Knowledge Translation of Economic Evaluations and Network Meta-Analyses

Shannon Sullivan

Thesis submitted to the
Faculty of Graduate and Postdoctoral Studies
in partial fulfillment of the requirements
for the Doctorate in Philosophy degree in Epidemiology

Department of Epidemiology and Community Medicine
Faculty of Medicine
University of Ottawa

© Shannon Sullivan, Ottawa, Canada, 2015

ABSTRACT

Economic evaluations and network meta-analyses (NMAs) are complex methodologies. Increasing their transparency and accessibility could enhance confidence in the legitimacy of policy decisions informed by these analyses.

Four systematic reviews were conducted to understand policymakers' informational needs and to determine what guidance researchers have on how to present economic evaluations and NMAs. Qualitative interviews were conducted with Canadian policymakers, i.e., knowledge users, to understand barriers and facilitators to using and communicating economic evaluations and NMAs and with individuals in international health technology assessment organizations, i.e. knowledge producers, to explore current approaches to communicating economic evaluations and NMAs. A toolkit for NMAs and economic evaluations was proposed based on an integrated review of these findings and guided by the Knowledge-to-Action framework. Examples of tools were developed and applied to an economic evaluation and NMA of osteoporosis therapies.

Systematic reviews and qualitative interviews found that communication approaches that provide robust content, identify contextual factors relevant to policy decisions and enhance clarity were valued. Twelve tools were proposed that enhance communication, education and access to resources for policymakers. Two of these tools were developed: Economic Guidance for Researchers and NMA Guidance for Researchers.

AUTHORIZATIONS

Written permission for the reproduction of the following figures and tables in this thesis has been obtained from publishers as follows:

1) The Knowledge to Action Framework (Chapter 2, Figure 1 and Chapter 7, Supplementary Figure 1) is used with permission from John Wiley and Sons.

This is a License Agreement between Shannon Sullivan ("You") and John Wiley and Sons ("John Wiley and Sons") provided by Copyright Clearance Center ("CCC"). The license consists of your order details, the terms and conditions provided by John Wiley and Sons, and the payment terms and conditions.

All payments must be made in full to CCC. For payment instructions, please see information listed at the bottom of this form.

License Number	3442180958175
License date	Aug 04, 2014
Licensed content publisher	John Wiley and Sons
Licensed content publication	Wiley Books
Licensed content title	Knowledge Translation in Health Care: Moving from Evidence to Practice, 2nd Edition
Book title	Knowledge Translation in Health Care: Moving from Evidence to Practice, 2nd Edition
Licensed copyright line	Copyright © 2013, John Wiley and Sons
Licensed content author	Sharon Straus (Editor), Jacqueline Tetroe (Editor), Ian D. Graham (Editor)
Licensed content date	Aug 1, 2013
Type of use	Dissertation/Thesis
Requestor type	University/Academic
Format	Print and electronic
Portion	Figure/table
Number of figures/tables	1
Original Wiley figure/table number(s)	Figure 1.1.1
Will you be translating?	No
Title of your thesis / dissertation	Knowledge Translation of Economic Evaluations and Network Meta-Analyses for Policymakers

2) Agency for Healthcare Research and Quality (AHRQ) Criteria for the Selection Of Effective Health Care Program Topics (Chapter 8, Table 1) is used with permission from the AHRQ.

"This document is in the public domain and may be used and reprinted without special permission. Citation of the source is appreciated."

3) AGREE II Criteria for the Critical Appraisal of Clinical Practice Guidelines (Chapter 4, Supplementary Table 1 and Chapter 6, Supplementary Table 1) is used with permission from the Agree Research Trust.

"Information may be cited with appropriate acknowledgement in scientific publications without obtaining further permissions." http://www.agreetrust.org/copyright/.

CONTRIBUTION OF THE AUTHORS

Five manuscripts (Chapters 3 to 7) were prepared as part of this thesis and are currently under review by journals. One additional manuscript is provided in draft form (Chapter 8) as part of this thesis. This will form the basis of a future publication that will include additional research done outside of the scope of this thesis; this publication strategy is being taken to ensure a robust and high impact publication.

All manuscripts are co-authored by the student (SMS) and at least one of her supervisors (George Wells and co-supervisor Doug Coyle). The student is the first author of all papers, having been primarily responsible for the conception, data collection, analysis, and writing of the manuscripts. All authors and/or thesis advisory committee members provided valuable feedback throughout the process.

Two additional manuscripts are provided in appendices to this thesis (Appendix D and E) on which the student (SMS) is a co-author. However, these are complementary to the thesis and do not represent work done as part of the thesis objectives.

TABLE OF CONTENTS

ABSTRACT..... ii
AUTHORIZATIONS..... Iii
CONTRIBUTION OF AUTHORS..... Iv
TABLE OF CONTENTS..... v
LIST OF TABLES..... viii
LIST OF FIGURES..... xiii
ABBREVIATIONS..... xv
ACKNOWLEDGEMENTS..... xvii

CHAPTER 1
Introduction..... 1
1.1 Rationale..... 1
1.2 Objectives..... 2
1.3 Chapter Outlines..... 3
1.4 References..... 7

CHAPTER 2
Background..... 8
2.1 Health Technology Assessment..... 8
2.2 Economic Evaluations..... 11
2.3 Network Meta-Analysis..... 13
2.4 Health Policy Decision-Making and Public Drug Funding..... 16
2.5 Knowledge Translation..... 19
2.6 References..... 25

CHAPTER 3
Knowledge, Barriers, Facilitators and Approaches to Communicating Economic Evaluations: A Systematic Review..... 31
Chapter Overview..... 31
Title page..... 33
Abstract..... 34
Introduction..... 36
Methods..... 37
Results..... 39
Discussion..... 42
References..... 49
Tables and Figures..... 55
Supplementary Materials. Additional Tables and Figures. 64

CHAPTER 4
Guideline Standards for Presenting Economic Evaluations: A Systematic Review 99
Chapter Overview..... 99
Title page..... 101
Abstract..... 103
Introduction..... 104
Methods..... 105
Results..... 107
Discussion..... 115
References..... 120
Tables and Figures..... 128
Supplementary Materials. Additional Tables and Figures. 136

CHAPTER 5
Knowledge, Barriers, Facilitators and Approaches to Communicating Network Meta-Analyses: A Systematic Review..... 151
Chapter Overview..... 151
Title page..... 153
Abstract..... 154
Introduction..... 155
Methods..... 156
Results..... 157
Discussion..... 167
References..... 172
Tables and Figures..... 176
Supplementary Materials. Additional Tables and Figures. 185

CHAPTER 6
Guideline Standards for Presenting Network Meta-Analyses: A Systematic Review 206
Chapter Overview..... 206
Title page..... 208
Abstract..... 209
Introduction..... 210
Methods..... 211
Results..... 213
Discussion..... 219
References..... 226
Tables and Figures..... 234
Supplementary Materials. Additional Tables and Figures. 238

CHAPTER 7
Canadian and International Perspectives on Knowledge Translation of Economic Evaluations and Network Meta-Analyses: Key Informant Interviews..... 247
Chapter Overview..... 247
Title page..... 249
Abstract..... 250
Introduction..... 251
Methods..... 252
Results..... 253
Discussion..... 261
References..... 265
Tables and Figures..... 268
Supplementary Materials. Additional Tables and Figures. 273

CHAPTER 8
Enhancing the Accessibility of Economic Evaluations and Network Meta-Analyses for Policymakers: Development of a Toolkit..... 286
Chapter Overview..... 286
Title page..... 288
Abstract..... 289
Introduction..... 290
Methods..... 291
Results..... 294
Discussion..... 299
References..... 306
Tables and Figures..... 309
Supplementary Materials 1. Additional Tables and Figures..... 316
Supplementary Materials 2. Economic Guidance for Researchers..... 341
Supplementary Materials 3. Network Meta-Analysis Guidance for Researchers..... 352

CHAPTER 9
Conclusions..... 363
9.1 Background..... 363
9.2 Discussion of Findings..... 365
9.3 Future Research and Conclusions..... 375
9.4 References..... 384

APPENDICES..... 389

Appendix A. Abstract for Oral Presentation, Cochrane Colloquium 2012..... 389

Appendix B. Ethics Approval..... 391

Appendix C. Interview Guides for Key Informant Interviews..... 394

Appendix D. Osteoporosis Economic Evaluation (Coyle et al., 2014)..... 407
Title page..... 408
Abstract..... 410
Introduction..... 411
Methods..... 412
Results..... 417
Discussion..... 418
References..... 421
Tables and Figures..... 425

Appendix E. Osteoporosis Network Meta-Analysis (Maher et al., 2014)..... 436
Title page..... 437
Abstract..... 438
Introduction..... 439
Methods..... 440
Results..... 442
Discussion..... 447
References..... 453
Tables and Figures..... 457
Supplementary Materials 1. Additional Tables and Figures..... 463
Supplementary Materials 2. Data Not Submitted for Publication..... 477

LIST OF TABLES

CHAPTER 3

Table 1.	Characteristics of Included Studies.	55
Table 2.	Themes Related to Communication and Presentation of Health Economic Information	59
Table 3.	Themes Related to Knowledge of Health Economic Information	60
Table 4.	Specific Barriers and Facilitators Related to Communicating Health Economic Information as Reported in Individual Studies	61

Supplementary Material

Supplemental Table 1.	Detailed Study Characteristics	65
Supplemental Table 2A.	Quality assessment of included systematic reviews based on AMSTAR criteria.	70
Supplemental Table 2B.	Quality assessment of included randomized controlled trials (RCTs) based on Cochrane Risk of Bias Tool	71
Supplemental Table 2C.	Quality assessment of qualitative or mixed methods studies using checklists from Greenhalgh et al. 2004	72
Supplemental Table 3.	Included studies highlighting preferences for the communication and presentation of health economic information among different target audiences.	80
Supplemental Table 4.	Included studies reporting outcomes related to knowledge and understanding of health economics among different target audiences.	84
Supplemental Table 5.	Specific Barriers and Facilitators Reported in Individual Studies and Grouped by Theme	90
Supplemental Table 6.	Types of barriers to using economic information and distribution across individual studies.	95
Supplemental Table 7.	Types of facilitators to using economic information and distribution across individual studies.	97

CHAPTER 4

Table 1.	Characteristics of Economic Evaluation Guidelines	128
Table 2.	Most Frequently (≥20%) Recommended Presentation Formats Across Economic Guidelines	134
Table 3.	Most Frequent Recommendations (≥20%) for Reporting and Conducting Economic Evaluation Guidelines	135

Supplementary Material

Supplemental Table 1.	AGREE II (Unmodified) for Critical Appraisal of Clinical Practice Guidelines	136
Supplemental Table 2.	Definitions of Variables and Data Extracted in Systematic Review	138
Supplemental Table 3.	Additional Guideline Characteristics	139
Supplemental Table 4.	Quality of Economic Guidelines as Assessed by modified AGREEII Instrument	141
Supplemental Table 5.	Advice on Presenting Economic Evaluations to Non-Economists	147
Supplemental Table 6.	Less Frequent (<20%) Reporting Recommendations in Economic Guidelines	148

CHAPTER 5

Table 1.	Characteristics of Included Studies.	176
Table 2.	Key Findings Related to Formatting Options for Communicating NMAs	177
Table 3.	Key Findings on Approaches to Communicating NMA Data Sources	178
Table 4.	Key Findings on Approaches to Communicating and Presenting NMA Assumptions	179
Table 5.	Key Findings on Approaches to Communicating NMA Results	180
Table 6.	Key Findings Related to Knowledge of NMAs	181

Supplementary Material

Supplemental Table 1A.	Quality assessment of included reviews based on AMSTAR criteria	185
Supplemental Table 1B.	Quality assessment of qualitative or mixed methods studies using checklists from Greenhalgh et al. 2004	188
Supplemental Table 1C.	Quality assessment of case studies adapted from checklists in Greenhalgh et al. 2004	190
Supplemental Table 2.	Detailed Included Study Characteristics	193
Supplemental Table 3.	Study-Level Details on Formatting Options for Presenting NMAs	195
Supplemental Table 4.	Study-Level Details on Approaches to Communicating NMA Data Sources, Assumptions and Results.	198
Supplemental Table 5.	Study-Level Details on NMA Knowledge Outcomes	201
Supplemental Table 6.	Study-Level Details of Barriers and Facilitators	202

