.

Stowe R, Ives N, Clarke CE, Deane K, van H, Wheatley K, et al. Evaluation of the efficacy and safety of adjuvant treatment to levodopa therapy in Parkinson's disease patients with motor complications. *Cochrane Database Syst Rev* 2010(7):007166.

Stowe R, Ives N, Clarke CE, van Hilten, Ferreira J, Hawker RJ, et al. Dopamine agonist therapy in early Parkinson's disease. *Cochrane Database of Systematic Reviews* 2011;1.

Strassels S, McNicol ED, Francia M, Belle D, Pepin J, Jackson KC. Meperidine (pethidine) for treatment of postoperative pain in adults. *Cochrane Database of Systematic Reviews* 2009;1.

Strippoli FMG, Bonifati C, Craig ME, Navaneethan SD, Craig JC. Angiotensin converting enzyme inhibitors and angiotensin II receptor antagonists for preventing the progression of diabetic kidney disease. *Cochrane Database of Systematic Reviews* 2011;1.

Sudlow CLM, Mason G, Maurice JB, Wedderburn CJ, Hankey GJ. Thienopyridine derivatives versus aspirin for preventing stroke and other serious vascular events in high vascular risk patients. *Cochrane Database of Systematic Reviews* 2010;2.

Tacklind J, Fink HA, MacDonald R, Rutks I, Wilt TJ. Finasteride for benign prostatic hyperplasia. *Cochrane Database of Systematic Reviews* 2011;2.

Tan A, Xia N, Gao F, Mo Z, Cao Y. Angiogenesis-inhibitors for metastatic thyroid cancer. *Cochrane Database of Systematic Reviews* 2010;6.

Tang W, Wong GWK, Perez MI, Wright JM. First-line drugs inhibiting the renin angiotensin system versus other first-line antihypertensive drug classes for hypertension. *Cochrane Database of Systematic Reviews* 2009;1.

Thomas RE, Russell M, Lorenzetti D. Interventions to increase influenza vaccination rates of those 60 years and older in the community. *Cochrane Database of Systematic Reviews* 2011;7.

Thomsen T, Villebro N, Moller AM. Interventions for preoperative smoking cessation. *Cochrane Database Syst Rev* 2010(7):002294.

Vardi M, Jacobson E, Nini A, Bitterman H. Intermediate acting versus long acting insulin for type 1 diabetes mellitus. *Cochrane Database of Systematic Reviews* 2011;1.

Vardi M, Zittan E, Bitterman H. Subcutaneous unfractionated heparin for the initial treatment of venous thromboembolism. *Cochrane Database of Systematic Reviews* 2010;2.

Visser K, van der Heijde D. Optimal dosage and route of administration of methotrexate in rheumatoid arthritis: a systematic review of the literature. *Ann Rheum Dis* 2009 Jul;68(7):1094-1099.

Wardlaw JM, Murray V, Berge E, del Zoppo GJ. Thrombolysis for acute ischaemic stroke. *Cochrane Database of Systematic Reviews* 2010;2.

Watanabe N, Churchill R, Hunot V, Furukawa TA. Psychotherapy for depression in children and adolescents. *Cochrane Database of Systematic Reviews* 2009;1.

Webster AC, Lee VWS, Chapman JR, Craig JC. Target of rapamycin inhibitors (TOR-I; sirolimus and everolimus) for primary immunosuppression in kidney transplant recipients. *Cochrane Database of Systematic Reviews* 2009;1.

Webster AC, Ruster LP, McGee R, Matheson SL, Higgins GY, Willis NS, et al. Interleukin 2 receptor antagonists for kidney transplant recipients. *Cochrane Database of Systematic Reviews* 2010;5.

Yuan Y, Ford AC, Khan KJ, Forman D, Oderda G, Calvet X, et al. Optimum duration of regimens for *Helicobacter pylori* eradication. *Cochrane Database of Systematic Reviews* 2009;1.

Did not use either Bayesian or Frequentist mixed treatment comparison method (n=41)

Abdullah AK, Khan S. Relative oral corticosteroid-sparing effect of 7 inhaled corticosteroids in chronic asthma: a meta-analysis. *Ann Allergy Asthma Immunol* 2008 Jul;101(1):74-81.

Abou-Setta AM. Firm embryo transfer catheters for assisted reproduction: a systematic review and meta-analysis using direct and adjusted indirect comparisons. *Reprod Biomed Online* 2006 Feb;12(2):191-198.

Abou-Setta AM. What is the best site for embryo deposition? A systematic review and meta-analysis using direct and adjusted indirect comparisons. *Reprod Biomed Online* 2007 May;14(5):611-619.

Berner MM, Kriston L, Harms A. Efficacy of PDE-5-inhibitors for erectile dysfunction. A comparative meta-analysis of fixed-dose regimen randomized controlled trials administering the International Index of Erectile Function in broad-spectrum populations. *Int J Impot Res* 2006 May-Jun;18(3):229-235.

Bottomley JM, Taylor RS, Rytov J. The effectiveness of two-compound formulation calcipotriol and betamethasone dipropionate gel in the treatment of moderately severe scalp psoriasis: a systematic review of direct and indirect evidence. *Curr Med Res Opin* 2011 Jan;27(1):251-268.

Buscemi N, Vandermeer B, Friesen C, Bialy L, Tubman M, Ospina M, et al. The efficacy and safety of drug treatments for chronic insomnia in adults: a meta-analysis of RCTs. *J Gen Intern Med* 2007 Sep;22(9):1335-1350.

Chen MHJ, Heran BS, Perez MI, Wright JM. Blood pressure lowering efficacy of beta-blockers as second-line therapy for primary hypertension. *Cochrane Database of Systematic Reviews* 2010;5.

Chen JMH, Heran BS, Wright JM. Blood pressure lowering efficacy of diuretics as second-line therapy for primary hypertension. *Cochrane Database of Systematic Reviews* 2010;2.

Chou R, Carson S, Chan BK. Gabapentin versus tricyclic antidepressants for diabetic neuropathy and post-herpetic neuralgia: discrepancies between direct and indirect meta-analyses of randomized controlled trials. *J Gen Intern Med* 2009 Feb;24(2):178-188.

Fakhoury W, Lockhart I, Kotchie RW, Aagren M, LeReun C. Indirect comparison of once daily insulin detemir and glargine in reducing weight gain and hypoglycaemic episodes when administered in addition to conventional oral anti-diabetic therapy in patients with type-2 diabetes. *Pharmacology* 2008;82(2):156-163.

Gartlehner G, Hansen RA, Jonas BL, Thieda P, Lohr KN. The comparative efficacy and safety of biologics for the treatment of rheumatoid arthritis: a systematic review and metaanalysis. *J Rheumatol* 2006 Dec;33(12):2398-2408.

Germani G, Pleguezuelo M, Gurusamy K, Meyer T, Isgro G, Burroughs AK. Clinical outcomes of radiofrequency ablation, percutaneous alcohol and acetic acid injection for hepatocellular carcinoma: a meta-analysis. *J Hepatol* 2010 Mar;52(3):380-388.

Hansen RA, Gartlehner G, Webb AP, Morgan LC, Moore CG, Jonas DE. Efficacy and safety of donepezil, galantamine, and rivastigmine for the treatment of Alzheimer's disease: a systematic review and meta-analysis. *Clin*

Interv Aging 2008;3(2):211-225.

Hauser W, Petzke F, Sommer C. Comparative efficacy and harms of duloxetine, milnacipran, and pregabalin in fibromyalgia syndrome. *J Pain* 2010 Jun;11(6):505-521.

Heran BS, Chen MHJ, Wang JJ, Wright JM. Blood pressure lowering efficacy of potassium-sparing diuretics (that block the epithelial sodium channel) for primary hypertension. *Cochrane Database of Systematic Reviews* 2010;5.

Heran BS, Galm BP, Wright JM. Blood pressure lowering efficacy of alpha blockers for primary hypertension. *Cochrane Database of Systematic Reviews* 2010;2.

Heran BS, Wong MMY, Heran IK, Wright JM. Blood pressure lowering efficacy of angiotensin converting enzyme (ACE) inhibitors for primary hypertension. *Cochrane Database of Systematic Reviews* 2009;1.

Heran BS, Wong MMY, Heran IK, Wright JM. Blood pressure lowering efficacy of angiotensin receptor blockers for primary hypertension. *Cochrane Database of Systematic Reviews* 2009;1.

Hofmeyr GJ, Gulmezoglu AM, Novikova N, Linder V, Ferreira S, Piaggio G. Misoprostol to prevent and treat postpartum haemorrhage: a systematic review and meta-analysis of maternal deaths and dose-related effects. *Bull World Health Organ* 2009 Sep;87(9):666-677.

Holley A, King C, Moores L, Jackson JL, Shorr A. The role of twice versus three times daily heparin dosing for thromboembolism prophylaxis in general medical patients admitted to the hospital. *Cochrane Database of Systematic Reviews* 2009;1.

Jalota L, Kalira V, George E, et al. Prevention of pain on injection of propofol: systematic review and meta-analysis. *BMJ* 2011;342:1110.

Kabra SK, Lodha R, Pandey RM. Antibiotics for community-acquired pneumonia in children. *Cochrane Database of Systematic Reviews* 2010;6.

Kalil AC, Mindru C, Florescu DF. Effectiveness of valganciclovir 900 mg versus 450 mg for cytomegalovirus prophylaxis in transplantation: direct and indirect treatment comparison meta-analysis. *Clin Infect Dis* 2011 Feb;52(3):313-321.

Kotze A, Scally A, Howell S. Efficacy and safety of different techniques of paravertebral block for analgesia after thoracotomy: a systematic review and metaregression. *Br J Anaesth* 2009 Nov;103(5):626-636.

Kumar S, Ahuja V, Vishnubhatla S, Prasad K, Kumar A. Curcumin for maintenance of remission in ulcerative colitis. *Cochrane Database of Systematic Reviews* 2011;2.

Kumar A, Hozo I, Wheatley K, Djulbegovic B. Thalidomide versus bortezomib based regimens as first-line therapy for patients with multiple myeloma: a systematic review. *Am J Hematol* 2011 Jan;86(1):18-24.

Latthe PM, Singh P, Foon R, Toozs-Hobson P. Two routes of transobturator tape procedures in stress urinary incontinence: a meta-analysis with direct and indirect comparison of randomized trials. *BJU Int* 2010 Jul;106(1):68-76.

Lim E, Harris G, Patel A, Adachi I, Edmonds L, Song F. Preoperative versus postoperative chemotherapy in patients with resectable non-small cell lung cancer: systematic review and indirect comparison meta-analysis of randomized trials. *J Thorac Oncol* 2009 Nov;4(11):1380-1388.

Lim E, Harris G, Patel A, Adachi I, Edmonds L, Song F. Preoperative versus postoperative chemotherapy in patients with resectable non-small cell lung cancer: Systematic review and indirect comparison meta-analysis of randomized trials [abstract no.7546]. *Journal of Clinical Oncology: ASCO annual meeting proceedings 2008*;26(15S part 1):408.

Loke YK, Kwok CS. Dabigatran and rivaroxaban for prevention of venous thromboembolism—systematic review and adjusted indirect comparison. *J Clin Pharm Ther* 2011 Feb;36(1):111-124.

Mills EJ, Rachlis B, O'Regan C, Thabane L, Perri D. Metastatic renal cell cancer treatments: an indirect comparison meta-analysis. *BMC Cancer* 2009;9:34.

Peterson K, McDonagh MS, Fu R. Comparative benefits and harms of competing medications for adults with attention-deficit hyperactivity disorder: a systematic review and indirect comparison meta-analysis. *Psychopharmacology (Berl)* 2008 Mar;197(1):1-11.

Piccini JP, Hasselblad V, Peterson ED, Washam JB, Califf RM, Kong DF. Comparative efficacy of dronedarone and amiodarone for the maintenance of sinus rhythm in patients with atrial fibrillation. *J Am Coll Cardiol* 2009 Sep 15;54(12):1089-1095.

Riemsma R, Forbes CA, Kessels A, Lykopoulos K, Amonkar MM, Rea DW, et al. Systematic review of aromatase inhibitors in the first-line treatment for hormone sensitive advanced or metastatic breast cancer. *Breast Cancer Res Treat* 2010 Aug;123(1):9-24.

Riemsma R, Simons JP, Bashir Z, Gooch CL, Kleijnen J. Systematic Review of topotecan (Hycamtin) in relapsed small cell lung cancer. *BMC Cancer* 2010;10:436.

Rostom A, Dube C, Jolicoeur E, Joyce JL, Muir K. Adverse gastrointestinal effects of COX-2 Inhibitors for

inflammatory diseases. Cochrane Database of Systematic Reviews 2009;1.

Salliot C, Finckh A, Katchamart W, Lu Y, Sun Y, Bombardier C, et al. Indirect comparisons of the efficacy of biological antirheumatic agents in rheumatoid arthritis in patients with an inadequate response to conventional disease-modifying antirheumatic drugs or to an anti-tumour necrosis factor agent: a meta-analysis. *Ann Rheum Dis* 2011 Feb;70(2):266-271.