CHAPTER 6

Table 1.	Characteristics of Network Meta-Analysis Guidelines	236
Table 2.	NMA Areas with Frequency of Presentation Formats Identified in Guidelines	237

Supplementary Material

Supplemental Table 1.	AGREE II (Unmodified) for Critical Appraisal of Clinical Practice Guidelines	238
Supplemental Table 2.	Definitions of Variables and Data Extracted in Systematic Review	240
Supplemental Table 3.	Purpose of NMA Guideline Documents	241
Supplemental Table 4.	Quality of Network Meta-Analysis Guidelines as Assessed by modified AGREEII Instrument	242
Other.	Completed PRISMA Checklist	244

CHAPTER 7

Table 1.	Characteristics of Canadian Provinces, Territories and Public Drug Programmes That are Represented in Key Informant Interviews	268
Table 2.	Overarching Themes from Canadian Key Informant Interviews, Organized by Question	269
Table 3.	Characteristics and Knowledge Translation Approaches of HTA Organizations Around the World (i.e. Knowledge Producers) That Are Represented in Key Informant Interviews	270
Table 4.	Approaches to the Knowledge Translation of Economic Evaluations and NMAs Identified by International Key Informants	272

Supplementary Material

Supplemental Box 1.	Additional Methodological Details and Considerations	274
Supplemental Box 2.	Additional Detailed Results	278
Supplemental Table 1.	Content Elements of Economic Evaluations and NMAs that are Valued by Canadian Policymakers	280
Supplemental Table 2.	Barriers and Facilitators to Canadian Policymakers' Use and Understanding of Economic Evaluations and NMAs	281
Supplemental Table 3.	Detailed Approaches to the Knowledge Translation of Economic Evaluations and NMAs Identified by International Key Informants	282

CHAPTER 8

Table 1.	Tool Selection Criteria Adapted from the Agency for Healthcare Research and Quality (AHRQ) Criteria for the Selection Of Effective Health Care Program Topics	309
Table 2.	Integrated Economic Findings to Inform Identification of Potential Economic Tools	311
Table 3.	Integrated Network Meta-Analysis (NMA) Findings to Inform Identification of Potential NMA Tools	313
Table 4.	Rankings to Inform Selection of Toolkit Components and Priority for Tool Development.	315

Supplementary Material 1

Supplementary Table 1.	Identification of Potential Tools Based on Individual Economic Study Findings	316
Supplementary Table 2.	Identification of Potential Tools Based on Individual NMA Study Findings	326
Supplementary Table 3.	Detailed Descriptions of Potential Tools	332
Supplementary Table 4.	Assessment of Selection Criteria and Rationale for Ranking Toolkit Components	335

APPENDIX D

Table 1.	Parameter Estimates	425
Table 2.	Results for base population	429
Table 3.	Sequential Cost Utility Analysis Results for Alternative Patient Populations	430
Table 4.	Sensitivity Analysis	432

APPENDIX E

Table 1.	Network Meta-Analysis Results: Vertebral fractures. Odds Ratio (95% Credible Interval)	459
Table 2.	Network meta-analysis results: non-vertebral fractures. Odds Ratio (95% Credible Interval)	460
Table 3.	Network meta-analysis results: Hip fractures. Odds Ratio (95% Credible Interval)	461
Table 4.	Network meta-analysis results: Withdrawals due to Adverse Events. Odds Ratio (95% Credible Interval)	462

Supplementary Material 1

Supplementary Table 1.	Patient and Trial Characteristics of Studies Included in NMA.	463
Supplementary Table 2.	Risk of Bias Assessment for Studies Included in Network Meta-Analysis	468
Supplementary Table 3.	Network meta-analysis results: Wrist fractures. Odds Ratio (95% Credible Interval)	472
Supplementary Table 4.	Number needed to treat for benefit or harm for therapies versus placebo for various outcomes	474
Supplementary Table 5A.	Random versus fixed effects comparisons for osteoporosis therapies versus placebo: Vertebral fractures.	475
Supplementary Table 5B.	Random versus fixed effects comparisons for osteoporosis therapies versus placebo: Non-vertebral fractures.	475
Supplementary Table 5C.	Random versus fixed effects comparisons for osteoporosis therapies versus placebo: Hip fractures.	476
Supplementary Table 5D.	Random versus fixed effects comparisons for osteoporosis therapies versus placebo: Wrist fractures.	476
Supplementary Table 5E.	Random versus fixed effects comparisons for osteoporosis therapies versus placebo: Withdrawals due to Adverse events (WDAEs)	476

Supplementary Material 2

Table 1A.	Bisphosphonates versus placebo, ≥ 3 years and < 3 years: Vertebral fractures.	477
Table 1B.	Bisphosphonates versus placebo, ≥ 3 years and < 3 years: Non-vertebral fractures.	477
Table 1C.	Bisphosphonates versus placebo, ≥ 3 years and < 3 years: Hip fractures.	478
Table 1D.	Bisphosphonates versus placebo, ≥ 3 years and < 3 years: Wrist fractures.	478
Table 1E.	Bisphosphonates versus placebo, ≥ 3 years and < 3 years: Adverse event withdrawals.	478
Table 2A.	Bisphosphonates versus placebo, secondary and primary prevention studies: Vertebral fractures.	479
Table 2B.	Bisphosphonates versus placebo, secondary and primary prevention studies: Non-vertebral fractures.	479
Table 2C.	Bisphosphonates versus placebo, secondary and primary prevention studies: Wrist fractures.	480
Table 2D.	Bisphosphonates versus placebo, secondary and primary prevention studies: Withdrawals Due to Adverse Events (WDAEs)	480

LIST OF FIGURES

CHAPTER 2

Figure 1.	Knowledge-to-Action Framework	20

CHAPTER 3

Figure 1.	Most Frequently Reported Types of Barriers and Facilitators to Using Economic Information.	63

Supplemental

Supplementary Figure 1.	Results of the Literature Search.	64

CHAPTER 4

Supplementary Material

Supplementary Figure 1.	Results of Systematic Review Search.	150

CHAPTER 5

Figure 1.	Results of the Literature Search	182
Figure 2A.	Most Frequently Reported Types of Barriers to Using and Communicating NMAs.	183
Figure 2B.	Most Frequently Reported Types of Facilitators to Using and Communicating NMAs.	184

CHAPTER 6

Figure 1.	PRISMA Flow Diagram for Systematic Review of NMA Guidelines.	235

CHAPTER 7

Supplementary Material

Supplemental Figure 1.	Knowledge to Action Framework	273
Supplemental Figure 2.	Key Informant Identification and Recruitment – Canadian Policymakers	276
Supplemental Figure 3.	Key Informant Identification and Recruitment – HTA Organizations	277

CHAPTER 8

Figure 1.	Knowledge-to-Action Framework: Action cycle and knowledge creation activities applied to the development of an economic evaluation and NMA toolkit.	310

APPENDIX D

Figure 1.	Design of Decision Model	434
Figure 2.	Cost Effectiveness Acceptability Curve	435

APPENDIX E

Figure 1.	PRISMA Flow Diagram for Network Meta-Analysis of Osteoporosis Therapies	457
Figure 2.	Evidence network for vertebral fractures.	458

Supplementary Material

Supplementary Figure 1.	Evidence network for non-vertebral fractures.	469
Supplementary Figure 2.	Evidence network for hip fractures.	470
Supplementary Figure 3.	Evidence network for wrist fractures.	471
Supplementary Figure 4.	Evidence network for adverse event withdrawals.	473

ABBREVIATIONS

AGREE	Appraisal of Guidelines for Research and Evaluation
AHRQ	Agency for Healthcare Research and Quality
AMSTAR	A Measurement Tool to Assess Systematic Reviews
ARR	absolute risk reduction
BMD	bone mineral density
BMJ	British Medical Journal
CADTH	Canadian Agency for Drugs and Technologies in Health
CaMos	Canadian Multicentre Osteoporosis Study
CCHS	Canadian Community Health Survey
CDEC	Canadian Drug Expert Committee
CEAC	cost effectiveness acceptability curve
CEAF	cost effectiveness acceptability frontier
CI	confidence interval
CIHI	Canadian Institute for Health Information
CIHR	Canadian Institutes for Health Research
COI	conflict of interest
CrI	credible interval
DB	double blind
DIC	deviance information criterion
EUnetHTA	European network for Health Technology Assessment
FE	fixed effects
GRADE	Grading of Recommendations Assessment, Development and Evaluation
HTA	health technology assessment
HTAi	Health Technology Assessment international
ICER	incremental cost effectiveness ratio
ICES	Institute for Clinical Evaluative Sciences
ICUR	incremental cost utility ratio
IDC	indirect comparisons
INAHTA	International Network of Agencies for Health Technology Assessment
ISPOR	International Society for Pharmacoeconomics and Outcomes Research
KT	knowledge translation
KTA	Knowledge to Action
LTC	long-term care
LY	life year
MA	meta-analysis
MDS	multi-dimensional space
MTC	mixed treatment comparisons
NA	not applicable
NICE	National Institute for Health and Care Excellence
NMA	network meta-analysis
NNB	number needed to benefit
NNH	number needed to harm
NNT	number needed to treat
NR	not reported
OECD	Organisation for Economic Co-operation and Development
OR	odds ratio
PARIHS	Promoting Action on Research Implementation in Health Services
PBAC	Pharmaceutical Benefits Advisory Committee
PCORI	Patient Centred Outcomes Research Institute
PCPA	Pan-Canadian Pricing Alliance
PrI	predictive interval
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
QALY	quality adjusted life year
Q&A	question and answer

RANK-L	RANK ligand
RCT	randomized controlled trials
RD	risk difference
RE	random effects
RNAO	Registered Nurses' Association of Ontario
RoB	risk of bias
RR	relative risk
SD	standard deviation
SR	systematic reviews
SUCRA	surface under the cumulative ranking curve
UK	United Kingdom
WDAE	withdrawal due to adverse event

ACKNOWLEDGEMENTS

I would like to acknowledge everyone who contributed to and supported me in doing this at different points along the way. They are too many to name, their roles were diverse and their impact was immeasurable.

CHAPTER 1.

INTRODUCTION

1.1 RATIONALE

Transparency is a key principle underlying fair and legitimate health technology assessment (HTA) processes and related policy decisions (Daniels & Sabin, 1997; Martin et al., 2002). Transparency requires not only providing sufficient information, but also providing information in an accessible and understandable format for end-users. This is especially relevant when complex methodologies, such as economic evaluations and network meta-analyses (NMA), form the basis of an HTA and are used to support evidence-informed policymaking.

The evaluation of new drugs takes place in the context of the safety and effectiveness of the current standard of care. When head-to-head trials with relevant comparators are not available, the comparative effectiveness of treatment strategies may be estimated through NMAs. NMAs can provide the basis of drug class and therapeutic reviews or they can inform estimates of clinical effects in economic models assessing the cost-effectiveness of treatments (Cooper et al., 2011). Economic evaluations of cost-effectiveness are frequently used in HTAs to provide an assessment of a drug or health technology's "value for money" when compared with other therapeutic options. Policymakers and clinicians may not have the same methodological background as researchers in these technical approaches but must use this information to inform policy and treatment decisions (Thurston et al., 2008).

Approaches that keep analyses and results as understandable as possible without altering their technical meaning can be challenging to achieve. Yet, the consequences of inaccessible HTAs can include misinterpretations of the evidence, challenges applying evidence to actual policy decisions or poor public acceptance of health policy decisions. Overall, increasing the transparency and accessibility of economic evaluations can contribute to increasing confidence in the legitimacy of decision-making process. It may also increase the use and impact of economic evaluations and NMAs to inform policy decisions on resource allocation (Niessen et al., 2012). Finally, more relevant, transparent and accessible information may allow policymakers to better communicate with stakeholders on issues related to the cost-effectiveness and comparative effectiveness of treatments.

As such, there is a need for trans-disciplinary research to bridge communication gaps between researchers and policymakers and enhance knowledge translation (KT) of key components of HTAs, including economic evaluations and NMAs. A systematic approach to the development of tools and approaches targeted to policymakers in these areas would be of value. KT principles

that have been applied in other areas of evidence-based healthcare may offer opportunities to improve the communication and presentation of economic evaluations and NMAs (Straus et al., 2013). While reporting guidelines used by researchers can contribute to enhancing the transparency of analyses (Moher et al., 2011), experience in the realm of clinical trials and evidence-based medicine has shown that they are insufficient for good communication to non-technical audiences and alternate tools and formats of sharing information such as decision-aids and clinical practice guidelines have been developed for patients and clinicians, respectively. Developing tools and alternate presentation formats that enhance the accessibility of economic evaluations and NMAs may also increase their impact and value to policymakers. In the current resource-constrained environment, with a growing number of new drugs and rising drug prices, there is an even greater need for policymakers to have transparent and accessible information on the cost-effectiveness and comparative effectiveness of new drugs (Sullivan et al., 2011; Niezen et al., 2006). Enhanced transparency and accessibility of economic evaluations and NMAs may have multiple downstream consequences including better-informed policy decisions that lead to improvements in health systems and health outcomes; better understanding of evidence and evidence-informed policy decisions among healthcare providers, patients and the public; and increased legitimacy and confidence among stakeholders in evidence-informed health policy decision-making processes.

1.2 OBJECTIVES

The objective of this thesis was to explore the information needs of policymakers when using economic evaluations and NMAs to inform drug funding decisions, as well as explore researchers' approaches to communicating this information. The intent was to use these findings to develop a toolkit that could enhance the relevance, transparency and accessibility of NMAs and economic evaluations for policymakers. More specifically, this thesis will aim to:

- Identify barriers and facilitators to communicating economic evaluations and NMAs. This will be done through systematic reviews of the literature (Chapters 3 and 5) and key informant interviews with Canadian policymakers (Chapter 7).
- Understand researchers approaches to presenting economic evaluations and NMAs. This will be done through a review of guidelines on conducting and/or reporting these analyses (Chapters 4 and 6).
- Identify a sample of international approaches to presenting economic evaluations and NMAs that could be adapted to the Canadian context. This will be done through key informant interviews with senior individuals from a sample of HTA organizations around the world (Chapter 7)

- Propose a toolkit, guided by knowledge translation principles, to enhance the relevance, transparency and accessibility of economic evaluations and NMAs for policymakers. This will be done by building on international approaches and integrating policymakers' information needs and researchers approaches to presenting analyses (Chapter 8)
- Develop components of the toolkit and illustrate their use. This will be done through applications of the tools to an economic evaluation and NMA of osteoporosis therapies (Chapter 8). Specific goals of the toolkit include:
  (1) Assisting policymakers in understanding and identifying key elements of complex methodologies such as economic evaluations and NMAs
  (2) Supporting policymakers in determining how to adapt complex analyses to their local drug funding policy context and
  (3) Enhancing the interface between researchers and policymakers when sharing information on economic evaluations and NMAs.

1.3 CHAPTER OUTLINES

This thesis was developed using a manuscript-based format. Six manuscripts (Chapters 3 to 8) have been prepared, along with an introductory background chapter (Chapter 2) and conclusion chapter integrating findings from this thesis (Chapter 9). Below is an overview of the chapters in this thesis, with a brief description of their content. Of note, the chapters on economic evaluations are presented before chapters on NMAs because the field is more established and the evidence-base is larger, which will provide a better foundation for describing NMA-related issues in this thesis.

Chapter 1. Introduction and Overview.

This is an introductory overview chapter providing an explanation of the purpose of the thesis, how it is organized and a brief description of each chapter.

Chapter 2. Background.

This is an introductory background chapter providing an overview of literature in content areas relevant to this thesis including: HTAs, economic evaluations, NMAs, health policy decision-making and KT.

Chapter 3. Knowledge, Barriers, Facilitators and Approaches to Communicating Economic Evaluations: A Systematic Review.

The objective of this first manuscript is to systematically review the literature on knowledge, barriers, facilitators and approaches to communicating and understanding health economic

information by non-technical audiences. Thirty-three studies were included, most of which were surveys and interviews of policymakers. Strategies that improved communication of health economic information included clarifying technical concepts, focusing on key principles, improving readability and tailoring the level of detail to end-users' needs.

Chapter 4. Guideline Standards for Presenting Economic Evaluations: A Systematic Review.

The objective of this second manuscript is to systematically review guidelines for economic evaluation in health care for information on how to both present economic evaluations and describe concepts commonly required for transparent reporting, which may need further explanation for non-technical audiences. This work will be used to identify what guidance researchers are given on how to communicate and present economic evaluations and consider how their perspective differs from the informational needs of end-users of economic evaluations. Thirty guidelines were included, many of which included a standard reporting format. Despite most guidelines being funded by policymakers, few guidelines provided advice on presenting economic evaluations to non-economists.

Chapter 5. Knowledge, Barriers, Facilitators and Approaches to Communicating Network Meta-Analyses: A Systematic Review.

The objective of this third manuscript is to systematically review the literature on knowledge, barriers, facilitators and approaches to communicating and understanding NMAs. This work will be used to identify the informational needs and optimal presentation of NMAs for non-technical audiences. Sixteen studies were included but most studies focused on researchers' perspectives and only one study developed NMA presentation approaches targeted to different end-users. Current communication approaches may be insufficient for non-technical audiences to understand what NMA results mean and how to use them.

Chapter 6. Guideline Standards for Presenting Network Meta-Analyses: A Systematic Review.

The objective of this fourth manuscript is to systematically review guidelines for conducting and/or reporting network meta-analyses (NMAs) for information on how to present NMAs. This work will be used to identify what guidance researchers are given on how to communicate and present NMAs and consider how their perspective differs from the informational needs of end-users of NMAs. Seven guidelines were included. Current

guidelines focus on how to conduct NMAs but provide limited guidance to researchers on how to best present analyses.

Chapter 7. Canadian and International Perspectives on Knowledge Translation of Economic Evaluations and Network Meta-Analyses: Key Informant Interviews.

The primary objective of this fifth manuscript is to conduct key informant interviews with Canadian policymakers to identify their informational needs related to understanding and using economic evaluations and NMAs to support policy decisions. A secondary objective is to conduct key informant interviews with senior individuals in HTA organizations around the world to identify a sample of existing knowledge tools related to economic evaluations and NMAs. Canadian policymakers (N=22) value information that includes robust content, contextual information so that analyses can be adapted to local and real-word settings and formatting approaches that enhance clarity and key messages. Internationally (N=17 key informants), few KT approaches specific to economic evaluations and NMAs exist, despite an acknowledged need and interest in better communication approaches.

Chapter 8. Enhancing the Accessibility of Economic Evaluations and Network Meta-Analyses for Policymakers: Development of a Toolkit.

The objective of this sixth manuscript is to outline the development of a toolkit that supports the use of economic evaluations and NMAs to inform policy decisions by enhancing their accessibility. This chapter integrates findings from previous systematic reviews and key informant interviews (Chapters 3 to 7) and proposes a toolkit to enhance the relevance, transparency and accessibility of economic evaluations and NMAs for policymakers. Three areas to be addressed by the toolkit were identified from systematic reviews and interviews: communication, education and accessibility. Six tools or approaches were proposed that could be applied to both economic evaluations and NMAs, i.e. 12 toolkit components, and that are targeted to either policymakers or researchers. These included: (1) guidance for researchers; (2) policymaker question and answer documents; (3) analysis templates; (4) learning modules; (5) tool repositories; and (6) lay peer-review processes. The first component of the toolkit, a guide for researchers on writing economic evaluations and a guide for researchers on writing NMAs for policymakers, was developed and illustrated through a case study of an economic evaluation and NMA of osteoporosis therapies for postmenopausal women.

Chapter 9. Conclusions

This summary chapter integrates findings from this thesis with published literature, considering implications for policy and practice, methodological research and theory

development, while considering strengths and weakness of the thesis, including its innovative nature. Next steps in the toolkit development and evaluation are identified, recommendations for a future research program are proposed and conclusions are drawn, noting the additional value that is provided by this thesis research.

Although each of the draft manuscripts (chapters 3 to 8) can be read independently, they should also be considered together in the context of this thesis, whose overall objective is to enhance the relevance, transparency and accessibility of economic evaluations and NMAs.

1.4 REFERENCES

Cooper NJ, Peters J, Lai MC, Juni P, Wandel S, Palmer S et al. How valuable are multiple treatment comparison methods in evidence-based health-care evaluation? Value Health 2011, 14(2): 371-80.

Daniels N, Sabin JE. Limits to health care: fair procedures, democratic deliberation and the legitimacy problem for insurers. Philosophy and Public Affairs 1997;26(4):303–50.

Martin DK, Giacomini M & Singer PA. Fairness, accountability for reasonableness, and the views of priority setting decision-makers. Health Policy 2002, 61: 279–290.

Moher D, Weeks L, Ocampo M, Seely D, Sampson M, Altman DG et al. Describing reporting guidelines for health research: a systematic review. J Clin Epidemiol 2011, 64: 718-742.

Niessen LW, Bridges J, Lau BD, Wilson RF, Sharma R, Walker DG, Frick KD, Bass EB. Assessing the Impact of Economic Evidence on Policymakers in Health Care. A Systematic Review. Methods Research Report (Prepared by the Johns Hopkins University Evidence-based Practice Center under contract No. 290-2007-10061-I). AHRQ Publication No. 12(13)-EHC133-EF. Rockville, MD: Agency for Healthcare Research and Quality. October 2012.

Niezen M, Stolk E, Steenhoek A, Uyl-De Groot C. Inequalities in oncology care: Economic consequences of high cost drugs. Eur J of Cancer, 2006, 42: 2887–2892.

Straus S, Tetroe J & Graham ID (Eds.), Knowledge translation in health care: Moving from evidence to practice, 2nd edition. Oxford: Wiley-Blackwell Publishing, 2013.

Sullivan R, Peppercorn J, Sikora K, Zalcberg J, Meropol N, Amir E. Delivering affordable cancer care in high-income countries. Lancet Oncology, 2011, 12: 933-980.

Thurston SJ, Craig D, Wilson P, Drummond MF. Increasing decision-makers' access to economic evaluations: alternative methods of communicating the information. Int J Technol Assess Health Care. 2008 Spring; 24:151-7.

CHAPTER 2.

BACKGROUND

This chapter provides an overview of the literature and background information on five areas that are integral to the conduct of this thesis: health technology assessment (HTA), economic evaluations, network meta-analyses (NMAs), health-policy decision-making and knowledge translation (KT). A more in-depth review of related literature is also provided in the systematic reviews conducted as part of this thesis (Chapters 3 to 6).

2.1 HEALTH TECHNOLOGY ASSESSMENT

What is Health Technology Assessment?

HTA is defined as "a multidisciplinary field that addresses the health impacts of technology, considering its specific healthcare context as well as available alternatives" and can be adapted to address and inform the needs of the health system. Other factors addressed by HTA may include economic, organizational, social, and ethical impacts (HTAi 2014; Cookson et al 2000).

The effects of new technologies may be assessed through evidence syntheses such as systematic reviews, meta-analyses and, more recently, NMAs. In the absence of head-to-head trials of relevant comparators, the comparative effectiveness of treatment strategies can be estimated by conducting an NMA (Salanti 2011). Economic evaluations are also often conducted as part of an HTA to inform the cost-effectiveness or "value for money" of drugs and health technologies when compared with existing treatment options (Drummond et al., 2005).

Based on HTAs and other policy-relevant factors, policymakers decide how to allocate resources and which drugs and technologies to fund (Tierney et al., 2008). HTA has supported the practice of evidence-based medicine and evidence-informed policy decision-making for approximately 20 years (Drummond et al., 2013). However, in the current era of high drug costs, personalized medicine and fiscal constraints, HTA takes on growing importance in informing resource allocation decisions. (Niezen et al., 2006; Sullivan et al., 2013). While advancements in medical sciences may increase, the collective resources that society can spend on health technologies may not correspondingly increase and decisions must be made on how finite resources are allocated and which technologies are funded.