Stettler C, Allemann S, Egger M, Windecker S, Meier B, Diem P. Efficacy of drug eluting stents in patients with and without diabetes mellitus: indirect comparison of controlled trials. *Heart* 2006 May;92(5):650-657.

Sultana A, Ghaneh P, Cunningham D, Starling N, Neoptolemos JP, Smith CT. Gemcitabine based combination chemotherapy in advanced pancreatic cancer-indirect comparison. *BMC Cancer* 2008;8:192.

Wilhelmus RK. Antiviral treatment and other therapeutic interventions for herpes simplex virus epithelial keratitis. *Cochrane Database of Systematic Reviews* 2011;7.

Wu P, Wilson K, Dimoulas P, Mills EJ. Effectiveness of smoking cessation therapies: a systematic review and meta-analysis. *BMC Public Health* 2006;6:300.

Zhou Z, Rahme E, Pilote L. Are statins created equal? Evidence from randomized trials of pravastatin, simvastatin, and atorvastatin for cardiovascular disease prevention. *Am Heart J* 2006 Feb;151(2):273-281.

Bayesian MTC but other than closed loop (n=13)

Bergman GJ, Hochberg MC, Boers M, Wintfeld N, Kielhorn A, Jansen JP. Indirect comparison of tocilizumab and other biologic agents in patients with rheumatoid arthritis and inadequate response to disease-modifying antirheumatic drugs. *Semin Arthritis Rheum* 2010;39(6):425-441

Edwards SJ, Clarke MJ, Wordsworth S, Welton NJ. Carbapenems versus other beta-lactams in the treatment of hospitalised patients with infection: a mixed treatment comparison. *Curr Med Res Opin* 2009;25(1):251-261

Freemantle N, Tharmanathan P, Herbrecht R. Systematic review and mixed treatment comparison of randomized evidence for empirical, pre-emptive and directed treatment strategies for invasive mould disease. *J Antimicrob Chemother* 2011 66(Suppl 1):25-35.

Jansen JP. Self-monitoring of glucose in type 2 diabetes mellitus: a Bayesian meta-analysis of direct and indirect comparisons. *Curr Med Res Opin* 2006;22(4):671-681.

Jansen JP, Bergman GJ, Huels J, Olson M. Prevention of vertebral fractures in osteoporosis: mixed treatment comparison of bisphosphonate therapies. *Curr Med Res Opin* 2009;25(8):1861-1868.

Jansen JP, Bergman GJ, Huels J, Olson M. The efficacy of bisphosphonates in the prevention of vertebral, hip, and nonvertebral-nonhip fractures in osteoporosis: a network meta-analysis. *Semin Arthritis Rheum* 2011;40(4):275-84.e1-2.

Mills EJ, Rachlis B, Wu P, Devereaux PJ, Arora P, Perri D. Primary prevention of cardiovascular mortality and events with statin treatments: a network meta-analysis involving more than 65,000 patients. *J Am Coll Cardiol* 2008;52(22):1769-1781.

Mills EJ, Perri D, Cooper C, Nachega JB, Wu P, Tleyjeh I, et al. Antifungal treatment for invasive Candida infections: a mixed treatment comparison meta-analysis. *Ann Clin Microbiol Antimicrob* 2009;8:23.

Mills EJ, Wu P, Chong G, et al. Efficacy and safety of statin treatment for cardiovascular disease: a network meta-analysis of 170,255 patients from 76 randomized trials. *QJM* 2011;104(2):109-124.

Quilici S, Abrams KR, Nicolas A, et al. Meta-analysis of the efficacy and tolerability of pramipexole versus ropinirole in the treatment of restless legs syndrome. *Sleep Med* 2008;9(7):715-726.

Quilici S, Chancellor J, Lothgren M, et al. Meta-analysis of duloxetine vs. pregabalin and gabapentin in the treatment of diabetic peripheral neuropathic pain. *BMC Neurol* 2009;9:6.

Stettler C, Allemann S, Wandel S, et al. Drug eluting and bare metal stents in people with and without diabetes: collaborative network meta-analysis. *BMJ* 2008;337:a1331.

Thompson Coon JS, Liu Z, Hoyle M, Rogers G, Green C, Moxham T, et al. Sunitinib and bevacizumab for first-line treatment of metastatic renal cell carcinoma: a systematic review and indirect comparison of clinical effectiveness. *Br J Cancer* 2009;101(2):238-243.

Protocol or abstract (n=3)

Le Cleach L, Trinquart L, LebrunVignes B, Giraudeau B, Chosidow O. Oral antiviral therapy for prevention of genital herpes outbreaks in immunocompetent and nonpregnant patients. *Cochrane Database of Systematic Reviews* 2009;1.

Filippini G, Vacchi L, D'Amico R, Di Pietrantonj C, Beecher D, Salanti G. Comparative efficacy and acceptability of immunomodulators and immunosuppressants for multiple sclerosis: overview of Cochrane systematic reviews and multiple-treatments meta-analysis. *Cochrane Database of Systematic Reviews* 2011;1.

Appendix F. Evidence Tables

TableF-1. Journal level characteristics

Journal	Included studies	Impact factor	Supplement or appendix; format	Word count limit	Table limit	Figure limit
Alimentary Pharmacology & Therapeutics	Edwards, 2009a	3.861	Y, online	N	N	N
Annals of Internal Medicine	Gross, 2011	16.792	Y, not specified	3,500-4,000	4 tables or figures	4 tables or figures
Archives of Internal Medicine	Sciarretta, 2011; Cooper, 2006	10.639	Y, online	3,500	6 to 8 tables or figures	6 to 8 tables or figures
British Medical Journal	Baldwin, 2011; Hartling, 2011; Trelle, 2011; Wandel, 2010; Lam, 2007	13.471	Y, online	N	N	N
British Medical Journal Psychiatry	Eckert 2006	13.471	Y, online	N	N	N
British Journal of Anaesthesia	Maund, 2011 [†]	4.224	Y, online	5,000	N	N
British Journal of Cancer	Coon, 2009	4.831	Y, online	5,000-5,500	1 table reduces word limit by 200	1 figure reduces word limit by 200
British Journal of Ophthalmology	Van den Bruel, 2011	2.934	Y, online	3,000	5 tables or figures	5 tables or figures
Cancer Treatment Reviews	Golfinopoulos, 2009	6.811	N	N	N	N
Clinical Therapeutics	Edwards, 2009b	2.551	Y, online	5,500-6,000	N	N
Cochrane Database of Systematic Reviews	Singh, 2011; Walsh, 2010	6.186	N	N	N	N
Current Medical Research and Opinion	van de Kerkhof, 2011; Orme, 2010; Uthman, 2010; Vissers, 2010	2.609*	Y, online	11,200	N	N
Dermatology	Bansback, 2009	2.714	Y, not specified	13 pages for text, tables, figures	Included in page count	Included in page count
Drug and Alcohol Dependence	Meader, 2009	3.365	Y, online	6,000	N	N
Europace	Freemantle, 2011	1.842	Y, not specified	5,000	5	5
Gastroenterology	Woo, 2010	12.023	Y, online	6,000	Minimum of 4 to 6 figures or illustrations	Minimum of 4 to 6 figures or illustrations

Journal	Included studies	Impact factor	Supplement or appendix; format	Word count limit	Table limit	Figure limit
Health technology assessment (Winchester, England)	Maund, 2011 [†]	4.197	N	N	N	N
International Clinical Psychopharmacology	Hansen, 2008	2.762	Y, online	7,500	N	N
The Journal of the American medical Association	Anothaisintawee, 2011; Phung, 2010	30	Y, online	3,500	4 tables or figures	4 tables or figures
Journal of Hospital Infection	Wang, 2010	3.078	N	5,000	N	N
Journal of Hypertension	Coleman, 2008	3.98	Y, online	N	N	N
Journal of the National Cancer Institute	Mauri, 2008; Kyrgiou, 2006	14.697	Y, online	6,000	8 table or figures	8 tables or figures
Lancet	Trikalinos, 2009; Elliot, 2007; Cipriani, 2009; Stettler, 2007	33.633	Y, online	4,500	“Should include about 5 illustrations”	“Should include about 5 illustrations”
Lancet Infectious Disease	Manzoli, 2009	16.144	Y, online	3,000-5,000	“Should include about 5 illustrations”	“Should include about 5 illustrations”
Lancet Neurology	Bangalore, 2011	21.659	Y, online	3,000-4,500	“Should include about 5 illustrations”	“Should include about 5 illustrations”
Lancet Oncology	Golfinopoulos, 2007	17.764	Y, online	3,000-5,000	“Should include about 5-6 illustrations”	“Should include about 5-6 illustrations”
Pharmacotherapy	Baker, 2009	2.631	N	7,000	N	N
Rheumatology	Nixon, 2007	4.171	Y, online	3,500	6 figures or tables	6 figures or tables
Thrombosis and Haemostasis	Roskell 2009	4.45*	Y, online	N	N	N
Value in Health	Dakin, 2010	2.342	Y, online	N	N	N

Abbreviations: Y: yes; N: no

*: The impact factor was obtained from Web of Science, except when the symbol appears for that journal the impact factor was not available in Web of Science and was taken from the journal's website.

†: Published as a manuscript and health technology assessment report, but counted as one unique publication

Table F-2. General characteristics of Bayesian mixed treatment comparisons

Author, year (N authors)	Method- ologist	Country	Funding	# printed pages	Affiliation	Supplement or appendix	Disease state evaluated	N and type of interventions compared	N trials, N patients	Network pattern
Baldwin, 2011 (4)†	Yes	UK	Industry	11	No	Yes	Behavioral Health (GAD)	10, Rx	27 3,989	Network with ≥1 closed loop
Bangalore, 2011 (10)	No	USA	Unfunded	18	No	Yes	Cardiology (anti- hypertensives)	8, Rx	70 324,168	Network with ≥1 closed loop
Gross, 2011 (9)	No	Brazil	Government /foundation	8	No	Yes	Endocrinology (T2DM)	6, Rx	18 4,535	Network with ≥1 closed loop
Hartling, 2011 (8)	Yes	Canada, Portugal	Government /foundation	10	No	Yes	Respiratory (acute bronchiolitis)	7, Rx	48 4897	Network with ≥1 closed loop
Maund, 2011 (6)	No	UK	Government /foundation	6 [‡]	HTA	Yes	Pain (major surgery)	4, Rx	60 5,236	Network with ≥1 closed loop
Sciarretta, 2011 (5)	Yes	Italy	NR	11	No	No	Cardiology (HTN and heart failure)	8, Rx	26 223,313	Network with ≥1 closed loop
Trelle, 2011 (8)	Yes	Switzerland	Government /foundation	11	No	Yes	Pain (NSAIDs)	8, Rx	31 116,429	Network with ≥1 closed loop
van de Kerkhof, 2011 (4)	No	UK	Industry	13	No	Yes	Dermatology (psoriasis)	17, Rx	19 9,134	Network with ≥1 closed loop
van den Bruel, 2011 (6)	No	Belgium	Unfunded	6	No	No	Ophthalmology (cataract surgery)	6, Devices	21 1,769	Network with ≥1 closed loop
Dakin, 2010 (3)	Yes	UK	Industry	12	No	Yes	Gastroenterology (chronic hepatitis B)	8, Rx	23 3,702	Network with ≥1 closed loop
Orme, 2010 (5)	No	UK	Industry	18	No	Yes	Ophthalmology (glaucoma)	10, Rx 19, Rx	93 16,898	Network with ≥1 closed loop
Phung, 2010 (4)	No	USA	Government /foundation	9	No	Yes	Endocrinology (T2DM)	7, Rx	27 11,198	Network with ≥1 closed loop

Author, year (N authors)	Method-ologist	Country	Funding	# printed pages	Affiliation	Supplement or appendix	Disease state evaluated	N and type of interventions compared	N trials, N patients	Network pattern
Uthman, 2010 (2)	Yes	UK	NR	7	No	No	Behavioral Health (anxiety)	6, Rx	16 IC	Network with ≥1 closed loop
Vissers, 2010 (5)	No	Netherlands	Industry	9	No	No	Pain (cancer)	5, Rx	6 594	Network with ≥1 closed loop [§]
Walsh, 2010 (6)	No	UK	Government /foundation	221 [¶]	Cochrane	No	Dental	7, Rx	75 105,969	Network with ≥1 closed loop
Wandel, 2010 (8)	Yes	Switzerland	Government /foundation	9	No	No	Rheumatology (OA)	4, Rx	10 3,803	Network with ≥1 closed loop
Wang, 2010 (9)	No	China	Government /foundation	11	No	No	Infectious Disease (CVCs for infections)	10, Device	48 11,525	Network with ≥1 closed loop
Woo, 2010 (10)	No	Canada	Industry	12	No	Yes	Gastroenterology (chronic hepatitis B)	10, Rx	20 8,624	Network with ≥1 closed loop [§]
Baker, 2009 (3)	No	USA	NR	15	No	No	Pulmonology (COPD)	5, Rx	43 31,020	Network with ≥1 closed loop [§]
Bansback, 2009 (6)	Yes	Canada	Industry	10	No	No	Dermatology (psoriasis)	8, Rx	22 9,917	Network with ≥1 closed loop [§]
Cipriani, 2009 (12)	Yes	Italy	Unfunded	13	No	No	Behavioral Health (depression)	12, Rx	117 25,928	Network with ≥1 closed loop
Edwards, 2009a (4)	No	UK	Industry	10	No	No	Gastroenterology (erosive esophagitis)	5, Rx	12 5,181	Network with ≥1 closed loop
Edwards, 2009b (2)	No	UK	Industry	14	No	Yes	Behavioral Health (bipolar and schizophrenia)	5, Rx	48 NR	Network with ≥1 closed loop
Golfinopoulos, 2009 (6)	Yes	Greece	NR	4	No	Yes	Oncology (unknown primary site)	5, Rx	10 683	Network with ≥1 closed loop [§]