Canadian and International HTA Processes

Many countries, regions and health systems (e.g., hospitals, insurance providers) have developed their own HTA processes (Morgan et al., 2006, Mathes et al., 2013; Kleinjen et al., 2012; Oortwin et al., 2013). While these processes share similarities, HTA organizations around the world vary in their size, structure, mandate and how they are integrated into the healthcare system. For example, while some organizations make actual policy decisions, others provide recommendations that policymakers may choose to follow or not. Other HTA organizations may provide only HTA reports and conclusions on the evidence-base, but make no specific recommendations on how the evidence should be applied to policy decisions. The evidence synthesis approach may also vary, depending, in part, on available expertise and resources. For some organizations, cost-effectiveness may be outside their mandate, while others may critique cost-effectiveness analyses conducted by other groups. In other cases, pharmaceutical companies with a vested interest in the HTA may provide the evidence base for an HTA (e.g. conducting an economic evaluation) and there may be concerns related to conflict of interest and bias (Bell et al., 2006; Garatini et al., 2010). Some organizations don't conduct NMAs but focus on more traditional evidence synthesis methods such as systematic reviews and meta-analyses. Development of economic models and multiple technology HTAs, in which NMAs may be conducted, can be time-consuming and resource intensive and may be limited by economic and statistical pools of expertise to draw upon (Thabane et al., 2011, Madden et al., 2009). In these cases, HTA organizations and/or policymakers may rely on already published analyses in the scientific literature or reports produced by other collaborations, e.g. Cochrane Collaboration, or by larger HTA organizations (Lavis et al., 2006; Chase et al., 2009). HTA organizations around the world can also benefit from shared knowledge and experiences through collaborations such as INAHTA and HTAi (INAHTA 2014, HTAi 2014).

In Canada, a number of regional or local HTA review processes exist, which are usually supported by academic units or research institutions. There is also a common national review process for drugs that is supported by the Canadian Agency for Drugs and Technologies in Health (CADTH), providing recommendations to nine Canadian provinces (excluding Quebec), the Canadian territories and federal drug plans (Tierney et al., 2008). A number of countries have HTA organizations and systems in which the drug-funding context is similar to Canada, e.g. the National Institute for Health and Care Excellence (NICE) in England and the Pharmaceutical Benefits Advisory Committee (PBAC) in Australia (Clement et al., 2009). These contextual factors include providing advice to policymakers, having a publicly funded health-care system and conducting or using cost-effectiveness analyses and/or NMAs to inform decisions.

Transparency and Accessibility of Health Technology Assessments

The use of public funds to cover drug costs requires that resource allocation decisions be made through a fair and legitimate process. One of the cornerstones of fair and legitimate policy decisions and underlying HTA processes is the provision of publicly accessible information on the rationale and reasons for a decision (Daniels & Sabin, 1997). This can have a number of positive consequences including improved consistency of decision-making and increased confidence in the legitimacy of decision-making processes. Clinicians, patients and the public are key stakeholders in HTA and drug-funding processes with a legitimate right to understand why limits may be placed on a drug or technology that they perceive could provide a benefit. (Daniels & Sabin, 1997) Therefore, ensuring that stakeholders are able to understand the overarching reasons for a policy decision is essential. While the amount of detailed evidence underlying a decision that a stakeholder requires in order to fully understand the rationale for a policy decision may vary, ensuring that evidence is accessible is important. However, accessibility and transparency involve not only making information available but also ensuring that information is provided in an understandable format. In addition, understanding and aligning the needs and perspectives of policymakers, clinicians and patients, i.e. knowledge users, with the needs and perspectives of researchers, i.e. knowledge producers, when enhancing the accessibility and transparency of HTAs is an important consideration (CIHR 2012).

HTA organizations have attempted to address transparency in different ways. For example, health technology reports and related recommendations may be widely disseminated either through organizational websites, companion publications in academic journals or other mechanisms (Scotland et al., 2011). Some organizations and collaborations have developed simplified evidence summaries for stakeholders, e.g., the Cochrane Collaboration, the Agency for Healthcare Research and Quality (AHRQ). However, these have primarily focused on traditional forms of evidence synthesis such as systematic reviews and meta-analyses. Attempts to summarize more complex scientific techniques and approaches such as economic evaluations and NMAs in plain language can be more challenging. Achieving the appropriate balance of making information understandable while maintaining the completeness and technical accuracy of economic evaluations and NMAs can be difficult, as important nuances may be lost by oversimplifying information.

Clarifying the information needs of Canadian policymakers, i.e. knowledge users, with respect to the communication of complex scientific techniques such as economic evaluations and NMAs that are used in HTAs, can provide valuable insights on how to enhance the transparency and accessibility of HTAs. Furthermore, learning from the experiences of HTA organizations around

the world on how they communicate and share information on economic evaluations and NMAs may inform the development of best practices that can be built upon in the Canadian setting.

2.2 ECONOMIC EVALUATIONS

Use of Economic Evaluations in Health Technology Assessments

Economic evaluations of cost-effectiveness are frequently used in HTAs to provide an assessment of a drug or health technology's "value for money" when compared with other therapeutic options (Drummond et al., 2005). As such, economic evaluations can aid policymakers in determining the optimal allocation of resources and how to maximize the health benefits for the population. As a result of resource constraints, choices must be made between therapeutic options and, as such, other benefits are lost because resources must be diverted from other opportunities, that is, there are 'opportunity costs' to resource allocation decisions (Mitton et al., 2002). Depending on how budgets are structured, these opportunities can be other drugs or health technologies, other improvements in the health care system or other societal benefits such as education and social welfare (Rawlins et al., 2011). Therefore, cost-effectiveness analysis cannot be considered in isolation by policymakers as resource allocation decisions must necessarily consider existing budgets. In a resource-scarce environment, in which, all drugs and technologies cannot be funded, economic evaluation is a valuable tool. It also has the advantage of providing a synthesis of much of the information relevant to decision-making, thereby, making any gaps in the evidence-base more apparent to policymakers. Although economic evaluations have been in use for the last 20 years, their impact on health policy is still uncertain (Drummond et al., 2013, Niesson et al., 2012). In this light, reflecting on how to increase the relevance, transparency and accessibility of economic evaluations to better demonstrate their value in making evidence-informed health policy decisions, merits consideration. Enhancing transparency such that there is a better understanding of the "value for money" of evidence-informed policy decisions, the related opportunity costs and how the use of limited resources attempts to maximize the health of the entire population can lead to greater confidence in the fairness and legitimacy of policy decisions, even when individuals may disagree with specific decisions (Daniels & Sabin 1997).

Economic Concepts Relevant to Health Policy and Resource Allocation Decisions

Different types of economic evaluations are frequently used to support policy decisions including cost effectiveness and cost utility analyses. In cost-effectiveness analyses, the units of effect are typically expressed in life-years, (i.e. survival gained or lost due to the intervention). In cost-utility

analyses, the units of effect are expressed as quality-adjusted life years (i.e. a measure of the quality of life that is gained or lost due to the intervention). Quality adjusted life-years (QALYs) are calculated by determining the life-years (LYs) that are gained and applying a utility value to them. In general, utility values are a measure of the preference that individuals have for an outcome, e.g. different states of health; the more preferable an outcome, the more utility is associated with it. Similarly, disutilities are a measure of the aversion that individuals have to an outcome, e.g. treatment side effects. Utilities can be obtained from population surveys or derived from quality of life data collected in clinical trials (Drummond et al., 2005).

Incremental QALYs (or incremental LYs) are used, along with incremental costs to calculate an incremental cost-effectiveness ratio (ICER), which is a measure of the "value" (i.e. clinical benefit expressed in QALYs or LYs) "for money" (i.e. the cost expressed in monetary units, e.g. Canadian dollars) of a treatment compared with alternatives (Drummond et al., 2005). In some regions when cost-effectiveness analyses are conducted, a standard level is set of how much people in that society are willing to pay for gains in survival or quality of life. This is referred to as a willingness-to-pay threshold. For example, in the UK, the general willingness-to-pay threshold for new technologies is between £20,000 and £30,000 per QALY and in Canada has been reported to be approximately $50,000 per QALY (Schlander 2007; Winquist et al., 2012). If a treatment has an incremental cost-effectiveness ratio that falls below this threshold, it is considered cost-effective and if the ICER is above this threshold, it is not considered cost-effective. However, there is considerable debate over the appropriateness of using a cost-effectiveness threshold to make final funding decisions as factors other than cost-effectiveness influence policy decisions. For example, in other regions, where no standard willingness-to-pay threshold is set, an ICER threshold is considered "not static" and is subject to "changes in the budget, the interventions funded, and the productivity of health care" (Cleemput et al., 2011)

When considering how to apply the results of economic evaluations to policy decisions, some factors related to economic modeling may be considered, including the quality of data sources, modelling approaches and exploration of uncertainty. Good quality data inputs can increase confidence in the results derived from an economic model. However, local data can sometimes be difficult to identify or collect. Good modeling practices dictate that models should be as simple as possible, therefore, trade-offs must be made between the complexity of the model and the precision that it offers. (Roberts et al., 2011) As a result, assumptions are sometimes made and verifying the appropriateness of these assumptions is critical. One of the advantages of cost-effectiveness analyses is that the costs and effects of therapies can be modeled and projected over time. This allows for the extrapolation of short-term trial data into long-term effects over a population's lifetime. However, the use of modeling techniques means that there is always some

inherent uncertainty in results produced by the economic analysis. Key sources of uncertainty in the analysis include structural uncertainty, parameter uncertainty and variability. These are often addressed through sensitivity and scenario analyses with the two main approaches being deterministic and probabilistic sensitivity analyses (Drummond et al., 2005).

Communicating and Presenting Economic Evaluations

Policymakers have sometimes expressed uncertainty regarding how best to interpret and apply the results of cost-effectiveness analyses to decisions (Williams et al., 2008). Furthermore, policy decisions can be controversial when limits are placed on patients' and clinicians' access to treatments and greater transparency and accessibility of economic evaluations in these situations is of value. While reporting guidelines exist for economic evaluations (CADTH 2006; Philips et al., 2004) and can contribute to clarity, they may be insufficient for optimal communication to lay audiences (Galani et al., 2008; Thurston et al., 2008). However, achieving the appropriate balance of providing basic explanations of economic concepts while maintaining the technical accuracy of the economic evaluation can be challenging, as important nuances may be lost in the simplification of information. Reviewing the literature on approaches for communicating and presenting economic evaluations (Williams et al., 2008, Erntoff et al., 2011, Niessen et al., 2012) and applying approaches from the fields of evidence-based medicine and knowledge translation may offer opportunities to enhance their transparency and accessibility. More transparent and accessible economic evaluations will provide greater confidence in the legitimacy of health policy decision-making processes and could increase awareness of the impact of economic evaluations on evidence-informed health policies.

2.3 NETWORK META-ANALYSIS

Use of Network Meta-Analyses in Health Technology Assessment

The ability of HTAs to inform policy decisions is dependent on the availability of evidence addressing the policy-relevant questions. Direct comparisons of a new drug with the current standard of care are ideally obtained through head-to-head trials of the treatments of interest or meta-analyses of these trials. However, many clinical trials are designed based on national regulatory agency requirements for market authorization and head-to-head trials are rarely required. When head-to-head trials are not available, the comparison of interest can be estimated indirectly through an NMA. NMAs are sometimes a key component of HTAs provided to

policymakers, as they can provide the evidence-base for a drug class or therapeutic review and for cost-effectiveness evaluations (CADTH 2009).

In the absence of formal statistical approaches like NMA, individuals may conduct informal 'back-of-the-envelope' assessments, indirectly comparing trial characteristics and results, without realizing that this is a form of indirect comparison. (Salanti et al., 2012) Applying principles and a methodology to guide how these indirect assessments are done, formalizes the process and provides a more robust approach to how indirect evidence can inform conclusions on comparative effectiveness. However, the validity of NMA results depends on the appropriate statistical conduct and clear presentation of the NMA results, such that policymakers and clinicians understand the limitations of NMAs and how to apply results to policy and treatment decisions.

Relevant NMA Background and Concepts

Network meta-analyses are an evolution of meta-analyses and share many of the same principles (Cooper et al 2011; Salanti et al 2012). However, a meta-analysis focuses on comparing two interventions at a time and there is no requirement for common comparators across trials or a network of treatments. This differs from network meta-analyses where a network of treatments connected by similar comparators exists and allows for the effect of multiple interventions to be compared simultaneously. The evidence included in an NMA can be represented by a network diagram and assessment of the network's shape can provide information on key characteristics of the network such as the relative presence of placebo-controlled and head-to-head trials (Ionnidis 2011, Nikolakopoulou et al., 2014).

Methodological standards for NMAs have been evolving over the past few years (Salanti et al., 2012). As a result, inadequate methodological approaches have sometimes been applied when conducting NMAs (Bafeta et al., 2013; Bafeta et al., 2014; Donegan et al., 2010; Song et al., 2009). NMAs are based on a number of assumptions, which must be explored and verified in order to ensure their methodological robustness (Salanti et al., 2012). These include (1) that the studies are similar in all ways other than the interventions being compared and that the indirect estimate of a comparison is a valid estimate of the unobserved head-to-head direct comparison estimate, i.e., the transitivity assumption; and (2) that when both direct and indirect estimates are available for a comparison, they are in agreement, i.e. the consistency assumption. These principles are similar to the assumption required for standard meta-analyses that trials are homogenous enough that they can be validly combined in an analysis. Methodological limitations in NMAs are related to the underlying assumptions required for conducting a network meta-analysis not being met or assessed. Therefore, the appropriateness of underlying assumptions

should be explored and clearly documented in an NMA, in order for a user to better understand if the results of a NMA are valid. For example, methodological problems identified in an analysis of 88 indirect comparisons by Song et al. (2009) included an unclear understanding of underlying assumptions; an inappropriate search for or selection of included trials; inappropriate or flawed methods; trial similarity not being objectively assessed; or an inappropriate combination of direct and indirect evidence.

NMAs are similar to traditional meta-analyses in that the final result is a treatment effect estimate (e.g., odds ratio, relative risk). However, in a NMA, an effect is available for each pair of comparators in the network versus one treatment effect that would be available for the single treatment comparison in a standard meta-analysis. When a NMA is conducted using a Bayesian framework, the rank of each treatment can be estimated, e.g. probability of being the best, based on the reported effect size across multiple simulations. Looking at the treatment with the highest probability of being ranked best may not always be most useful, though. For example, a treatment may have a low probability of being best but a high probability of being second or third best, which may be acceptable in certain decision-making contexts (Dias et al., 2012). The uncertainty in a treatment effect estimate can also affect it's ranking. For example, a high level of uncertainty in the treatment effect can lead to a high probability of being best, which is misleading. In these cases, cumulative rankings may provide additional insights by indicating the probability a treatment is best, second-best, third-best, and so on (Salanti et al., 2011). The probability of a treatment being best or second-best or third-best must also be interpreted carefully since in some cases the difference between treatments may be small and not clinically meaningful (Li et al., 2011).

Communicating and Presenting NMAs

One frequently cited concern with NMA is that the complex statistical methods do not permit the end user to understand how the results were obtained or if they are valid (Pocock 2007, Abdelhamid et al., 2012). Guidance documents outlining general requirements for the conduct and interpretation of NMAs exist for different audiences and different purposes (CADTH 2009; Dias et al 2012 [TSD1, TSD2, TSD3, TSD4]; Hoaglin et al., 2011; Jansen et al., 2011; Mills et al., 2012) and reporting guidelines that will enhance the transparency of NMAs are currently under development. (Hutton et al., 2014) While good reporting practices are important for researchers to follow and can contribute to clarity, they may be insufficient for optimal communication to lay audiences such as health policymakers. Optimal approaches to communicating NMAs are currently being explored by researchers (Tan et al., 2014) but general principles and experience related to how to best share information with knowledge users can also be gleaned from research in evidence-based medicine and knowledge translation. Having knowledge users such as

policymakers and clinicians be able to understand and accurately interpret the results of NMAs in the context of real-world practice and policy decisions can increase the impact of NMAs. Similar to economic evaluations, creating and presenting NMAs in a transparent and accessible format may also lead to greater confidence in the legitimacy of health policy decisions informed by NMAs.

2.4 HEALTH POLICY DECISION-MAKING AND PUBLIC DRUG FUNDING

Canadian Public Drug Funding Policy Context

The Canadian policy environment for public drug funding is a complex system made up of federal, provincial, territorial (FPT) and regional decision-makers due to the fact that healthcare in Canada is a shared responsibility across FPT governments (Daw & Morgan, 2012; National Pharmaceutical Strategy, 2006). While provinces and territories provide public drug funding for their residents, the federal government is responsible for drug funding in populations under its jurisdiction (e.g. First Nations, veterans, Canadian Forces, federal inmates). Each of these Canadian jurisdictions varies in size, available resources, populations served, political mandates, and other key factors that can influence health policies. As a result, the local context in which policy decisions are made is unique to each jurisdiction and how evidence is applied to policy decisions also varies across jurisdictions. Within each jurisdiction, different types and levels of policymakers exist, and not all make final decisions. In this thesis, policymakers were defined broadly as those individuals with knowledge of how evidence is used to inform or influence drug policy and funding decisions. As it applies to public drug funding, jurisdictions also differ in the populations that are covered and the types of drugs that are funded. Finally, drug pricing may differ between jurisdictions and the nature of confidential contracts with pharmaceutical manufacturers means that these effective prices are often unknown to others. Recent collaborative initiatives among provinces, such as the Pan-Canadian Pricing Alliance, may lead to less provincial disparity in drug pricing, however, prices will likely still remain confidential (Morgan et al., 2013). Taken together, all of these factors create challenges when trying to develop economic evaluations and NMAs that are applicable and relevant to the local context of all Canadian provinces, territories and federal drug programmes.

Health policy and resource allocation decisions try to maximize the health gains of the population. Therefore, although clinical research is conducted in defined patient populations and clinicians treat individual patients, policy decisions must be considered at a population level. Health policy decisions can be complex, multi-factorial decisions that consider, not only comparative and cost-effectiveness, but also ethical factors, competing resources and unmet needs in the population

(Lehoux et al., 2007). At times, policy decisions may be controversial and ensuring the legitimacy and transparency of these decisions and the evidence upon which they are based is important (Daniels & Sabin, 1997).

Policymakers' Use of Evidence

Policymakers may experience barriers to the practical use and application of evidence, including economic evaluations and NMAs, to inform health policy decisions. Systematic reviews have been conducted to identify factors that improve policymakers' use of evidence (Chambers et al., 2011; Innvaer et al., 2002; Lavis et al., 2005; Murthy et al., 2012; Perrier et al., 2011). In the majority of these reviews, findings were limited by the paucity of evidence and heterogeneity of study designs. One finding, however, was that providing a single clear message could lead to greater use of systematic review evidence, when there is awareness that change in policy or practice is required (Murthy et al., 2012). Systematic reviews that focused on interview studies with health policymakers identified a number of facilitators to research uptake. For example, Innvaer et al., (2002), found that research that included a summary with clear recommendations enhanced evidence uptake. Similarly, Lavis et al., (2005) found that providing details on the applicability of research findings to the local context and providing information in a format that could be scanned quickly was valued by policymakers.

In a single study in which public health decision-makers in Ontario were interviewed about preferences for receiving evidence, it was found that having a clear statement relating to both short and long-term implications for practice was very helpful (Dobbins et al., 2007). In this study, decision-makers emphasized that they were looking not only for facts but also for guidance on implementing research findings in practice and how those findings may apply to their specific situation, e.g. subpopulations and settings. Policymakers in this study also indicated that they only want limited information on methodology. Abbreviated summaries of high-quality research containing these elements would allow policymakers to determine the value of the research and then they could access more details as needed in full technical documents.

One of the key principles of knowledge translation is that interventions should be tailored to the target audiences (Wensing et al, 2013). Policymakers may not have the same methodological background as researchers, yet, they must still use the results of complex methodologies such as economic evaluations and NMAs to inform policy decisions. Also, policy makers have indicated they have limited time to review information (Innvaer et al 2002; Lavis et al 2005) and may be required to communicate and defend policy decisions to the public. Even among policymakers, preferences for different formats have been expressed, with one study reporting that legislative

staff prefer briefs presenting information as anecdotes while legislators preferred data-focused policy briefs (Brownson et al., 2011). Finally, while both policy makers and other knowledge users may express preferences for the information they receive or how they receive it, these preferences must be critically examined as preferences do not always lead to improved and accurate understanding (Ancker et al., 2006; Treveana et al., 2012).

Policymakers' use of evidence is also influenced by key stakeholders who are impacted by policy decisions such as clinicians and patients. Translating evidence that is understandable to patients is taking on greater importance given the growing role patients are taking in health-related research and treatment decisions, especially when access to treatments may be impacted by health policies. For example, the Patient-Centered Outcomes Research Institute (PCORI) requires that results of funded research be communicated in a way patients understand, including information about the limitations of the research and need for further research (Leonard 2010). A review of best practices for communicating numerical concepts to patients emphasized simplification, clear formatting, the omission of distracting information, appropriate framing and the use of visuals (Apter et al 2008). One of the most important things patients want to know from evidence is if the results are applicable to their situation and, if not, why not (Leonard 2010). This may require understanding subtle aspects of the economic evaluations and NMAs, which can be challenging to accurately but simply communicate to patients with a range of health literacy levels (Hildon et al., 2012). A systematic review on communicating with patients about evidence (Trevena et al., 2006) reported that communication tools that were structured, tailored or interactive were more likely to increase patients' understanding, regardless of the format that was used. Further consideration of how approaches to communicating economic evaluations and NMAs could be tailored to patients may be warranted in future (Wensing et al., 2013).

Understanding Canadian policymakers information needs and experiences with using economic evaluations and NMAs, as well as identifying approaches HTA organizations and researchers have taken to tailor and target economic evaluations, and more recently NMAs, to policymakers is important. Enhancing the transparency and accessibility of evidence informing health policy decisions may increase confidence in the legitimacy of evidence-informed health policy decisions. It may also lead to greater understanding and engagement of stakeholders in HTA processes, which is a growing arena in this era of greater patient and public involvement in HTA processes.

2.5 KNOWLEDGE TRANSLATION

Knowledge Translation Principles

The overall goal in the field of knowledge translation (KT) is the successful implementation of evidence into policy or practice. This field has been described using many different terms, such as knowledge exchange, knowledge transfer, knowledge dissemination, knowledge mobilization, implementation science or research utilization (Graham et al., 2006). The Canadian Institutes of Health Research (CIHR), defines KT as "a dynamic and iterative process that includes synthesis, dissemination, exchange and ethically-sound application of knowledge to improve health, provide more effective health services and products and strengthen the health care system." (CIHR 2012) This is a comprehensive definition that incorporates important concepts and nuances that may not be suggested by alternate terminology (CIHR 2014). For example, knowledge is a broad term that encompasses not only research but also other forms of evidence such as patient values and stakeholder feedback. "Translation" recognizes that producing knowledge is not always sufficient to create changes in behavior, practice or policy and that action must be taken to transform or translate this knowledge in some way for the end-user. "Synthesis" recognizes that research from individual studies must be brought together and integrated into the broader context of key stakeholders and practice and policy environments. "Dissemination" involves identifying appropriate target audiences and tailoring knowledge for their use. This includes both traditional communication approaches as well as more active approaches that engage key stakeholders. For example, "exchange" implies a two-way interactive process between knowledge producers and knowledge users, whereas other terms such as "diffusion" and "transfer" imply one-way processes. This definition also recognizes that implementing evidence can occur in rapidly and continuously changing environments and so must employ a "dynamic and iterative" approach to that implementation. The importance of ethical principles, social values and legal regulations is also integral to how knowledge is translated. Finally, the definition encompasses the ultimate goal of knowledge translation, which is to effectively use knowledge to change behaviors, practices and policies in ways that improve health outcomes and health systems.