Author, year (N authors)	Methodologist	Country	Funding	# printed pages	Affiliation	Supplement or appendix	Disease state evaluated	N and type of interventions compared	N trials, N patients	Network pattern
Manzoli, 2009 (6)	Yes	Italy	NR	11	No	Yes	Infectious disease (avian flu vaccine)	13, Other	13 8,382	Network with ≥ 1 closed loop
Meader, 2009 (1)	No	UK	Unknown	5	No	Yes	Substance Abuse (opiod detoxification)	4, Rx	20 2,112	Network with ≥ 1 closed loop [§]
Coleman, 2008 (4)	No	USA	NR	8	No	No	Cardiology (anti-hypertensives)	6, Rx	27 126,137	Network with ≥ 1 closed loop
Mauri, 2008 (5)	Yes	Greece	Unfunded	12	No	Yes	Oncology (breast cancer)	22, Rx	128 26,031	Network with ≥ 1 closed loop
Stettler, 2008 (29)	Yes	Switzerland	Government /foundation	11	No	Yes	Cardiology (stents)	3, Device	35 14,799	Closed loop [§]
Golfinopoulos, 2007 (4)	Yes	Greece	Unfunded	14	No	Yes	Oncology (colorectal cancer)	12, Rx	40 15,802	Network with ≥ 1 closed loop
Lam, 2007 (2)	No	China	Unfunded	10	No	No	Cardiology (left ventricular dysfunction)	5, Device and Rx	12 8,307	Network with ≥ 1 closed loop
Nixon, 2007 (3)	Yes	UK	Other	8	No	No	Rheumatology (RA)	9, Rx	13 6,694	Network with ≥ 1 closed loop [§]
Cooper, 2006 (4)	No	England	Government /foundation	7	No	Yes	Cardiology (stroke prevention)	8, Rx	19 17,833	Network with ≥ 1 closed loop
Kyrgiou, 2006 (5)	Yes	Greece	NR	9	No	Yes	Oncology (ovarian cancer)	8, Rx	60 16,478	Network with ≥ 1 closed loop [§]

Abbreviations: CODP: chronic obstructive pulmonary disease; CVC: central venous catheter; DM: diabetes mellitus; HTA: Health Technology Assessment; HTN: hypertension; NSAID: non-steroidal anti-inflammatory drugs; RA: rheumatoid arthritis; RCC: renal cell carcinoma; RLS: restless leg syndrome; Rx: pharmacologic; T2DM: type 2 diabetes mellitus;

*: A methodologist was considered an individual with affiliation to a department of statistics, biostatistics, epidemiology, clinical epidemiology, or public health services, as determined by author information and affiliations listed in the publication.

†: Includes both a Bayesian MTC model and a Frequentist MTC model therefore appears in both tables.

- ‡: Published as a manuscript and report, with the manuscript serving as the primary data source.
- §: Diagram was not provided, pattern determined from study characteristics reported
- ||: Two models reported
- ¶: Cochrane report

Table F-3. Methodological characteristics of Bayesian mixed treatment comparisons

Author, year	Meta-analysis model	MCMC details and Priors	Heterogeneity, inconsistency, and claims of equivalence or non-inferiority
Baldwin, 2011*	Was traditional meta-analysis run? Yes	Was the code available? No	Heterogeneity assessment in traditional meta-analysis: NR
	Model(s): Random effects	Was the raw data available? No	Heterogeneity assessment in Bayesian meta-analysis: NR
	Adjustment for multiple arms: NR	Starting values: NR	Evaluation of inconsistency: Compared consistency of results from mixed treatment meta-analysis and the direct comparative meta-analysis
	Adjustment for covariates: NR	Number of chains: NR	Equivalence claims: NR
	Model fit tested: NR	Number of iterations and burn-in: 21,000 iterations; 1,000 burn-in	Non-inferiority claims: NR
	Was there graphical representation of the posterior distribution? Yes	Convergence statistics evaluated: Yes, considering kernel density plots	Minimally important difference defined: NR
	Did authors rank order interventions? Yes	Prior distribution of d: NR	
	Software used: WinBUGS	Prior distribution for σ : NR	
		Were priors justified: No; "Vague prior parameters were chosen"	
		Was a sensitivity analysis conducted based on priors? NR	

Author, year	Meta-analysis model	MCMC details and Priors	Heterogeneity, inconsistency, and claims of equivalence or non-inferiority
Bangalore, 2011	<p>Was traditional meta-analysis run? Yes</p> <p>Model(s): Fixed and random effects</p> <p>Adjustment for multiple arms: NR</p> <p>Adjustment for covariates: NR</p> <p>Model fit tested: NR</p> <p>Was there graphical representation of the posterior distribution? No</p> <p>Did authors rank order interventions? Yes</p> <p>Software used: WinBUGS 1.4.3</p>	<p>Was the code available? No</p> <p>Was the raw data available? No</p> <p>Starting values: NR</p> <p>Number of chains: NR</p> <p>Number of iterations and burn-in: NR</p> <p>Convergence statistics evaluated: NR</p> <p>Prior distribution of d: NR</p> <p>Prior distribution for σ: NR</p> <p>Were priors justified: Yes; vague priors were used for comparisons of treatments so the findings were close to those obtained with frequentist models</p> <p>Was a sensitivity analysis conducted based on priors? NR</p>	<p>Heterogeneity assessment in traditional meta-analysis: I^2</p> <p>Heterogeneity assessment in Bayesian meta-analysis: NR</p> <p>Evaluation of inconsistency: Used various statistical modeling (traditional MA, network MA, and trial sequential analyses) to assess for consistency in the magnitude and direction of effect size</p> <p>Equivalence claims: NR</p> <p>Non-inferiority claims: NR</p> <p>Minimally important difference defined: “Our meta-analysis refutes a 5.0% to 10.0% relative risk increase in either cancer or cancer related death with most antihypertensive drug classes.”</p> <p>“10% relative risk increase because this small increase in cancer risk is likely to be clinically meaningful”</p>

Author, year	Meta-analysis model	MCMC details and Priors	Heterogeneity, inconsistency, and claims of equivalence or non-inferiority
Gross, 2011	<p>Was traditional meta-analysis run? Yes</p> <p>Model(s): NR</p> <p>Adjustment for multiple arms: NR</p> <p>Adjustment for covariates: NR</p> <p>Model fit tested: NR</p> <p>Was there graphical representation of the posterior distribution? No</p> <p>Did authors rank order interventions? Yes</p> <p>Software used: WinBUGS 1.4.3</p>	<p>Was the code available? Option to email author although no reply</p> <p>Was the raw data available? Yes, in manuscript</p> <p>Starting values: NR</p> <p>Number of chains: NR</p> <p>Number of iterations and burn-in: NR</p> <p>Convergence statistics evaluated: NR</p> <p>Prior distribution of d: NR</p> <p>Prior distribution for σ: NR</p> <p>Were priors justified: NR and did not specify if vague priors used or not</p> <p>Was a sensitivity analysis conducted based on priors? NR</p>	<p>Heterogeneity assessment in traditional meta-analysis: Cochrane Q-statistic, I^2</p> <p>Heterogeneity assessment in Bayesian meta-analysis: NR</p> <p>Evaluation of inconsistency: Compare findings with prior meta-analysis and between traditional and network meta-analyses</p> <p>Equivalence claims: NR</p> <p>Non-inferiority claims: NR</p> <p>Minimally important difference defined: NR</p>

Author, year	Meta-analysis model	MCMC details and Priors	Heterogeneity, inconsistency, and claims of equivalence or non-inferiority
Hartling, 2011	<p>Was traditional meta-analysis run? Yes</p> <p>Model(s): Random effects</p> <p>Adjustment for multiple arms: NR</p> <p>Adjustment for covariates: NR</p> <p>Model fit tested: NR</p> <p>Was there graphical representation of the posterior distribution? No</p> <p>Did authors rank order interventions? Yes</p> <p>Software used: WinBUGS</p>	<p>Was the code available? No</p> <p>Was the raw data available? No</p> <p>Starting values: NR</p> <p>Number of chains: NR</p> <p>Number of iterations and burn-in: 220,000 iterations; burn-in 20,000</p> <p>Convergence statistics evaluated: NR</p> <p>Prior distribution of d: Normal, 0 to 10,000</p> <p>Prior distribution for σ: Uniform, 0 to 2 (admissions) or 0 to 10 (length of stay)</p> <p>Were priors justified: NR, consider these non-informative priors</p> <p>Was a sensitivity analysis conducted based on priors? Yes</p>	<p>Heterogeneity assessment in traditional meta-analysis: I^2</p> <p>Heterogeneity assessment in Bayesian meta-analysis: NR</p> <p>Evaluation of inconsistency: Cross validation of all contrasts that had direct evidence</p> <p>Equivalence claims: NR</p> <p>Non-inferiority claims: NR</p> <p>Minimally important difference defined: NR</p>

Author, year	Meta-analysis model	MCMC details and Priors	Heterogeneity, inconsistency, and claims of equivalence or non-inferiority
Maund, 2011	<p>Was traditional meta-analysis run? Yes</p> <p>Model(s): Random effects</p> <p>Adjustment for multiple arms: Yes</p> <p>Adjustment for covariates: Yes, baseline morphine consumption</p> <p>Model fit tested: Yes, residual deviation and DIC</p> <p>Was there graphical representation of the posterior distribution? No</p> <p>Did authors rank order interventions? Yes</p> <p>Software used: WinBUGS</p>	<p>Was the code available? Yes, referral to another website/source</p> <p>Was the raw data available? Yes, in the full report published as a HTA</p> <p>Starting values: NR</p> <p>Number of chains: NR</p> <p>Number of iterations and burn-in: 105,000 iterations; Burn-in 5,000</p> <p>Convergence statistics evaluated: Yes, NR</p> <p>Prior distribution of d: dnorm(0,0.0001)</p> <p>Prior distribution for σ: dunif(0,2)</p> <p>Were priors justified: Yes, used uninformative priors</p> <p>Was a sensitivity analysis conducted based on priors? No</p>	<p>Heterogeneity assessment in traditional meta-analysis: Consideration of PICO, visualization of results, Chi^2, I^2</p> <p>Heterogeneity assessment in Bayesian meta-analysis: NR</p> <p>Evaluation of inconsistency: Compared with direct evidence synthesis within this report and with prior reports.</p> <p>Equivalence claims: NR</p> <p>Non-inferiority claims: NR</p> <p>Minimally important difference defined: NR</p>

Author, year	Meta-analysis model	MCMC details and Priors	Heterogeneity, inconsistency, and claims of equivalence or non-inferiority
Sciarretta, 2011	Was traditional meta-analysis run? Yes	Was the code available? No	Heterogeneity assessment in traditional meta-analysis: Chi-squared
	Model(s): Random effects	Was the raw data available? Yes, in the manuscript	Heterogeneity assessment in Bayesian meta-analysis: NR
	Adjustment for multiple arms: NR	Starting values: NR	Evaluation of inconsistency: Inconsistency was calculated as suggested by Lu and Ades, σ_w^2
	Adjustment for covariates: NR	Number of chains: 2	Equivalence claims: NR
	Model fit tested: NR	Number of iterations and burn-in: 105,000 Iterations; 5,000 burn-in	Non-inferiority claims: NR
	Was there graphical representation of the posterior distribution? No	Convergence statistics evaluated: Yes, Gelman Rubin statistics to determine burn in	Minimally important difference defined: NR
	Did authors rank order interventions? No	Prior distribution of d: NR	
	Software used: WinBUGS	Prior distribution for σ : NR	
		Were priors justified: NR; We also used noninformative priors that represented complete lack of credible prior information.	
		Was a sensitivity analysis conducted based on priors? No	

Author, year	Meta-analysis model	MCMC details and Priors	Heterogeneity, inconsistency, and claims of equivalence or non-inferiority
Trelle, 2011	<p>Was traditional meta-analysis run? Yes</p> <p>Model(s): Random effects</p> <p>Adjustment for multiple arms: Yes</p> <p>Adjustment for covariates: NR</p> <p>Model fit tested: Yes, residual deviance</p> <p>Was there graphical representation of the posterior distribution? Yes</p> <p>Did authors rank order interventions? No</p> <p>Software used: WinBUGS</p>	<p>Was the code available? Yes, in the online supplement</p> <p>Was the raw data available? Yes, in the manuscript</p> <p>Starting values: NR</p> <p>Number of chains: NR</p> <p>Number of iterations and burn-in: 100,000 iterations; 50,000 burn-in</p> <p>Convergence statistics evaluated: Yes, Gelman Rubin statistic</p> <p>Prior distribution of d: dnorm(0, 0.001)</p> <p>Prior distribution for σ: dunif(0,2)</p> <p>Were priors justified: NR, used minimally informative priors</p> <p>Was a sensitivity analysis conducted based on priors? No</p>	<p>Heterogeneity assessment in traditional meta-analysis: NR</p> <p>Heterogeneity assessment in Bayesian meta-analysis: Tau²</p> <p>Evaluation of inconsistency: Inconsistency factors, defined as the difference in log rate ratios derived from direct and indirect comparisons</p> <p>Equivalence claims: NR</p> <p>Non-inferiority claims: NR</p> <p>Minimally important difference defined: Pre-specified rate ratio of 1.3 as the primary threshold for outcomes evaluated</p>

Author, year	Meta-analysis model	MCMC details and Priors	Heterogeneity, inconsistency, and claims of equivalence or non-inferiority
van de Kerkhof, 2011	<p>Was traditional meta-analysis run? No</p> <p>Model(s): Fixed and random effects</p> <p>Adjustment for multiple arms: NR</p> <p>Adjustment for covariates: NR</p> <p>Model fit tested: Yes, but method NR</p> <p>Was there graphical representation of the posterior distribution? No</p> <p>Did authors rank order interventions? Yes</p> <p>Software used: WinBUGS 1.4</p>	<p>Was the code available? No</p> <p>Was the raw data available? Yes, in the manuscript</p> <p>Starting values: NR</p> <p>Number of chains: NR</p> <p>Number of iterations and burn-in: NR</p> <p>Convergence statistics evaluated: NR</p> <p>Prior distribution of d: NR</p> <p>Prior distribution for σ: NR</p> <p>Were priors justified: Yes, In order not to influence the estimates by the choice of the prior distribution, a non-informative (i.e., 'flat') distribution was used for the parameters of the model. With such a prior distribution, results as reflected with the posterior distribution are solely driven by the data.</p> <p>Was a sensitivity analysis conducted based on priors? No</p>	<p>Heterogeneity assessment in traditional meta-analysis: NA</p> <p>Heterogeneity assessment in Bayesian meta-analysis: NR</p> <p>Evaluation of inconsistency: Compare results with previously conducted meta-analyses</p> <p>Equivalence claims: NR</p> <p>Non-inferiority claims: NR</p> <p>Minimally important difference defined: NR</p>

Author, year	Meta-analysis model	MCMC details and Priors	Heterogeneity, inconsistency, and claims of equivalence or non-inferiority
Van den Bruel, 2011	<p>Was traditional meta-analysis run? Yes</p> <p>Model(s): Fixed and random effects</p> <p>Adjustment for multiple arms: NR</p> <p>Adjustment for covariates: NR</p> <p>Model fit tested: Yes, DIC</p> <p>Was there graphical representation of the posterior distribution? No</p> <p>Did authors rank order interventions? Yes</p> <p>Software used: WinBUGS 1.4</p>	<p>Was the code available? No</p> <p>Was the raw data available? Yes, in the manuscript</p> <p>Starting values: NR</p> <p>Number of chains: NR</p> <p>Number of iterations and burn-in: NR</p> <p>Convergence statistics evaluated: NR</p> <p>Prior distribution of d: NR</p> <p>Prior distribution for σ: NR</p> <p>Were priors justified: NR and did not report if vague priors used or not</p> <p>Was a sensitivity analysis conducted based on priors? No</p>	<p>Heterogeneity assessment in traditional meta-analysis: NR</p> <p>Heterogeneity assessment in Bayesian meta-analysis: NR</p> <p>Evaluation of inconsistency: Consistency of evidence sources was assessed by calculating the posterior mean residual deviance for each individual treatment arm</p> <p>Equivalence claims: NR</p> <p>Non-inferiority claims: NR</p> <p>Minimally important difference defined: A difference of more than 100 cells/mm³ was clinically relevant difference, further justified in text</p>

Author, year	Meta-analysis model	MCMC details and Priors	Heterogeneity, inconsistency, and claims of equivalence or non-inferiority
Dakin, 2010	<p>Was traditional meta-analysis run? Yes</p> <p>Model(s): Fixed and random effects</p> <p>Adjustment for multiple arms: Yes</p> <p>Adjustment for covariates: NR</p> <p>Model fit tested: Yes, DIC and residual deviance</p> <p>Was there graphical representation of the posterior distribution? No</p> <p>Did authors rank order interventions? Yes</p> <p>Software used: WinBUGS 1.4</p>	<p>Was the code available? Yes, in the online supplement</p> <p>Was the raw data available? Yes, in the manuscript</p> <p>Starting values: NR</p> <p>Number of chains: 2</p> <p>Number of iterations and burn-in: 520,000 to 945,000 Iterations; 500,000 to 925,000 burn-in</p> <p>Convergence statistics evaluated: Yes, NR</p> <p>Prior distribution of d: dnorm(0, 0.001)</p> <p>Prior distribution for σ: dnorm(0,K)I(0,) or unif(0,10)</p> <p>Were priors justified: Yes, "Sensitivity analyses suggested that the posterior estimates of the uncertainty around treatment effects (but not the posterior means) were sensitive to the priors used. Informative half-normal priors were therefore used for the between-studies SD in order to allow this external data to help inform the between-studies SD; these distributions were based on a meta-analysis of interferon trials identified in a published systematic review."</p> <p>Was a sensitivity analysis conducted based on priors? Yes, alternate priors NR</p>	<p>Heterogeneity assessment in traditional meta-analysis: NR</p> <p>Heterogeneity assessment in Bayesian meta-analysis: Between study SD</p> <p>Evaluation of inconsistency: NR</p> <p>Equivalence claims: NR</p> <p>Non-inferiority claims: NR</p> <p>Minimally important difference defined: NR</p>

Author, year	Meta-analysis model	MCMC details and Priors	Heterogeneity, inconsistency, and claims of equivalence or non-inferiority
Orme, 2010	<p>Was traditional meta-analysis run? No</p> <p>Model(s): Random effects</p> <p>Adjustment for multiple arms: Yes</p> <p>Adjustment for covariates: Not in main model but sensitivity analysis conducted to adjust for baseline intra-ocular pressure</p> <p>Model fit tested: NR</p> <p>Was there graphical representation of the posterior distribution? No</p> <p>Did authors rank order interventions? No</p> <p>Software used: WinBUGS 1.4</p>	<p>Was the code available? No</p> <p>Was the raw data available? No</p> <p>Starting values: NR</p> <p>Number of chains: 2</p> <p>Number of iterations and burn-in: 120,000 Iterations; 100,000 burn-in</p> <p>Convergence statistics evaluated: Yes, visual plot inspection</p> <p>Prior distribution of d: Normal</p> <p>Prior distribution for σ: Uniform(0,10)</p> <p>Were priors justified: No, "Uninformative normal priors were used for all model parameters except for the between-studies SD for which an uninformative uniform prior."</p> <p>Was a sensitivity analysis conducted based on priors? No</p>	<p>Heterogeneity assessment in traditional meta-analysis: NA</p> <p>Heterogeneity assessment in Bayesian meta-analysis: Between study SD</p> <p>Evaluation of inconsistency: NR</p> <p>Equivalence claims: NR</p> <p>Non-inferiority claims: NR</p> <p>Minimally important difference defined: NR</p>

Author, year	Meta-analysis model	MCMC details and Priors	Heterogeneity, inconsistency, and claims of equivalence or non-inferiority
Phung, 2010	Was traditional meta-analysis run? Yes	Was the code available? No	Heterogeneity assessment in traditional meta-analysis: I^2
	Model(s): Random effects	Was the raw data available? No	Heterogeneity assessment in Bayesian meta-analysis: NR
	Adjustment for multiple arms: NR	Starting values: NR	Evaluation of inconsistency: Qualitative comparison of results from traditional and network analyses
	Adjustment for covariates: NR	Number of chains: NR	Equivalence claims: NR
	Model fit tested: Yes, residual deviance	Number of iterations and burn-in: NR	Non-inferiority claims: Drugs investigated produce similar glucose lowering when applying minimally important difference
	Was there graphical representation of the posterior distribution? No	Convergence statistics evaluated: NR	Minimally important difference defined: 0.4% non-inferiority margin used by the Food and Drug Administration
	Did authors rank order interventions? No	Prior distribution of d: NR	
	Software used: WinBUGS	Prior distribution for σ : NR	
		Were priors justified: NR and did not report if vague priors used or not	
		Was a sensitivity analysis conducted based on priors? No	

Author, year	Meta-analysis model	MCMC details and Priors	Heterogeneity, inconsistency, and claims of equivalence or non-inferiority
Uthman, 2010	Was traditional meta-analysis run? No	Was the code available? No	Heterogeneity assessment in traditional meta-analysis: NR
	Model(s): NR	Was the raw data available? Yes, in the manuscript	Heterogeneity assessment in Bayesian meta-analysis: NR
	Adjustment for multiple arms: NR	Starting values: NR	Evaluation of inconsistency: NR
	Adjustment for covariates: NR	Number of chains: NR	Equivalence claims: NR
	Model fit tested: NR	Number of iterations and burn-in: NR	Non-inferiority claims: NR
	Was there graphical representation of the posterior distribution? No	Convergence statistics evaluated: NR	Minimally important difference defined: NR
	Did authors rank order interventions? Yes	Prior distribution of d: NR	
	Software used: WinBUGS	Prior distribution for σ : NR	
		Were priors justified: NR and did not report if vague priors used or not	
		Was a sensitivity analysis conducted based on priors? No	

Author, year	Meta-analysis model	MCMC details and Priors	Heterogeneity, inconsistency, and claims of equivalence or non-inferiority
Vissers, 2010	<p>Was traditional meta-analysis run? No</p> <p>Model(s): Fixed effects</p> <p>Adjustment for multiple arms: NR</p> <p>Adjustment for covariates: NR</p> <p>Model fit tested: Yes, method NR</p> <p>Was there graphical representation of the posterior distribution? No</p> <p>Did authors rank order interventions? Yes</p> <p>Software used: WinBUGS 1.4</p>	<p>Was the code available? No</p> <p>Was the raw data available? No</p> <p>Starting values: NR</p> <p>Number of chains: NR</p> <p>Number of iterations and burn-in: NR</p> <p>Convergence statistics evaluated: NR</p> <p>Prior distribution of d: NR</p> <p>Prior distribution for σ: NR</p> <p>Were priors justified: NR and did not specify if vague priors used or not</p> <p>Was a sensitivity analysis conducted based on priors? No</p>	<p>Heterogeneity assessment in traditional meta-analysis: NA</p> <p>Heterogeneity assessment in Bayesian meta-analysis: NR</p> <p>Evaluation of inconsistency: NR</p> <p>Equivalence claims: NR</p> <p>Non-inferiority claims: NR</p> <p>Minimally important difference defined: NR</p>

Author, year	Meta-analysis model	MCMC details and Priors	Heterogeneity, inconsistency, and claims of equivalence or non-inferiority
Walsh, 2010	<p>Was traditional meta-analysis run? Yes</p> <p>Model(s): Random effects</p> <p>Adjustment for multiple arms: Yes</p> <p>Adjustment for covariates: NR</p> <p>Model fit tested: Yes, median sum deviation</p> <p>Was there graphical representation of the posterior distribution? No</p> <p>Did authors rank order interventions? No</p> <p>Software used: WinBUGS</p>	<p>Was the code available? Yes, external website</p> <p>Was the raw data available? No</p> <p>Starting values: NR</p> <p>Number of chains: NR</p> <p>Number of iterations and burn-in: NR</p> <p>Convergence statistics evaluated: NR</p> <p>Prior distribution of d: dnorm(0,0.001)</p> <p>Prior distribution for σ: dunif(0,2)</p> <p>Were priors justified: NR, code says "vague priors"</p> <p>Was a sensitivity analysis conducted based on priors? No</p>	<p>Heterogeneity assessment in traditional meta-analysis: Chi² and I²</p> <p>Heterogeneity assessment in Bayesian meta-analysis: Tau</p> <p>Evaluation of inconsistency: NR</p> <p>Equivalence claims: NR</p> <p>Non-inferiority claims: NR</p> <p>Minimally important difference defined: NR</p>