The Knowledge-to-Action Framework

Many different frameworks have been developed to guide the implementation of evidence into practice (Rycroft-Malone & Bucknall, 2010). In Canada, the Knowledge-to-Action (KTA) framework has been developed and is often followed by Canadian researchers. This is a comprehensive framework that is integrated into decision-making at prominent evidence-based organizations such as the Canadian Institutes of Health Research (CIHR) and the Registered Nurses' Association of Ontario (RNAO) (CIHR 2012; RNAO 2012). It recognizes that creating new knowledge, on its own, may not lead to improved health outcomes and more effective health

systems, and may require accompanying actions or activities to facilitate change. Therefore, the KTA framework (see Figure 1) is presented in two parts: knowledge creation and the action cycle (Graham et al., 2006).

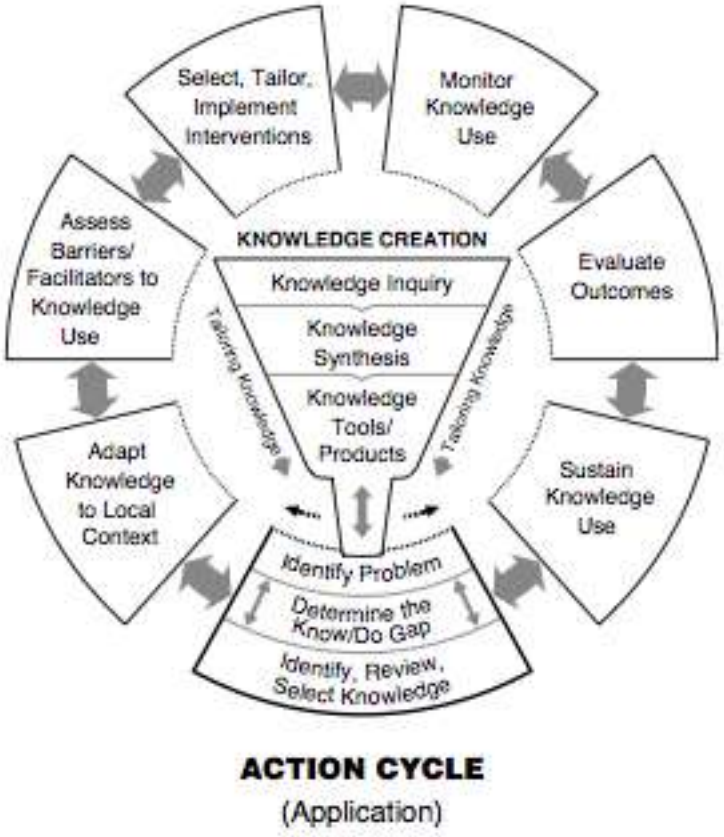


Figure 1. Knowledge-to-Action (KTA) Framework
From Straus et al., 2013 and used with permission from John Wiley & Sons.

Knowledge creation is represented as a funnel in the KTA framework with each layer in the funnel becoming more refined than the previous layer. The funnel implies that knowledge created by researchers, or other knowledge producers, must be distilled into forms that are more useful to knowledge users and allow the user to more easily apply research findings to improve health outcomes or health systems. This includes synthesizing research findings as well as tailoring the knowledge to specific types of users (e.g. policymakers, healthcare providers, patients) and adapting knowledge to the environment in which it will be implemented (e.g. local policy and practice settings). The purpose of producing tools is “to present knowledge in clear, concise and

user-friendly formats and ideally to provide explicit recommendations with the intent of influencing what stakeholders do and to meet the stakeholders' knowledge or informational needs, thereby facilitating the uptake and application of knowledge" (Graham et al., 2006). Examples of tools that have been developed to enhance the use of health-related evidence include clinical practice guidelines, patient decision aids and plain-language summaries (Brouwers et al., 2013). If more than one tool is developed to address the same problem, these tools can be assembled into a toolkit. Economic evaluations and NMAs are forms of knowledge creation. Determining how these analyses could be further distilled and tailored to target audiences such as policymakers merits consideration. Therefore, development of a toolkit to enhance the relevance, transparency and accessibility of economic evaluations and NMAs is the focus of this thesis.

The action cycle represents those activities that are needed in order to apply knowledge to improving health systems and health outcomes. The steps in the action cycle are derived from an analysis of planned-action models. Planned-action models are those models that are intended to bring about change (Graham et al., 2006). These steps represent key principles underlying knowledge translation and are presented as a cycle because this is a dynamic process, whereby activities may influence each other, may occur iteratively or may occur in varying sequences. In the KTA framework, the first step is identifying the problem or the knowledge-to-action gap, i.e., what knowledge can be better implemented in policy or practice to improve health outcomes and/or health systems. Once the problem has been identified, barriers and facilitators to knowledge use are assessed. These findings can be used to select and develop tools that are both tailored to specific target audiences, e.g. policymakers, healthcare providers or patients, and that allow knowledge to be adapted to local policy and practice settings.

Later steps in the KTA cycle include implementation, evaluation, monitoring and sustainability. This thesis, which develops a toolkit for NMAs and economic evaluations, does not focus on these latter steps of implementation, evaluation, monitoring and sustainability. However, they will be integral to the ultimate success of the toolkit.

Approaches to Communicating Health Risks and Evidence

In considering the CIHR definition of knowledge translation and its underlying principles, optimizing communication, i.e., conveying information to others, is only one aspect to consider when trying to change behaviours, practice or policy. However, optimal communication is integral to enhancing the relevance, transparency and accessibility of economic evaluations and NMAs and end-users' understanding of these complex analyses and, therefore, is the focus of this thesis. Optimal communication should consider not only what information is provided to target audiences, but also the format of how information is presented and the medium of how

information is shared with the end-user. Applying principles from the KTA framework such as identifying barriers and facilitators to communication, tailoring presentation formats to target audiences and ensuring content can be adapted to local settings will assist in enhancing the relevance, transparency and accessibility of economic evaluations and NMAs for end-users. While these principles have not been extensively applied to presenting economic evaluations and NMAs, to-date,, a body of evidence exists in other fields that may be applicable when optimizing the communication and presentation of economic evaluations and NMAs for policymakers. This includes a large body of literature on best practices for communicating health risks and evidence to knowledge users, of which a few reviews are summarized below.

One review reporting on different visual formats of displaying data found that, in many cases, tables were preferred over bar graphs and led to more accurate comprehension of concepts (Hildon et al., 2012). There are a number of studies that have indicated that how results are framed, i.e., framing effects, can influence decisions (Apter et al., 2008). A systematic review including 35 studies on framing effects (Akl et al., 2011) found that framing results positively versus negatively (e.g., remaining free of a sore arm versus getting a sore arm) or in terms of gain versus loss (e.g. chance of survival versus chance of dying) had a significant impact in individual studies but not overall. This may suggest that framing has an effect only in select circumstances. Jargon is frequently cited as a concern when public or patients try to understand complex technical information and attempts to inject plain language into methods and results to make them accessible for non-technical audiences have been explored (Elwyn et al., 2006). Finally, authors may overload the user with too much detail when simplicity is preferred (Hildon et al., 2012).

Well-designed visual displays are appealing as they can decrease the amount of mental computation required to understand a concept. However, a number of narrative and systematic reviews have reported that graphs that provide simplicity or provide audiences with familiar formats, do not necessarily result in accurate comprehension (Ancker et al., 2006; Trevena et al., 2012). A certain level of expertise and instruction may be required to interpret the graphs (Ancker et al., 2006) or individuals may interpret graphs in ways that weren't originally intended. For example, individuals may be more focused on the pattern of data than precise values. Therefore, graphs may be best used to display general information and bottom-line messages; more detailed and precise information could be presented in other ways (Trevena et al., 2012). In many cases, text to accompany graphics may be necessary, either as a written description or in a figure legend (Visschers et al., 2009). Graphs have been shown to improve quantitative understanding if the display helps individuals understand the relationship between the numerator and the denominator (Ancker et al., 2006; Visschers et al., 2009). Also graphical displays that capitalize

on basic graphical perceptions such as height of bar graphs, angles in a pie chart or colour gradients have been shown to improve quantitative understanding. Finally, the numerical formats employed, such as using a common denominator have been shown to facilitate quantitative comparisons (Ancker et al., 2006; McCaffery et al., 2012).

Evidence suggests that individuals interpret probabilities better when they are presented at the discrete level rather than as proportions or percentages (Ancker et al., 2006; Burkell 2004). Providing discrete arrays of 1000 faces (Cates calculator) may better demonstrate the relationship between numerator and denominator and thereby lead to improved understanding (Ancker et al., 2006). Images other than faces have been used (e.g., asterisks, dots, circles), however patients have reported finding faces or human figures more meaningful (Ancker et al., 2006). Individuals also have challenges understanding very small risks as probabilities, therefore, in these cases, large icon arrays such as the 1000 faces may be helpful (Burkell 2004; Visschers et al., 2009).

Systematic reviews have reported that, among other factors, adding icon arrays to numerical information improved health literacy (McCaffery et al 2012; Hildon et al 2012). However, icon arrays are just one of a number of graphical formats that may reduce biases such as denominator neglect, framing effects and over-emphasis of anecdotes (Apter et al., 2008). Visual representations such as the 1000 faces array can be helpful for individuals with high graphical literacy or individuals who have challenges understanding and applying numbers (Trevena et al., 2012; Burkell 2004). Many researchers have struggled with how best to present the concept of uncertainty and its impact on decision-making (Hildon et al., 2012). A systematic review on how best to present probabilities indicates that additional research is required to determine optimal formats for conveying uncertainty that is a result of either randomness or a result of ambiguity (Trevena et al., 2012). There are reports that when risk probabilities are expressed only in words, they are not well-understood by individuals (Apter et al., 2008; Visschers et al., 2009). Part of the rationale for this poor understanding is that pictures require less computational effort, so providing pictures rather than words benefits individuals with low health literacy (Hildon et al., 2012). However, in many cases, it is important that text also accompanies the pictures as graphics may not always be accurately understood, especially the first time they are examined (Visschers et al., 2009). Although many individuals express a preference for pictures, these preferences may not always equate with improved comprehension. However, in support of these preferences, systematic reviews have found that, among other factors, presenting numerical information in tables or pictographs rather than text improves health literacy (McCaffery et al., 2012; Apter et al., 2008). A review of strategies for communicating numerical concepts to patients also reported

that pictographs can improve attention and recall when they are linked to text as well as decrease the influence of framing effects and anecdotes (Apter et al., 2008).

Knowledge Translation Applications to Economic Evaluations and NMAs

While economic evaluations and NMAs inform HTAs in many processes and organizations around the world, it is unknown if additional approaches to knowledge translation are also taken to help stakeholders better understand and use this type of evidence. Many important stakeholders generate, use or are impacted by NMAs and economic evaluations including, researchers, policymakers, healthcare providers, patients and the public. Each of these stakeholders have unique characteristics, points-of-view and knowledge bases that should be considered when developing tools to enhance the transparency and accessibility of NMAs and economic evaluations. While recognizing the importance of all these stakeholder groups, this thesis focuses on policymakers to maintain a reasonable scope of the thesis. Policymakers were selected because: (1) considerable literature exists on policymakers' use of other types of evidence syntheses, which may be applicable to the use of economic evaluations and NMAs; (2) policymakers are often requestors of economic evaluations and NMAs, therefore, it is important to ensure analyses are sufficiently transparent and accessible enough to meet their needs before considering other stakeholders; and (3) policymakers are directly linked to each of the other stakeholder groups, therefore, information on other stakeholder groups may be obtained indirectly as policymakers consider their interactions with researchers, healthcare providers, patients and the public. The long-term intent is to build on this research with policymakers and ultimately expand the applicability of the toolkit to other stakeholder groups.

Exploring how KT principles can be applied to communicating economic evaluations and NMAs to policymakers has the end-goal of developing a toolkit to enhance the transparency and accessibility of these analyses for policymakers. Creating the toolkit following the KTA framework provides a systematic and established approach to formally develop best practices for sharing information on economic evaluations and NMAs with policymakers. Although some approaches may exist and have been informally developed, the evidence-base underlying these approaches remains undefined. Using a mixed-methods approach to develop and define issues in these areas will draw together academic literature, guideline standards for researchers, Canadian policymaker perspectives and international experiences to inform best practices to increase the relevance, transparency and accessibility of economic evaluations and NMAs.