Author, year	Meta-analysis model	MCMC details and Priors	Heterogeneity, inconsistency, and claims of equivalence or non-inferiority
Wandel, 2010	<p>Was traditional meta-analysis run? Yes</p> <p>Model(s): Random effects</p> <p>Adjustment for multiple arms: No</p> <p>Adjustment for covariates: Yes, time, allocation concealment, intention-to-treat, quality, glucosamine type, quality control of preparation, joint type</p> <p>Model fit tested: Yes, Q-Q plots</p> <p>Was there graphical representation of the posterior distribution? No</p> <p>Did authors rank order interventions? No</p> <p>Software used: WinBUGS 1.4</p>	<p>Was the code available? No</p> <p>Was the raw data available? No</p> <p>Starting values: NR</p> <p>Number of chains: NR</p> <p>Number of iterations and burn-in: 150,000 iterations; 50,000 burn-in</p> <p>Convergence statistics evaluated: Yes, Gelman Rubin statistic</p> <p>Prior distribution of d: NR</p> <p>Prior distribution for σ: NR</p> <p>Were priors justified: No, model used minimally informative prior distributions</p> <p>Was a sensitivity analysis conducted based on priors? Yes</p>	<p>Heterogeneity assessment in traditional meta-analysis: NR</p> <p>Heterogeneity assessment in Bayesian meta-analysis: Tau², calculated p-value for heterogeneity</p> <p>Evaluation of inconsistency: Inconsistency factors</p> <p>Equivalence claims: “None of the pooled estimates crossed the pre-specified boundary of a minimal clinically important difference of -0.9 cm on a 10cm visual analogue scale.”</p> <p>Non-inferiority claims: “The lower end of the credible intervals did not cross the pre-specified boundaries.”</p> <p>Minimally important difference defined: Minimal clinically important difference of 0.37 SD units corresponding to 0.9cm on a 10cm visual analogue scale. Based on recent studies in patients with OA.</p>

Author, year	Meta-analysis model	MCMC details and Priors	Heterogeneity, inconsistency, and claims of equivalence or non-inferiority
Wang, 2010	Was traditional meta-analysis run? No	Was the code available? No	Heterogeneity assessment in traditional meta-analysis: NA
	Model(s): Fixed and random effects	Was the raw data available? Yes, in the manuscript	Heterogeneity assessment in Bayesian meta-analysis: Tau ²
	Adjustment for multiple arms: NR	Starting values: NR	Evaluation of inconsistency: Results of head-to-head comparisons from previous conventional meta-analyses were concordant with results from our network meta-analysis, indicating that the network of trials was consistent
	Adjustment for covariates: In sensitivity analysis only (methodological quality and no central venous catheter per patient)	Number of chains: NR	
	Model fit tested: Yes, DIC and residual deviance	Number of iterations and burn-in: NR	
	Was there graphical representation of the posterior distribution? No	Convergence statistics evaluated: NR	Equivalence claims: NR
	Did authors rank order interventions? No	Prior distribution of d: NR	Non-inferiority claims: NR
	Software used: WinBUGS 1.4.3	Prior distribution for σ : NR	Minimally important difference defined: NR
		Were priors justified: NR and did not specify if vague priors used or not	
		Was a sensitivity analysis conducted based on priors? No	

Author, year	Meta-analysis model	MCMC details and Priors	Heterogeneity, inconsistency, and claims of equivalence or non-inferiority
Woo, 2010	<p>Was traditional meta-analysis run? Yes</p> <p>Model(s): Random effects</p> <p>Adjustment for multiple arms: NR</p> <p>Adjustment for covariates: NR</p> <p>Model fit tested: Yes, residual deviance</p> <p>Was there graphical representation of the posterior distribution? No</p> <p>Did authors rank order interventions? Yes</p> <p>Software used: WinBUGS 1.4.3</p>	<p>Was the code available? Yes, in online supplement</p> <p>Was the raw data available? Yes, in the manuscript</p> <p>Starting values: NR</p> <p>Number of chains: 3</p> <p>Number of iterations and burn-in: 25,000 iterations; 5,000 burn-in</p> <p>Convergence statistics evaluated: Yes, Gelman Rubin Brooke statistic</p> <p>Prior distribution of d: dnorm(0, 0.1)</p> <p>Prior distribution for σ: dt(0,1,2) l(0,)</p> <p>Were priors justified: NR and did not specify if vague priors used or not</p> <p>Was a sensitivity analysis conducted based on priors? No</p>	<p>Heterogeneity assessment in traditional meta-analysis: NR</p> <p>Heterogeneity assessment in Bayesian meta-analysis: Between study SD in log OR</p> <p>Evaluation of inconsistency: NR</p> <p>Equivalence claims: NR</p> <p>Non-inferiority claims: NR</p> <p>Minimally important difference defined: NR</p>

Author, year	Meta-analysis model	MCMC details and Priors	Heterogeneity, inconsistency, and claims of equivalence or non-inferiority
Baker, 2009	Was traditional meta-analysis run? Yes	Was the code available? No	Heterogeneity assessment in traditional meta-analysis: Cochrane Q statistic
	Model(s): NR	Was the raw data available? No	Heterogeneity assessment in Bayesian meta-analysis: NR
	Adjustment for multiple arms: NR	Starting values: NR	Evaluation of inconsistency: NR
	Adjustment for covariates: NR	Number of chains: NR	Equivalence claims: NR
	Model fit tested: NR	Number of iterations and burn-in: NR	Non-inferiority claims: NR
	Was there graphical representation of the posterior distribution? No	Convergence statistics evaluated: NR	Minimally important difference defined: NA
	Did authors rank order interventions? No	Prior distribution of d: NR	
	Software used: WinBUGS + BUGSXLTA Wrapper	Prior distribution for σ : NR	
		Were priors justified: NR and did not specify if vague priors used or not	
		Was a sensitivity analysis conducted based on priors? No	

Author, year	Meta-analysis model	MCMC details and Priors	Heterogeneity, inconsistency, and claims of equivalence or non-inferiority
Bansback, 2009	<p>Was traditional meta-analysis run? No</p> <p>Model(s): NR</p> <p>Adjustment for multiple arms: NR</p> <p>Adjustment for covariates: Yes, placebo response rate</p> <p>Model fit tested: NR</p> <p>Was there graphical representation of the posterior distribution? No</p> <p>Did authors rank order interventions? Yes</p> <p>Software used: WinBUGS 1.4.1</p>	<p>Was the code available? No</p> <p>Was the raw data available? Yes, in manuscript</p> <p>Starting values: NR</p> <p>Number of chains: NR</p> <p>Number of iterations and burn-in: 15,000 iterations; 5,000 burn-in</p> <p>Convergence statistics evaluated: NR</p> <p>Prior distribution of d: NR</p> <p>Prior distribution for σ: NR</p> <p>Were priors justified: NR, uninformative prior distributions for each treatment.</p> <p>Was a sensitivity analysis conducted based on priors? No</p>	<p>Heterogeneity assessment in traditional meta-analysis: NA</p> <p>Heterogeneity assessment in Bayesian meta-analysis: NR</p> <p>Evaluation of inconsistency: NR</p> <p>Equivalence claims: NR</p> <p>Non-inferiority claims: NR</p> <p>Minimally important difference defined: NR</p>

Author, year	Meta-analysis model	MCMC details and Priors	Heterogeneity, inconsistency, and claims of equivalence or non-inferiority
Cipriani, 2009	<p>Was traditional meta-analysis run? Yes</p> <p>Model(s): Random effects</p> <p>Adjustment for multiple arms: NR</p> <p>Adjustment for covariates: NR</p> <p>Model fit tested: NR</p> <p>Was there graphical representation of the posterior distribution? Yes</p> <p>Did authors rank order interventions? Yes</p> <p>Software used: WinBUGS</p>	<p>Was the code available? No</p> <p>Was the raw data available? Yes, referred to external website</p> <p>Starting values: NR</p> <p>Number of chains: NR</p> <p>Number of iterations and burn-in: NR</p> <p>Convergence statistics evaluated: NR</p> <p>Prior distribution of d: NR</p> <p>Prior distribution for σ: NR</p> <p>Were priors justified: NR and did not specify if vague prior used or not</p> <p>Was a sensitivity analysis conducted based on priors? No</p>	<p>Heterogeneity assessment in traditional meta-analysis: I^2 and visual inspection of forest plots</p> <p>Heterogeneity assessment in Bayesian meta-analysis: NR</p> <p>Evaluation of inconsistency: Calculated ratio of odds ratios for indirect versus direct evidence whenever indirect estimates could be constructed with a single common comparator. Incoherence was defined as the disagreement between direct and indirect evidence with a 95% confidence interval excluding 1</p> <p>Equivalence claims: NR</p> <p>Non-inferiority claims: NR</p> <p>Minimally important difference defined: NR</p>

Author, year	Meta-analysis model	MCMC details and Priors	Heterogeneity, inconsistency, and claims of equivalence or non-inferiority
Edwards, 2009a	<p>Was traditional meta-analysis run? No</p> <p>Model(s): Fixed and random effects</p> <p>Adjustment for multiple arms: NR</p> <p>Adjustment for covariates: Yes, publication year</p> <p>Model fit tested: Yes, DIC</p> <p>Was there graphical representation of the posterior distribution? No</p> <p>Did authors rank order interventions? Yes</p> <p>Software used: WinBUGS 1.4.3</p>	<p>Was the code available? No</p> <p>Was the raw data available? Yes, in manuscript</p> <p>Starting values: NR</p> <p>Number of chains: NR</p> <p>Number of iterations and burn-in: NR</p> <p>Convergence statistics evaluated: NR</p> <p>Prior distribution of d: NR</p> <p>Prior distribution for σ: NR</p> <p>Were priors justified: NR and did not specify if vague priors used or not</p> <p>Was a sensitivity analysis conducted based on priors? No</p>	<p>Heterogeneity assessment in traditional meta-analysis: NA</p> <p>Heterogeneity assessment in Bayesian meta-analysis: SD</p> <p>Evaluation of inconsistency: Posterior mean residual deviance</p> <p>Equivalence claims: NR</p> <p>Non-inferiority claims: NR</p> <p>Minimally important difference defined: NR</p>

Author, year	Meta-analysis model	MCMC details and Priors	Heterogeneity, inconsistency, and claims of equivalence or non-inferiority
Edwards, 2009b	Was traditional meta-analysis run? No	Was the code available? No	Heterogeneity assessment in traditional meta-analysis: NA
	Model(s): Random effects	Was the raw data available? No	Heterogeneity assessment in Bayesian meta-analysis: Pre-specified SD values within pairwise comparisons
	Adjustment for multiple arms: Yes	Starting values: NR	Evaluation of inconsistency: Estimated by assessing the posterior mean residual deviance
	Adjustment for covariates: NR	Number of chains: NR	Equivalence claims: NR
	Model fit tested: Yes, residual deviance	Number of iterations and burn-in: NR	Non-inferiority claims: NR
	Was there graphical representation of the posterior distribution? No	Convergence statistics evaluated: NR	Minimally important difference defined: NR
	Did authors rank order interventions? Yes	Prior distribution of d: NR	
	Software used: WinBUGS	Prior distribution for σ : NR	
		Were priors justified: Yes, Vague prior distributions were used for comparisons of treatments so that the findings would be close to those obtained with frequentist methods.	
		Was a sensitivity analysis conducted based on priors? No	

Author, year	Meta-analysis model	MCMC details and Priors	Heterogeneity, inconsistency, and claims of equivalence or non-inferiority
Golfinopoulos, 2009	Was traditional meta-analysis run? Yes	Was the code available? No	Heterogeneity assessment in traditional meta-analysis: I^2
	Model(s): NR	Was the raw data available? No	Heterogeneity assessment in Bayesian meta-analysis: NR
	Adjustment for multiple arms: NR	Starting values: NR	Evaluation of inconsistency: State "there was no clear evidence for incoherence" although do not report methods used to determine this
	Adjustment for covariates: NR	Number of chains: 3	
	Model fit tested: NR	Number of iterations and burn-in: 70,000 iterations; 20,000 burn-in	Equivalence claims: NR
	Was there graphical representation of the posterior distribution? No	Convergence statistics evaluated: Yes, ensured after observing mixing of 3 chains	Non-inferiority claims: NR
	Did authors rank order interventions? Yes	Prior distribution of d : NR	Minimally important difference defined: NR
	Software used: WinBUGS	Prior distribution for σ : NR	
		Were priors justified: NR, used approximately vague normal priors for all location parameters	
		Was a sensitivity analysis conducted based on priors? No	

Author, year	Meta-analysis model	MCMC details and Priors	Heterogeneity, inconsistency, and claims of equivalence or non-inferiority
Manzoli, 2009	<p>Was traditional meta-analysis run? Yes</p> <p>Model(s): Random effects</p> <p>Adjustment for multiple arms: NR</p> <p>Adjustment for covariates: NR</p> <p>Model fit tested: NR</p> <p>Was there graphical representation of the posterior distribution? No</p> <p>Did authors rank order interventions? No</p> <p>Software used: WinBUGS 1.4.3</p>	<p>Was the code available? No</p> <p>Was the raw data available? No</p> <p>Starting values: NR</p> <p>Number of chains: 3</p> <p>Number of iterations and burn-in: 250,000 iterations; 50,000 burn-in</p> <p>Convergence statistics evaluated: Yes, observing mix of 3 chains and Brooks Gelman Rubin diagnostic tool</p> <p>Prior distribution of d: NR</p> <p>Prior distribution for σ: NR</p> <p>Were priors justified: NR, we used approximately vague normal priors for all location parameters</p> <p>Was a sensitivity analysis conducted based on priors? No</p>	<p>Heterogeneity assessment in traditional meta-analysis: I^2</p> <p>Heterogeneity assessment in Bayesian meta-analysis: NR</p> <p>Evaluation of inconsistency: Assessed inconsistency between different sources of evidence (direct and indirect) in each closed loop, as previously described</p> <p>Equivalence claims: NR</p> <p>Non-inferiority claims: NR</p> <p>Minimally important difference defined: Absolute immunogenicity of 70% considered satisfactory based on the Committee for Proprietary Medicinal Products, and used to make superiority claims.</p>

Author, year	Meta-analysis model	MCMC details and Priors	Heterogeneity, inconsistency, and claims of equivalence or non-inferiority
Meader, 2009	<p>Was traditional meta-analysis run? Yes</p> <p>Model(s): Fixed and random effects</p> <p>Adjustment for multiple arms: NR</p> <p>Adjustment for covariates: NR</p> <p>Model fit tested: Yes, residual deviance</p> <p>Was there graphical representation of the posterior distribution? No</p> <p>Did authors rank order interventions? Yes</p> <p>Software used: WinBUGS</p>	<p>Was the code available? Yes, online supplement</p> <p>Was the raw data available? Yes, in manuscript</p> <p>Starting values: NR</p> <p>Number of chains: 2</p> <p>Number of iterations and burn-in: 100,000 iterations; 20,000 burn-in</p> <p>Convergence statistics evaluated: Yes, Brooks Gelman Rubin diagnostic plot</p> <p>Prior distribution of d: dnorm(0, .001)</p> <p>Prior distribution for σ: sdunif(0,2)</p> <p>Were priors justified: NR and did not specify if vague prior used or not</p> <p>Was a sensitivity analysis conducted based on priors? No</p>	<p>Heterogeneity assessment in traditional meta-analysis: NR</p> <p>Heterogeneity assessment in Bayesian meta-analysis: NR</p> <p>Evaluation of inconsistency: State that direct and mixed treatment comparison were largely consistent for most data although do not report methods to determine consistency</p> <p>Equivalence claims: NR</p> <p>Non-inferiority claims: NR</p> <p>Minimally important difference defined: NR</p>

Author, year	Meta-analysis model	MCMC details and Priors	Heterogeneity, inconsistency, and claims of equivalence or non-inferiority
Coleman, 2008	Was traditional meta-analysis run? Yes	Was the code available? No	Heterogeneity assessment in traditional meta-analysis: Cochrane Q statistic
	Model(s): Random effects	Was the raw data available? Yes, in the manuscript	Heterogeneity assessment in Bayesian meta-analysis: NR
	Adjustment for multiple arms: Yes	Starting values: NR	Evaluation of inconsistency: Compare results of network and traditional meta-analyses
	Adjustment for covariates: NR	Number of chains: NR	Equivalence claims: NR
	Model fit tested: NR	Number of iterations and burn-in: NR	Non-inferiority claims: NR
	Was there graphical representation of the posterior distribution? No	Convergence statistics evaluated: NR	Minimally important difference defined: NR
	Did authors rank order interventions? NR	Prior distribution of d: NR	
	Software used: WinBUGS	Prior distribution for σ : NR	
		Were priors justified: NR and did not specify if vague priors used or not	
		Was a sensitivity analysis conducted based on priors? No	

Author, year	Meta-analysis model	MCMC details and Priors	Heterogeneity, inconsistency, and claims of equivalence or non-inferiority
Mauri, 2008	<p>Was traditional meta-analysis run? Yes</p> <p>Model(s): Random effects</p> <p>Adjustment for multiple arms: NR</p> <p>Adjustment for covariates: NR</p> <p>Model fit tested: NR</p> <p>Was there graphical representation of the posterior distribution? No</p> <p>Did authors rank order interventions? No</p> <p>Software used: WinBUGS 1.4</p>	<p>Was the code available? No</p> <p>Was the raw data available? Yes, in the online supplement</p> <p>Starting values: NR</p> <p>Number of chains: NR</p> <p>Number of iterations and burn-in: NR</p> <p>Convergence statistics evaluated: NR</p> <p>Prior distribution of d: NR</p> <p>Prior distribution for σ: NR</p> <p>Were priors justified: NR and did not specify if vague priors used or not</p> <p>Was a sensitivity analysis conducted based on priors? No</p>	<p>Heterogeneity assessment in traditional meta-analysis: I^2</p> <p>Heterogeneity assessment in Bayesian meta-analysis: Tau^2</p> <p>Evaluation of inconsistency: Estimated incoherence in each closed loop, none found except one loop with modest incoherence</p> <p>Equivalence claims: NR</p> <p>Non-inferiority claims: NR</p> <p>Minimally important difference defined: NR</p>

Author, year	Meta-analysis model	MCMC details and Priors	Heterogeneity, inconsistency, and claims of equivalence or non-inferiority
Stettler, 2008	<p>Was traditional meta-analysis run? Yes</p> <p>Model(s): Random effects</p> <p>Adjustment for multiple arms: NR</p> <p>Adjustment for covariates: Yes, trial covariates</p> <p>Model fit tested: Yes, residual deviance</p> <p>Was there graphical representation of the posterior distribution? No</p> <p>Did authors rank order interventions? No</p> <p>Software used: WinBUGS 1.4.1</p>	<p>Was the code available? No</p> <p>Was the raw data available? Yes, referral to online appendix</p> <p>Starting values: NR</p> <p>Number of chains: NR</p> <p>Number of iterations and burn-in: 100,000 iterations; 60,000 burn-in</p> <p>Convergence statistics evaluated: Yes, Gelman Rubin statistic</p> <p>Prior distribution of d: N~(0,1000)</p> <p>Prior distribution for σ: NR</p> <p>Were priors justified: NR</p> <p>Was a sensitivity analysis conducted based on priors? No</p>	<p>Heterogeneity assessment in traditional meta-analysis: NR</p> <p>Heterogeneity assessment in Bayesian meta-analysis: Tau²</p> <p>Evaluation of inconsistency: Calculated inconsistency factors- the estimated difference between the log hazard ratios from direct comparisons within randomized trials and the log hazard ratios from indirect comparisons between randomized trials in common. Compare results from traditional and network meta-analyses.</p> <p>Equivalence claims: NR</p> <p>Non-inferiority claims: NR</p> <p>Minimally important difference defined: NR</p>

Author, year	Meta-analysis model	MCMC details and Priors	Heterogeneity, inconsistency, and claims of equivalence or non-inferiority
Golfinopoulos, 2007	<p>Was traditional meta-analysis run? Yes</p> <p>Model(s): NR</p> <p>Adjustment for multiple arms: NR</p> <p>Adjustment for covariates: NR</p> <p>Model fit tested: NR</p> <p>Was there graphical representation of the posterior distribution? No</p> <p>Did authors rank order interventions? Yes</p> <p>Software used: WinBUGS + S-Plus</p>	<p>Was the code available? No</p> <p>Was the raw data available? Yes, in the manuscript</p> <p>Starting values: NR</p> <p>Number of chains: NR</p> <p>Number of iterations and burn-in: NR</p> <p>Convergence statistics evaluated: NR</p> <p>Prior distribution of d: NR</p> <p>Prior distribution for σ: NR</p> <p>Were priors justified: NR and do not specify of vague priors used or not</p> <p>Was a sensitivity analysis conducted based on priors? No</p>	<p>Heterogeneity assessment in traditional meta-analysis: I^2</p> <p>Heterogeneity assessment in Bayesian meta-analysis: NR</p> <p>Evaluation of inconsistency: "No direct evidence was available and evaluation of incoherence was therefore impossible" method</p> <p>Equivalence claims: NR</p> <p>Non-inferiority claims: NR</p> <p>Minimally important difference defined: NR</p>

Author, year	Meta-analysis model	MCMC details and Priors	Heterogeneity, inconsistency, and claims of equivalence or non-inferiority
Lam, 2007	Was traditional meta-analysis run? Yes	Was the code available? No	Heterogeneity assessment in traditional meta-analysis: Chi ² /Cochran Q test, I ² , L'Abbe plots
	Model(s): Random effects	Was the raw data available? No	Heterogeneity assessment in Bayesian meta-analysis: NR
	Adjustment for multiple arms: Yes	Starting values: NR	Evaluation of inconsistency: Present results of traditional and network analyses together showing consistency of data
	Adjustment for covariates: NR	Number of chains: 3	Equivalence claims: NR
	Model fit tested: NR	Number of iterations and burn-in: 55,000 iterations, 5,000 burn-in	Non-inferiority claims: NR
	Was there graphical representation of the posterior distribution? No	Convergence statistics evaluated: NR	Minimally important difference defined: NR
	Did authors rank order interventions? Yes	Prior distribution of d: Normal (0, 10,000)	
	Software used: WinBUGS 1.4.1	Prior distribution for σ : Uniform (0, 2)	
		Were priors justified: Yes, "To ensure that overall effects were dominated by data from the trials and not influenced by choice of initial distribution we used low information (noninformative) prior distributions."	
		Was a sensitivity analysis conducted based on priors? Yes	

Author, year	Meta-analysis model	MCMC details and Priors	Heterogeneity, inconsistency, and claims of equivalence or non-inferiority
Nixon, 2007	<p>Was traditional meta-analysis run? No</p> <p>Model(s): Random effects</p> <p>Adjustment for multiple arms: Yes</p> <p>Adjustment for covariates: Yes, model includes average disease duration and average baseline HAQ for each study</p> <p>Model fit tested: NR</p> <p>Was there graphical representation of the posterior distribution? No</p> <p>Did authors rank order interventions? No</p> <p>Software used: WinBUGS 1.4.3</p>	<p>Was the code available? No</p> <p>Was the raw data available? Yes, in manuscript</p> <p>Starting values: NR</p> <p>Number of chains: NR</p> <p>Number of iterations and burn-in: NR</p> <p>Convergence statistics evaluated: NR</p> <p>Prior distribution of d: NR</p> <p>Prior distribution for σ: NR</p> <p>Were priors justified: NR and do not specify if vague priors used or not</p> <p>Was a sensitivity analysis conducted based on priors? No</p>	<p>Heterogeneity assessment in traditional meta-analysis: NA</p> <p>Heterogeneity assessment in Bayesian meta-analysis: NR</p> <p>Evaluation of inconsistency: NR</p> <p>Equivalence claims: NR</p> <p>Non-inferiority claims: NR</p> <p>Minimally important difference defined: NR</p>

Author, year	Meta-analysis model	MCMC details and Priors	Heterogeneity, inconsistency, and claims of equivalence or non-inferiority
Cooper, 2006	Was traditional meta-analysis run? No	Was the code available? Yes, in the appendix	Heterogeneity assessment in traditional meta-analysis: NA
	Model(s): Random effects	Was the raw data available? Yes, in the manuscript	Heterogeneity assessment in Bayesian meta-analysis: NR
	Adjustment for multiple arms: Yes	Starting values: NR	Evaluation of inconsistency: Generic comparison to previously conducted meta-analyses
	Adjustment for covariates: NR	Number of chains: NR	Equivalence claims: NR
	Model fit tested: Yes, residual deviance	Number of iterations and burn-in: NR	Non-inferiority claims: NR
	Was there graphical representation of the posterior distribution? No	Convergence statistics evaluated: NR	Minimally important difference defined: NR
	Did authors rank order interventions? No	Prior distribution of d : Uniform (-10,10)	
	Software used: NR	Prior distribution for σ : Uniform (0,2)	
		Were priors justified: No, all priors were intended to be vague	
		Was a sensitivity analysis conducted based on priors? No	

Author, year	Meta-analysis model	MCMC details and Priors	Heterogeneity, inconsistency, and claims of equivalence or non-inferiority
Kyrgiou, 2006	Was traditional meta-analysis run? Yes	Was the code available? No	Heterogeneity assessment in traditional meta-analysis: I^2
	Model(s): Random effects	Was the raw data available? Yes, in manuscript	Heterogeneity assessment in Bayesian meta-analysis: NR
	Adjustment for multiple arms: NR	Starting values: NR	Evaluation of inconsistency: Calculated incoherence values
	Adjustment for covariates: NR	Number of chains: NR	Equivalence claims: NR
	Model fit tested: NR	Number of iterations and burn-in: NR	Non-inferiority claims: NR
	Was there graphical representation of the posterior distribution? No	Convergence statistics evaluated: NR	Minimally important difference defined: NR
	Did authors rank order interventions? Yes	Prior distribution of d: NR	
	Software used: WinBUGS	Prior distribution for σ : NR	
		Were priors justified: NR and did not specify if vague priors used or not	
		Was a sensitivity analysis conducted based on priors? No	

Abbreviations: DIC=deviance information criterion; HTA=health technology assessment; MCMC=Markov-chain Monte Carlo; MTC=mixed treatment comparison; NA=not applicable; NR=not reported; OR=odds ratio; SD=standard deviation

*: Includes both a Bayesian MTC model and a Frequentist MTC model therefore appears in both tables.