2.6 REFERENCES

Abdelhamid A, Loke Y, Parekh-Bhurke S, Chen Y, Sutton A, Eastwood A, Holland R, Song F. Use of indirect comparison methods in systematic reviews: a survey of Cochrane review authors. Research Synthesis Methods 2012, 3: 71-79.

Akl EA, Oxman AD, Herrin J, Vist GE, Terrenato I, Sperati F, Costiniuk C, Blank D, Schünemann H. Framing of health information messages. Cochrane Database of Systematic Reviews 2011, Issue 12. Art. No.: CD006777. DOI: 10.1002/14651858.CD006777.pub2.

Ancker et al 2006. Design features of graphs in health risk communication. A systematic review. JAMIA 13(6): 608

Apter A, Paasche-Orlow M, Remillard J, Bennett I, Ben-Joseph E, Batista R, Hyde J, Rudd R. Numeracy and Communication with Patients: They Are Counting on Us. J Gen Intern Med 2008, 23(12): 2117–24.

Bafeta A, Trinquart L, Seror R, Ravaud P. Analysis of the systematic reviews process in reports of network meta-analyses: methodological systematic review. BMJ 2013, 347:f3675.

Bafeta A, Trinquart L, Seror R, Ravaud P. Reporting of results from network meta-analyses: methodological systematic review. BMJ 2014, 348:g1741.

Bell CM, Urbach DR, Ray JG, Bayoumi A, Rosen AB, Greenberg D, Neumann PJ. Bias in published cost effectiveness studies: systematic review. BMJ 2006, 332:699-703.

Brownson R, Dodson E, Stamatakis K, Casey C, Elliott M, Luke D, Wintrode C, Kreuter M (2011). Communicating Evidence-Based Information on Cancer Prevention to State-Level Policy Makers. J Natl Cancer Inst; 103:306–316.

Brouwers, M. C., Stacey, D., & O'Connor, A. M. (2013). Knowledge translation tools. In S. Straus, J. Tetroe & I. D. Graham (Eds.), *Knowledge translation in health care: Moving from evidence to practice* (pp. 50-62). Chichester: Wiley Blackwell BMJ Books.

Burkell J. What are the chances? Evaluating risk and benefit information in consumer health materials. J Med Libr Assoc 2004, 92(2): 200-208.

CADTH. Guidelines for the economic evaluation of health technologies: Canada [3rd Edition]. Ottawa: Canadian Agency for Drugs and Technologies in Health; 2006.

CADTH Health Technology Assessment Report. Indirect Evidence. Indirect treatment comparisons in meta-analysis. Ottawa: Canadian Agency for Drugs and Technologies in Health; 2009.

Canadian Institutes of Health Research. More About Knowledge Translation. http://www.cihr-irsc.gc.ca/e/39033.html [Last accessed: August 2014]

Canadian Institutes of Health Research. Guide to Knowledge Translation Planning at CIHR: Integrated and End-of-Grant Approaches. Ottawa: CIHR, 2012

Canadian Institutes of Health Research. Guide to Researcher and Knowledge User Collaboration in Health Research. Ottawa: CIHR.

Cates Plot and Calculator. Available at: http://www.nntonline.net/visualrx/cates_plot/ [Last accessed: August 2014.]

Chambers et al. Maximizing the impact of systematic reviews in health care decision making. A systematic scoping review of knowledge translation resources. Milbank Q, 2011, 89: 131-156.

Chase D, Rosten C, Turner S, Hicks N, Milne R. Development of a toolkit and glossary to aid in the adaptation of health technology assessment (HTA) reports for use in different contexts. Health Technol Assess 2009; 13(59).

Cleemput I, Neyt M, Thiry N, De Laet C, and Leys M. Using threshold values for cost per quality adjusted life year gained in healthcare decisions. Intl J. Technology Assessment in Health Care, 2011, 21: 71-76.

Cookson R, Maynard A. Health technology assessment in Europe: Improving clarity and performance. Int J Technol Assess Health Care 2000;16: 639-650.

Cooper NJ, Peters J, Lai MC, Juni P, Wandel S, Palmer S et al. How valuable are multiple treatment comparison methods in evidence-based health-care evaluation? Value Health 2011, 14(2): 371-80.

Daniels N & Sabin J. Limits to health care: Fair procedures, democratic deliberations and the legitimacy problem for insurers. Philosophy & Public Affairs 1997; 303-350.

Daw JR & Morgan S. Stitching the gaps in the Canadian public drug coverage patchwork? A review of provincial pharmacare policy changes from 2000 to 2010. Health Policy 2012, 104: 19–26.

Dias S, Welton NJ, Sutton AJ, Ades AE. NICE DSU Technical Support Document 1: Introduction to evidence synthesis for decision making. 2011; last updated April 2012; available from http://www.nicedsu.org.uk [Last accessed: January 2014]

Dias S, Welton NJ, Sutton AJ, Ades AE. NICE DSU Technical Support Document 2: A generalised linear modelling framework for pair-wise and network meta-analysis of randomised controlled trials. 2011; last updated April 2012; available from http://www.nicedsu.org.uk [Last accessed: January 2014]

Dias S, Sutton AJ, Welton NJ, Ades AE. NICE DSU Technical Support Document 3: Heterogeneity: subgroups, meta-regression, bias and bias-adjustment. 2011; last updated April 2012; available from http://www.nicedsu.org.uk [Last accessed: January 17, 2014]

Dias S, Welton NJ, Sutton AJ, Caldwell DM, Lu G. & Ades AE. NICE DSU Technical Support Document 4: Inconsistency in Networks of Evidence Based on Randomised Controlled Trials. 2011; last updated April 2012; available from http://www.nicedsu.org.uk [Last accessed: January 2014]

Dobbins M, Jack S, Thomas H, Kothari A. Public Health Decision-Makers' Informational Needs and Preferences for Receiving Research Evidence. Worldviews on Evidence-Based Nursing 2007; 4(3), 156–163.

Donegan S, Williamson P, Gamble C, Tudor-Smith C. Indirect comparisons. A review of reporting and methodological quality. PLoS One 2010, 5(11):e11054.

Drummond M. Twenty years of using economic evaluations for drug reimbursement decisions: what has been achieved? J Health Polit Policy Law. 2013;38(6): 1081-102.

Drummond MF, Sculpher MJ, Torrance GW, O'Brien BJ, and Stoddard GL (2005). *Methods for the Economic Evaluation of Health Care Programmes.* 3rd edn. Oxford University Press Inc., New York.

Elwyn G, O'Connor A, Stacey D, Volk R, Edwards A, Coulter A et al., on behalf of the International Patient Decision Aids Standards (IPDAS) Collaboration. A quality criteria framework for patient decision aids: online international Delphi consensus process. BMJ, doi:10.1136/bmj.38926.629329.AE (published 14 August 2006)

Erntoft S. Pharmaceutical priority setting and the use of health economic evaluations: a systematic literature review. Value in Health 2011;14: 587-599.

Galani C and Rutten FFH. Self-reported health care decision makers' attitudes towards economic evaluations of medical technologies Cur Med Res Op 2008, 24: 3049-3058

Garattini L, Koleva D, Casadei G. Modeling in pharmacoeconomic studies: funding sources and outcomes. Int J Technol Assess Health Care 2010; 26: 330-333.

Graham ID, Logan J, Harrison MB, Straus SE, Tetroe J, Casell W et al. Lost in knowledge translation: time for a map: J Contin Ed Health Prof 2006; 26(1):13-24.

Health Technology Assessment international (HTAi). Available at: http://www.htai.org [Last accessed: August 7, 2014]

Health Technology Assessment international (HTAi). Definition of 'health technology assessment'. Available at: http://www.htai.org [Last accessed, August 2014]

Hildon Z, Allwood D, Black N (2012). Impact of format and content of visual display of data on comprehension, choice and preference: a systematic review. Int J Qual Health Care, 24 (1): 55-64.

Hoaglin D, Hawkins N, Jansen J, Scott D, Itzler R, Cappelleri J, Boersma C, Thompson D, Larholt K, Diaz M, Barrett A (2011). 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 in Health 14:429-43.

Hutton B, Salanti G, Chaimani A, Caldwell DM, Schmid C, Thorlund K et al. The quality of reporting methods and results in network meta-analyses. An overview of reviews and suggestions for improvement. PLoS One 2014, 9(3): e92508.

Ioannidis J. Geometry of the evidence. Agenda-wide views of research. Presented at: Plenary session. The Challenges of Reviewing Science in the XXI Century. 19th Cochrane Colloquium 2011; 2011 Oct 19-22; Madrid, Spain. Available from http://www.cochrane.org/multimedia/multimedia-cochrane-colloquia-and-meetings/colloquium-madrid-2011 [Last accessed: August 2014]

Innvaer, S., Vist, G., Trommald, M., & Oxman, A. (2002). Health policy-makers' perceptions of their use of evidence: A systematic review. Journal of Health Services Research and Policy, 7, 239-244.

International Network of Agencies for Health Technology Assessment (INAHTA). Available at: http://www.inahta.org [Last accessed: August 2014]

Jansen J, Fleurence R, Devine B, Itzler R, Barrett A, Hawkins N, Lee K, Boersma C, Annemans L, Cappelleri J. 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 in Health, 2011, 14: 417–428.

Kleijnen S, George E, Goulden S, d'Andon A, Vitré P, Osinska B. Relative Effectiveness Assessment of Pharmaceuticals: Similarities and Differences in 29 Jurisdictions. Value in Health, 2012, 15: 954–960.

Lavis JN, Davies HT, Gruen RL, Walshe K, Farquhar CM. Working within and beyond the Cochrane Collaboration to make systematic reviews more useful to healthcare managers and policy makers. Healthc Policy, 2006 1: 21-33.

Lavis, J., Davies, H., Oxman, A., Denis, J. L., Golden-Biddle, K., & Ferlie, E. Towards systematic reviews that inform health care management and policy-making. Journal of Health Services Research and Policy, 2005, 10 (Suppl 1), 35-48.

Lehoux P & Williams-Jones B. Mapping the integration of social and ethical issues in health technology assessment. International Journal of Technology Assessment in Health Care, 2007, 23: 9-16.

Leonard D. Translating and communicating patient-centred research to the public. Am J Pharm Benefits, 2010, 2: 408-410.

Li T, Puhan M, Vedula S, Singh S, Dickersin K. Network meta-analysis. Highly attractive but more methodological research is needed. BMC Medicine 2011 9:79.

Madden L, King L, Shiell A. How do government health departments in Australia access health economics advice to inform decisions for health? A survey. Aust New Zealand Health Policy. 2009, 6:6

Mathes T, Jacobs E, Morfeld JC, Pieper D. Methods of international health technology assessment agencies for economic evaluations--a comparative analysis. BMC Health Serv Res. 2013 Sep 30;13:371.

McCaffery K, Sheridan S, Nutbeam D, Clayman M, Kelly Blake K, Holmes Rovner M, Rovner D, Smith S, Wolf M. (2012). Addressing health literacy. In Volk R & Llewellyn Thomas H (editors). 2012 Update of the International Patient Decision Aids Standards (IPDAS) Collaboration's Background Document. Chapter J Available at: http://ipdas.ohri.ca/resources.html.[Last accessed: August 9, 2014]

Mills E, Ionnidis J, Thorlund K, Schunemann H, Pulan M, Guyatt G. How to use an article reporting a multiple treatment comparison meta-analysis. JAMA 2012; 308: 1246-1253.

Mitton CR. Priority seeting for decision makers: using health economics in practice. Eur J Health Econom 2002, **3:** 240-243.

Morgan SG, McMahon M, Mitton C, et al. Centralized drug review processes in Australia, Canada, New Zealand, and the United Kingdom. Health Aff (Millwood) 2006;25:337–47.

Morgan S, Thomson P, Daw J, Friesen M. Inter-jurisdictional cooperation on pharmaceutical product listing agreements: views from Canadian provinces. BMC Health Services Research 2013; 13:34.