Table F-4. Reporting of outcomes in Bayesian mixed treatment comparisons

Author, year (N authors)	Outcome type	Measure of effect	Measure of variance	Mean or median of distribution	Presentation of results
Baldwin, 2011*	Binary	OR	95% CrI	NR	Text and table
Bangalore, 2011	Binary	OR	95% CrI	NR	Table
Gross, 2011	Continuous	WMD	95% CrI	NR	Text and table
Hartling, 2011	Binary	OR	95% CrI	NR	Text and figure
	Continuous	WMD	95% CrI	NR	Text and figure
Maund, 2011	Binary	OR	95% CrI	NR	Text, table, and figure
	Continuous	WMD	95% CrI	NR	Text, table, and figure
Sciarretta, 2011	Binary	OR	95% CrI	Median	Text, table, and figure
Trelle, 2011	Binary	RR	95% CrI	Median	Text and figure
van de Kerkhof, 2011	Binary	RR	95% CrI	NR	Text and figure
	Continuous	WMD	95% CrI	NR	Text and figure
Van den Bruel, 2011	Continuous	WMD	95% CrI	NR	Table
Dakin, 2010	Binary	OR	95% CrI	NR	Text, table, and figure
Orme, 2010	Binary	OR	SE	NR	Table
	Continuous	WMD	95% CrI and SE	NR	Text and table
Phung, 2010	Binary	RR	95% CrI	NR	Text, table, and figure
	Continuous	WMD	95% CrI	NR	Text, table, and figure
Uthman, 2010	Binary	RR	95% CrI	NR	Text and figure
Vissers, 2010	Continuous	WMD	95% CrI	NR	Text and figure
Walsh, 2010	Continuous	SMD and prevention fraction	95% CrI	NR	Text, table, and figure
Wandel, 2010	Binary	OR	95% CrI	NR	Text
	Continuous	WMD and SMD	95% CrI	NR	Text and figure
Wang, 2010	Binary	OR	95% CrI	NR	Text, table, and figure
Woo, 2010	Binary	OR	95% CrI	Median	Text and table
Baker, 2009	Binary	OR	95% CrI	NR	Text, table, and figure
Bansback, 2009	Categorical non-binary	RR	95% CrI	NR	Text, table, figure
Cipriani, 2009	Binary	OR	95% CrI	NR	Text, table, figure
Edwards, 2009a	Binary	OR	95% CrI	NR	Text, table, and figure
Edwards, 2009b	Binary	OR	95% CrI	NR	Table
Golfinopoulos, 2009	Binary	HR	95% CrI	NR	Text and table
Manzoli, 2009	Binary	OR and RD	95% CrI	NR	Table
Meader, 2009	Binary	OR	95% CrI	NR	Text and table
Coleman, 2008	Binary	OR	95% CrI	NR	Table and figure
Mauri, 2008	Binary	HR	95% CrI	NR	Text and table
Stettler, 2008	Binary	HR	95% CrI	Median	Text, table, and figure
Golfinopoulos, 2007	Binary	HR	95% CrI	NR	Text and table
Lam, 2007	Binary	OR	95% CrI	Mean	Text and figure
Nixon, 2007	Binary	OR	95% CrI	NR	Text, table, and figure

Author, year (N authors)	Outcome type	Measure of effect	Measure of variance	Mean or median of distribution	Presentation of results
Cooper, 2006	Binary	RR	95% CrI	NR	Text and figure
Kyrgiou, 2006	Binary	RR and HR	95% CrI	NR	Text and table

Abbreviations: CrI=credible interval; NR=not reported; OR=odds ratio; RD=risk difference; RR=relative risk

* Includes both a Bayesian MTC model and a Frequentist MTC model therefore appears in both tables.

Table F-5. Characteristics of frequentist mixed treatment comparisons

Author, year (N authors)	Method- ologist*	Country	Funding	# printed pages	Affiliation	Supplement or appendix	Disease state evaluated	N and type of interventions compared	N trials, N patients	Network pattern
Anothais- intawee, 2011 (7)	Yes	Thailand	Government/ foundation	9	No	Yes	Genitourinary (Chronic prostatitis)	9, Rx	23 NR	Network with ≥1 closed loop
Baldwin, 2011 (4) [†]	Yes	UK	Industry	11	No	Yes	Behavioral Health (GAD)	10, Rx	27 3,989	Network with ≥1 closed loop
Freemantle, 2011 (5)	No	France	Industry	17	No	Yes	Cardiology (AF)	5, Rx	39 174,662	Network with ≥1 closed loop [‡]
Singh, 2011 (20)	Yes	USA	Other	58	Yes, Cochrane	No	Rheumatology (Biologics)	9, Rx	163 50,010	Network with ≥1 closed loop
Roskell, 2009 (5)	No	UK	Industry	10	No	Yes	Cardiology (AF)	12, Rx	21 NR	Network with ≥1 closed loop
Trikalinos, 2009 (5)	No	USA	Government/ foundation	8	No	Yes	Cardiology (Stents)	4, Procedure, device and Rx	61 25,388	Network with ≥1 closed loop
Hansen, 2008 (6)	Yes	USA	Government/ foundation	10	HTA	Yes	Behavioral Health (Social anxiety disorder)	7, Rx	18 5,172	Network with ≥1 closed loop [‡]
Elliot, 2007 (2)	No	USA	Government/ foundation	7	No	No	Cardiology (anti- hypertensives)	6, Rx	22 143,513	Network with ≥1 closed loop
Eckert, 2006 (2)	No	France	Unknown	15	No	No	Behavioral Health (MDD)	4, Rx	39 14,573	Network with ≥1 closed loop

Abbreviations: AF=atrial fibrillation; GAD=generalized anxiety disorder; HTA=health technology assessment; MDD=major depressive disorder; NR=not reported; Rx=pharmacologic; UK=United Kingdom

*: A methodologist was considered an individual with affiliation to a department of statistics, biostatistics, epidemiology, clinical epidemiology, or public health services, as determined by author information and affiliations listed in the publication.

†: Includes both a Bayesian MTC model and a Frequentist MTC model therefore appears in both tables.

Table F-6. Methodological characteristics of Frequentist mixed treatment comparisons

Author, year	Network model characteristics	Measure of heterogeneity, inconsistency and claims of equivalence or non-inferiority
Anothaisin-tawee, 2011	<p>Was traditional meta-analysis run? Yes</p> <p>Model(s): Mixed-effect hierarchical model with a log-link function using the “xtpoisson” command</p> <p>Weighting of studies: Inverse variance</p> <p>Adjustment for covariates: Yes, effects of study were included as covariates</p> <p>Was the raw data available? Yes, in manuscript</p> <p>Software used: Stata 11.0</p>	<p>Heterogeneity assessment in traditional meta-analysis: Cochrane Q-statistic, I^2</p> <p>Heterogeneity assessment in network meta-analysis: NR</p> <p>Evaluation of inconsistency: Compare results from traditional and network meta-analyses</p> <p>Equivalence claims: NR</p> <p>Non-inferiority claims: NR</p> <p>Minimally important difference defined: NA</p>
Baldwin, 2011*	<p>Was traditional meta-analysis run? Yes</p> <p>Model(s): Frequentist framework using random effects</p> <p>Weighting of studies: NR</p> <p>Adjustment for covariates: NR</p> <p>Was the raw data available? No</p> <p>Software used: Stata 9</p>	<p>Heterogeneity assessment in traditional meta-analysis: NR</p> <p>Heterogeneity assessment in network meta-analysis: NR</p> <p>Evaluation of inconsistency: Test for consistency between results of the direct meta-analysis and those of the mixed treatment meta-analyses by subtracting the odds ratios and using a t-test to identify differences in effect estimates between the two models</p> <p>Equivalence claims: NR</p> <p>Non-inferiority claims: NR</p> <p>Minimally important difference defined: NA</p>

Author, year	Network model characteristics	Measure of heterogeneity, inconsistency and claims of equivalence or non-inferiority
Freemantle, 2011	<p>Was traditional meta-analysis run? Yes</p> <p>Model(s): Random effects, non-linear mixed model based upon psuedolikelihood</p> <p>Weighting of studies: NR</p> <p>Adjustment for covariates: NR</p> <p>Was the raw data available? Yes, in manuscript</p> <p>Software used: SAS</p>	<p>Heterogeneity assessment in traditional meta-analysis: NR</p> <p>Heterogeneity assessment in network meta-analysis: Covariance statistic and SE</p> <p>Evaluation of inconsistency: Compare results from traditional and network meta-analyses</p> <p>Equivalence claims: NR</p> <p>Non-inferiority claims: NR</p> <p>Minimally important difference defined: NA</p>
Singh, 2011	<p>Was traditional meta-analysis run? Yes</p> <p>Model(s): Bayes Framework</p> <p>Weighting of studies: NR</p> <p>Adjustment for covariates: NR</p> <p>Was the raw data available? Yes, in report</p> <p>Software used: NR</p>	<p>Heterogeneity assessment in traditional meta-analysis: NR</p> <p>Heterogeneity assessment in network meta-analysis: Tau²</p> <p>Evaluation of inconsistency: NR</p> <p>Equivalence claims: NR</p> <p>Non-inferiority claims: NR</p> <p>Minimally important difference defined: NA</p>

Author, year	Network model characteristics	Measure of heterogeneity, inconsistency and claims of equivalence or non-inferiority
Roskell, 2009	<p>Was traditional meta-analysis run? No</p> <p>Model(s): Bayes Framework</p> <p>Weighting of studies: NR</p> <p>Adjustment for covariates: Length of follow-up</p> <p>Was the raw data available? Yes, in online appendix</p> <p>Software used: SAS</p>	<p>Heterogeneity assessment in traditional meta-analysis: NA</p> <p>Heterogeneity assessment in network meta-analysis: NR</p> <p>Evaluation of inconsistency: Compare results from MTC to previously published literature</p> <p>Equivalence claims: NR</p> <p>Non-inferiority claims: NR</p> <p>Minimally important difference defined: NA</p>
Trikalinos, 2009	<p>Was traditional meta-analysis run? Yes</p> <p>Model(s): Two level linear mixed-effects model with heteroscedastic errors</p> <p>Weighting of studies: NR</p> <p>Adjustment for covariates: NR</p> <p>Was the raw data available? Yes, in online appendix</p> <p>Software used: R 2.6.0 nlme package</p>	<p>Heterogeneity assessment in traditional meta-analysis: I^2</p> <p>Heterogeneity assessment in network meta-analysis: NR</p> <p>Evaluation of inconsistency: Measured and reported network incoherence values</p> <p>Equivalence claims: NR</p> <p>Non-inferiority claims: NR</p> <p>Minimally important difference defined: NA</p>

Author, year	Network model characteristics	Measure of heterogeneity, inconsistency and claims of equivalence or non-inferiority
Hansen, 2008	<p>Was traditional meta-analysis run? Yes</p> <p>Model(s): Frequentist mixed-effects meta-regression</p> <p>Weighting of studies: NR</p> <p>Adjustment for covariates: NR</p> <p>Was the raw data available? Yes, in online appendix</p> <p>Software used: R code using Metafor package</p>	<p>Heterogeneity assessment in traditional meta-analysis: I^2</p> <p>Heterogeneity assessment in network meta-analysis: NR</p> <p>Evaluation of inconsistency: Compare results from network meta-analysis to previously published literature</p> <p>Equivalence claims: NR</p> <p>Non-inferiority claims: NR</p> <p>Minimally important difference defined: NA</p>
Elliot, 2007	<p>Was traditional meta-analysis run? Yes</p> <p>Model(s): "online program published by Lumely"</p> <p>Weighting of studies: NR</p> <p>Adjustment for covariates: NR</p> <p>Was the raw data available? Yes, in manuscript</p> <p>Software used: R 1.14 framework 2.21</p>	<p>Heterogeneity assessment in traditional meta-analysis: Riley-Day test</p> <p>Heterogeneity assessment in network meta-analysis: NR</p> <p>Evaluation of inconsistency: Measured and reported incoherence values</p> <p>Equivalence claims: NR</p> <p>Non-inferiority claims: NR</p> <p>Minimally important difference defined: NA</p>

Author, year	Network model characteristics	Measure of heterogeneity, inconsistency and claims of equivalence or non-inferiority
Eckert, 2006	Was traditional meta-analysis run? Yes	Heterogeneity assessment in traditional meta-analysis: NR
	Model(s): Bayes Framework	Heterogeneity assessment in network meta-analysis: NR
	Weighting of studies: NR	Evaluation of inconsistency: Compare results from MTC to previously published literature
	Adjustment for covariates: NR	Equivalence claims: NR
	Was the raw data available? Yes, in manuscript	Non-inferiority claims: NR
	Software used: SAS	Minimally important difference defined: NA

Abbreviations: NA= not applicable; NR=not reported; SE=standard error

*: Includes both a Bayesian MTC model and a Frequentist MTC model therefore appears in both tables.

Table F-7. Reporting of outcomes in frequentist mixed treatment comparisons

Author, year (N authors)	Outcome type	Measure of effect	Measure of variance	Presentation of results
Baldwin, 2011*	Binary	OR	95% CI	Text and table
Freemantle, 2011	Binary	OR	95% CI	Text and figure
Anothaisintawee, 2001	Continuous	WMD	95% CI	Text
Singh, 2011	Binary	OR	95% CI	Text and table
Roskell, 2009	Binary	RR	95% CI	Text and figure
Trikalinos, 2009	Binary	RR	95% CI	Text, table, and figure
Hansen, 2008	Binary	Relative benefit	95% CI	Figure
Elliott, 2007	Binary	OR	95% CI	Text and figure
Eckert, 2006	Binary	Log OR	95% CI	Text and figure
	Continuous	Standardized effect size	95% CI	Text and figure

Abbreviations: CI=confidence interval; NR=not reported; OR=odds ratio; RR=relative risk; WMD=weighted-mean difference

*: Includes both a Bayesian MTC model and a Frequentist MTC model therefore appears in both tables.

References for Evidence Tables

1. Bangalore S, Kumar S, Kjeldsen E et al. Antihypertensive drugs and risk of cancer: network meta-analyses and trial sequential analysis of 324168 participants from randomized trials. *Lancet Oncol* 2011;12:65-82. PMID:21123111
2. Gross JL, Kramer CK, Leitao CB, et al. Effect of antihyperglycemics agents added to metformin and a sulfonyleurea on glycemc control and weight gain in type 2 diabetes: a network meta-analysis. *Ann Intern Med* 2011;154(10):672-679. PMID:21576535
3. Hartling L, Fernandes RM, Bialy L, et al Steroids and bronchodilators for acute bronchiolitis in the first two years of life: systematic review and meta-analysis. *BMJ* 2011;342:1714. PMID:21471175
4. Maund E, McDaid C, Rice S, Wright K, Jenkins B, Woolacott N. Paracetamol and selective and non-selective non-steroidal anti-inflammatory drugs for the reduction in morphine-related side-effects after major surgery: a systematic review. *Br J Anaesth* 2011;106(3):292-297. PMID:21285082
5. McDaid C, Maund E, Rice S, Wright K, Jenkins B, Woolacott N. Paracetamol and selective and non-selective non-steroidal anti-inflammatory drugs (NSAIDs) for the reduction of morphine-related side effects after major surgery: a systematic review. *Health Technol Assess* 2010;14(17):1-153. PMID:20346263
6. Sciarretta S, Palano F, Tocci G, Baldini R, Volpe M. Antihypertensive treatment and development of heart failure in hypertension: a Bayesian network meta-analysis of studies in patients with hypertension and high cardiovascular risk. *Arch Intern Med* 2011;171(5):384-394. PMID:21059964
7. Trelle S, Reichenback S, Wandel S, et al. Cardiovascular safety of non-steroidal anti-inflammatory drugs: network meta-analysis. *BMJ* 2011;342:c7086. PMID:21224324
8. van de Kerkhof P, de Peuter R, Rytto J, Jansen JP. Mixed treatment comparison of a two-compound formulation (TCF) product containing calcipotriol and betamethasone dipropionate with other topical treatments in psoriasis vulgaris. *Curr Med Res Opin* 2011;27(1):225-238. PMID:21142833
9. Van den Bruel A, Gailly J, Devriese S, Welton NJ, Shortt AJ, Vrijens F. The protective effect of ophthalmic viscoelastic devices on endothelial cell loss during cataract surgery: a meta-analysis using mixed treatment comparisons. *Br J Ophthalmol* 2011;95(1):5-10. PMID:19965836
10. Dakin H, Fidler C, Harper C. Mixed treatment comparison meta-analysis evaluating the relative efficacy of nucleos(t)ides for treatment of nucleos(t)ide-naive patients with chronic hepatitis B. *Value Health* 2010;13(8):934-945. PMID:20825624
11. Orme M, Collins S, Dakin H, Kelly S, Loftus J. Mixed treatment comparison and meta-regression of the efficacy and safety of prostaglandin analogues and comparators for primary open-angle glaucoma and ocular hypertension. *Curr Med Res Opin* 2010;26(3):511-528. PMID:20014995
12. Phung OJ, Scholle JM, Talwar M, Coleman CI. Effect of noninsulin antidiabetic drugs added to metformin therapy on glycemc control, weight gain, and hypoglycemia in type 2 diabetes. *JAMA* 2010;303(14):1410-1418. PMID:20388897

13. Uthman OA, Abdulmalik J. Comparative efficacy and acceptability of pharmacotherapeutic agents for anxiety disorders in children and adolescents: a mixed treatment comparison meta-analysis. *Curr Med Res Opin* 2010;26(1):53-59. PMID:19905879
14. Vissers D, Stam W, Nolte T, Lenre M, Jansen J. Efficacy of intranasal fentanyl spray versus other opioids for breakthrough pain in cancer. *Curr Med Res Opin* 2010;26(5):1037-1045. PMID:20199140
15. Walsh T, Worthington HV, Glenny A, Appelbe P, Marinho CCV, Shi X. Fluoride toothpastes of different concentrations for preventing dental caries in children and adolescents. *Cochrane Database of Systematic Reviews* 2011;1 Art No CD007868. PMID:20091655
16. Wandel S, Juni P, Tendal B, Nuesch E, Villiger PM, Welton NJ, et al. Effects of glucosamine, chondroitin, or placebo in patients with osteoarthritis of hip or knee: network meta-analysis. *BMJ* 2010;341:4675. PMID:20847017
17. Wang H, Huang T, Jing J, Jin J, Wang P, Yang M, et al. Effectiveness of different central venous catheters for catheter-related infections: a network meta-analysis. *J Hosp Infect* 2010;76(1):1-11. PMID:20638155
18. Woo G, Tomlinson G, Nishikawa Y, Kowgier M, Sherman M, Wong DK, et al. Tenofovir and entecavir are the most effective antiviral agents for chronic hepatitis B: a systematic review and Bayesian meta-analyses. *Gastroenterology* 2010;139(4):1218-1229. PMID:20600036
19. Baker WL, Baker EL, Coleman CI. Pharmacologic treatments for chronic obstructive pulmonary disease: a mixed-treatment comparison meta-analysis. *Pharmacotherapy* 2009;29(8):891-905. PMID:19637942
20. Bansback N, Sizto S, Sun H, Feldman S, Willian MK, Anis A. Efficacy of systemic treatments for moderate to severe plaque psoriasis: systematic review and meta-analysis. *Dermatology* 2009;219(3):209-218. PMID:19657180
21. Cipriani A, Furukawa TA, Salanti G, Geddes JR, Higgins JP, Churchill R, et al. Comparative efficacy and acceptability of 12 new-generation antidepressants: a multiple-treatments meta-analysis. *Lancet* 2009;373(9665):746-758. PMID:19185342
22. Edwards SJ, Lind T, Lundell L, Das R. Systematic review: standard- and double-dose proton pump inhibitors for the healing of severe erosive oesophagitis -- a mixed treatment comparison of randomized controlled trials. *Aliment Pharmacol Ther* 2009;30(6):547-556. PMID:19558609
23. Edwards SJ, Smith CJ. Tolerability of atypical antipsychotics in the treatment of adults with schizophrenia or bipolar disorder: a mixed treatment comparison of randomized controlled trials. *Clin Ther* 2009;31(Pt 1):1345-1359. PMID:19698898
24. Golfiopoulos V, Pentheroudakis G, Salanti G, Nearchou AD, Ioannidis JP, Pavlidis N. Comparative survival with diverse chemotherapy regimens for cancer of unknown primary site: multiple-treatments meta-analysis. *Cancer Treat Rev* 2009;35(7):570-573. PMID:19539430
25. Manzoli L, Georgia S, De Vito C, Boccia A, Ionnidis JPA, Villari P. Immunogenicity and adverse events of avian influenza A H5N1 vaccine in healthy adults: multiple-treatments meta-analysis. *Lancet Infect Dis* 2009;9:481-92. PMID:19628173

26. Meader N. A comparison of methadone, buprenorphine and alpha2 adrenergic agonists for opioid detoxification: a mixed treatment comparison meta-analysis. *Drug and Alcohol Dependence* 2010;108:110-114. PMID:20074867
27. Coleman CI, Baker WL, Kluger J, White CM. Antihypertensive medication and their impact on cancer incidence: a mixed treatment comparison meta-analysis of randomized controlled trials. *J Hypertens* 2008;26(4):622-629. PMID:18327066
28. Mauri D, Polyzos NP, Salanti G, Pavlidis N, Ionnidis JPA. Multiple-treatments meta-analysis of chemotherapy and targeted therapies in advanced breast cancer. *J Natl Cancer Inst* 2008;100:1780-1791. PMID:19066278
29. Stettler C, Allemann S, Wandel S, et al. Drug eluting and bare metal stents in people with and without diabetes: collaborative network meta-analysis. *BMJ* 2008;337:a1331. PMID:18757996
30. Golfinopoulos V, Salanti G, Pavlidis N, Ioannidis JP. Survival and disease-progression benefits with treatment regimens for advanced colorectal cancer: a meta-analysis. *Lancet Oncol* 2007;8(10):898-911. PMID:17888735
31. Lam SK, Owen A. Combined resynchronisation and implantable defibrillator therapy in left ventricular dysfunction: Bayesian network meta-analysis of randomised controlled trials. *BMJ* 2007;335(7626):925. PMID:17932160
32. Nixon R, Bansback N, Brennan A. The efficacy of inhibiting tumour necrosis factor alpha and interleukin 1 in patients with rheumatoid arthritis: a meta-analysis and adjusted indirect comparisons. *Rheumatology* 2007;46(7):1140-1147. PMID:17478472
33. Cooper NJ, Sutton AJ, Lu G, Khunti K. Mixed comparison of stroke prevention treatments in individuals with nonrheumatic atrial fibrillation. *Arch Intern Med* 2006;166(12):1269-1275. PMID:16801509
34. Kyrgiou M, Salanti G, Pavlidis N, Paraskevaidis E, Ioannidis JP. Survival benefits with diverse chemotherapy regimens for ovarian cancer: meta-analysis of multiple treatments. *J Natl Cancer Inst* 2006;98(22):1655-1663. PMID:17105988
35. Baldwin D, Woods R, Lawson R, Taylor D. Efficacy of drug treatments for generalised anxiety disorder: systematic review and meta-analysis. *BMJ* 2011;342:1199. PMID:21398351
36. Freemantle N, Lafuente-Lafuente C, Mitchell S, Eckert L, Reynolds M. Mixed treatment comparison of dronedarone, amiodarone, sotalol, flecainide, and propafenone, for the management of atrial fibrillation. *Europace* 2011;13(3):329-345. PMID:21227948
37. Anothaisintawee T, Attia J, Nickel JC, et al. Management of chronic prostatitis/chronic pelvic pain syndrome: a systematic review and network meta-analysis. *JAMA* 2011;305(1):78-86. PMID:21205969
38. Hansen RA, Gaynes BN, Gartlehner G, Moore CG, Tiwari R, Lohr KN. Efficacy and tolerability of second-generation antidepressants in social anxiety disorder. *Int Clin Psychopharmacol* 2008;23(3):170-179. PMID:18408531
39. Trikalinos TA, Alsheikh-Ali AA, Tatsioni A, Nallamothu BK, Kent DM. Percutaneous coronary interventions for non-acute coronary artery disease: a quantitative 20-year synopsis and a network meta-analysis. *Lancet* 2009;373(9667):911-918. PMID:19286090
40. Elliott WJ, Meyer PM. Incident diabetes in clinical trials of antihypertensive drugs: a network meta-analysis. *Lancet* 2007;369(9557):201-207. PMID:17240286

Appendix G. Glossary

Closed loop network of evidence: A network of evidence where greater than two interventions are being compared indirectly, and at least one pair of interventions is being compared both directly and indirectly.

Lumley's network meta-analysis approach: A Frequentist approach to conduct a MTC originally described by Lumley et al. whereby both direct and indirect evidence are combined when there is at least one closed loop of evidence connecting two interventions of interest using a mixed model

Meta-Analysis: The process of extracting and pooling data from several studies investigating a similar topic to synthesize a final outcome

Mixed treatment comparison (MTC): A statistical approach used to analyze a network of evidence with more than two interventions which are being compared indirectly, and at least one pair of interventions compared both directly and indirectly

Network meta-analysis: The simultaneous synthesis of evidence of all pairwise comparisons across more than two interventions

Markov chain Monte Carlo (MCMC) methods: Simulation-based methods which can be used for the analysis of complex statistical models and to obtain estimates from distributions.