Murthy L, Shepperd S, Clarke MJ, Garner SE, Lavis JN, Perrier L, Roberts NW, Straus SE. Interventions to improve the use of systematic reviews in decision-making by health system managers, policy makers and clinicians. Cochrane Database of Systematic Reviews 2012, Issue 9. Art. No.: CD009401. DOI: 10.1002/14651858.CD009401.pub2

National Pharmaceuticals Strategy, Federal/Provincial/Territorial Ministerial Task Force. National Pharmaceuticals Strategy Progress Report. Health Canada: Her Majesty the Queen in Right of Canada. Ottawa, Canada. June 2006.

Niessen LW, Bridges J, Lau BD, Wilson RF, Sharma R, Walker DG, Frick KD, Bass EB. Assessing the Impact of Economic Evidence on Policymakers in Health Care. A Systematic Review. Methods Research Report (Prepared by the Johns Hopkins University Evidence-based Practice Center under contract No. 290-2007-10061-I). AHRQ Publication No. 12(13)-EHC133-EF. Rockville, MD: Agency for Healthcare Research and Quality. October 2012. www.effectivehealthcare.ahrq.gov/reports/final.cfm.

Niezen M, Stolk E, Steenhoek A, Uyl-De Groot C. Inequalities in oncology care: Economic consequences of high cost drugs. Eur J of Cancer, 2006, 42: 2887–2892.

Nikolakopoulou A. Chaimani A. Veroniki AA. Vasiliadis HS. Schmid CH. Salanti G. Characteristics of networks of interventions: a description of a database of 186 published networks. PLoS ONE 2014, 9(1):e86754.

Oortwijn W1, Broos P, Vondeling H, Banta D, Todorova L Mapping of health technology assessment in selected countries. Int J Technol Assess Health Care. 2013, 29(4):424-34.

Perrier L, Mrklas K, Lavis J, Straus S. Interventions encouraging the use of systematic reviews by health policymakers and managers: A systematic review. Implementation Science 2011, 6:43.

Pocock S. Safety of drug-eluting stents. Demystifying network meta-analysis. Lancet, 2007, 370:2099-100

Rawlins MD, Chalkidou K. The opportunity cost of cancer care: a statement from NICE. Lancet Oncol. 2011 Sep;12(10):931-2.

Registered Nurses' Association of Ontario. (2012). Toolkit: Implementation of best practice guidelines (2nd ed.). Toronto, ON: Registered Nurses' Association of Ontario (RNAO).

Roberts M, Russell LB, Paltiel AD, Chambers M, McEwan P, Krahn M; ISPOR-SMDM Modeling Good Research Practices Task Force. Conceptualizing a model: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force--2. Value Health 2012, 15(6):804-11.

Salanti G, Ades AE, Ioannidis JPA. Graphical methods and numerical summaries for presenting results from multiple-treatment meta-analysis: an overview and tutorial. J Clin Epidemiol 2011;64:163-71.

Salanti G, Del Giovane C, Chaimani A, Caldwell DM, Higgins JP. Evaluating the quality of evidence from a network meta-analysis. PLoS One 2014, 9(7): e99682.

Schlander M. The use of cost-effectiveness by the National Institute for Health and Clinical Excellence (NICE): no(t yet an) exemplar of a deliberative process. J Med Ethics 2008; 34: 534-539.

Scotland G, Waugh N, Royle P, McNamee P, Henderson R, Hollic R. Denosumab for the Prevention of Osteoporotic Fractures in Post-Menopausal Women A NICE Single Technology Appraisal. Pharmacoeconomics 2011; 29 (11): 951-961.

Song F, Loke YK, Walsh T, Glenny AM, Eastwood AJ, Altman DG. Methodological problems in the use of indirect comparisons for evaluating healthcare interventions: survey of published systematic reviews. BMJ 2009, 338: b1147.

Straus S, Tetroe J & Graham ID (Eds.), Knowledge translation in health care: Moving from evidence to practice, 2nd edition. Oxford: Blackwell Publishing, 2013.

Sullivan R, Peppercorn J, Sikora K, Zalcberg J, Meropol N, Amir E. Delivering affordable cancer care in high-income countries. Lancet Oncology, 2011, 12: 933-980.

Thurston SJ, Craig D, Wilson P, Drummond MF. Increasing decision-makers' access to economic evaluations: alternative methods of communicating the information. Int J Technol Assess Health Care. 2008, 24(2):151-7.

Tierney M, Manns B; Members of the Canadian Expert Drug Advisory Committee. Optimizing the use of prescription drugs in Canada through the Common Drug Review. CMAJ. 2008, 178(4):432-5.

Thabane L, Wells G, Cook R, Platt R, Pogue J, Pullenayegum E, et al. Canadian-led capacity-building in biostatistics and methodology in cardiovascular and diabetes trials: the CANNeCTIN Biostatistics and Methodological Innovation Working Group. Trials 2011, 12:48.

Trevena L, Davey H, Barratt A, Butow P, Caldwell P. A systematic review on communicating with patients about evidence. J Evalu Clin Practice, 2006, 12: 13-23.

Trevena L, Zikmund-Fisher B, Edwards A, Gaissmaier W, GAlesic M, Han P, King J, LAwson M, Linders S, Lipkus I, Ozanne E, Peters E, Timmermans D, Woloshin S (2012). Presenting Probabilities. In Volk R & Llewellyn Thomas H (editors). 2012 Update of the International Patient Decision Aids Standards (IPDAS) Collaboration's Background Document. Chapter J. Available at: http://ipdas.ohri.ca/resources.html

Visschers V, Meertens R, Passchier W, de Vries N. Probability Information in Risk Communication: A Review of the Research Literature. Risk Analysis, 2009, 29: 267-287.

Wensing, M., Bosch, M., Grol R. (2013). Developing and selecting knowledge translation interventions. In S. Straus, J. Tetroe & I. D. Graham (Eds.), *Knowledge translation in health care: Moving from evidence to practice*. Chichester: Wiley Blackwell BMJ Books.

Williams I, McIver S, Moore D, Bryan S. The use of economic evaluations in NHS decision-making: a review and empirical investigation. Health Technol Assess 2008; 12: 1-175.

Winquist E, Bell CM, Clarke JTR, Evans G, Martin J, Sabharwal M, Gadhok A, Stevenson H, Coyle D. An evaluation framework for funding drugs for rare disease. Value in Health 2012:15; 982–986.

CHAPTER 3

KNOWLEDGE, BARRIERS, FACILITATORS AND APPROACHES TO COMMUNICATING ECONOMIC EVALUATIONS: A SYSTEMATIC REVIEW

CHAPTER OVERVIEW

Purpose and Relationship to Other Manuscripts:

One of the goals of this thesis is to understand barriers and facilitators to communicating economic evaluations from the perspective of end-users. Therefore, the objective of this first manuscript is to systematically review the literature on knowledge, barriers, facilitators and approaches to communicating and understanding health economic information by non-technical audiences, i.e. knowledge users. This work will be used to identify the informational needs and optimal presentation of economic evaluations for lay audiences. The majority of studies identified in the systematic review were related to policymakers; therefore, findings are more applicable to policymakers than other non-technical audiences.

Supporting Material:

Appendix A. Abstract presented at Cochrane Colloquium 2012.

- Note: This abstract included interim findings from this systematic review. Interim results reported in this abstract were revised and updated (to December 2013) in the final version of this thesis.

Manuscript Status:

Submitted for review and awaiting a decision. All information in this chapter, including the supplementary materials, has been submitted to the journal.

Authors:

Shannon M Sullivan (SMS), Zahra Akbar (ZA) Chaim Bell (CB), Mona Sabharwal (MS), Doug Coyle (DC)

Author Roles (SMS, ZA, CB, MS, DC) and Contributions:

SMS, CB, MS, ZA were responsible for the study conception.

SMS, ZA were responsible for the search for trials.

SMS, ZA, DC were responsible for the selection of trials.

SMS, ZA were responsible for the data extraction.

SMS, ZA were responsible for the data analysis.

SMS, ZA, CB, MS, DC were responsible for the interpretation of results.

SMS, ZA were responsible for drafting the manuscript.

All authors (SMS, ZA, CB, MS, DC) provided comments on drafts and revisions to the manuscript, agree with the manuscript's results and conclusions, and read and approved the final manuscript.

TITLE:

Presenting and Communicating Health Economic Information: A Systematic Review.

SHORT TITLE:

Communicating Health Economic Information

ABSTRACT:

Objective: To conduct a systematic review of barriers, facilitators and approaches to understanding and communicating health economic information to lay audiences (i.e., non-economists), for the purpose of developing knowledge translation tools.

Methods: Electronic databases and supplementary sources were searched for studies in lay audiences that reported approaches to presenting or communicating health economic information, understanding and knowledge of economics and barriers and facilitators to using health economic information. Frequencies were tabulated and qualitative data synthesis conducted.

Results: Thirty-three studies were included, most of which reported on surveys and interviews. Participants were primarily policy-makers but also included clinicians and the public. Strategies that improved communication of health economic information included clarifying technical concepts, focusing on key principles, improving readability and tailoring the level of detail to end-users' needs. Despite a strong interest, participants' practical experience using health economic information was low. Individuals had high-level knowledge and understanding of economic concepts but had difficulty applying concepts to real-world decisions. Facilitators to using economic information included supportive policies, good communication approaches, economic training and having high-quality and relevant information. Barriers to using economic information included limited economic knowledge, perceived lack of methodological credibility, poor communication, and limited relevance.

Conclusions: Effective strategies for communicating economic evaluations to a lay audience require the development of novel approaches and presentation formats. Tools that disaggregate results, focus on practical applications and are tailored to end-users could make economic evaluations more accessible and increase confidence in and legitimacy of health technology assessment decision-making processes.

KEY WORDS:

economic evaluation; decision-making; communication; knowledge translation; systematic review

ACKNOWLEDGEMENTS:

George Wells (University of Ottawa Heart Institute), Barbara Davies (University of Ottawa) and Elaine McPhail (Canadian Agency for Drugs and Technologies in Health) provided comments on the design, analysis and draft manuscripts of this study.

SOURCES OF FUNDING:

This study was funded, in part, by the pan-Canadian Oncology Drug Review and the Canadian Institutes of Health Research.

INTRODUCTION

Health technology assessments (HTAs) incorporate information about effectiveness, safety, and sometimes cost-effectiveness of drugs and health technologies to support evidence-informed health policy and decision-making (1). Economic evaluations of these drugs and health technologies provide policy makers with an assessment of the technology's "value for money" and are one factor that can help determine the optimal allocation of resources (2). While advancements in medical sciences may increase, the collective resources that society can spend on health technologies may not increase correspondingly. Therefore, decisions must be made on how finite resources are allocated and which technologies are funded or not funded. Increasingly, lay audiences (i.e. non-economists), are looking to better understand economic evaluations and how public funds are being allocated. This may include the public, patients or bodies set up by democratic societies to provide checks and balances in public policy (e.g. ombudsmans, auditors, human rights tribunals, citizen panels) (3) As such, audiences who do not have a technical background in health economics require tools that will help them to interpret, understand, and use economic evaluations.

Policy-makers must also clearly communicate economic concepts and the rationale for why a particular drug or technology is funded or not funded to clinicians and the public. As such, they should have access to understandable health economic information, even when it is based on complex modeling and technical concepts such as incremental cost effectiveness ratios and quality adjusted life years. Although economic evaluations have been used for the past twenty years, their policy impact is still unclear (4). More effective presentation of economic evaluations could increase understanding of how economic evaluations inform policy decisions, improve policy makers' communication of resource allocation decisions, and enhance patient and clinician understanding of the role of health economics in policy decision making. All of this would increase the confidence and perceived legitimacy of HTA decision-making processes (5). However, achieving the appropriate balance of making information understandable while maintaining the

completeness and technical accuracy of the economic evaluation can be challenging as important nuances may be lost by oversimplifying information.

In order to provide more understandable health economic information, a review of current approaches to communicating and presenting economic information is required. In addition, a review of lay audiences' knowledge of health economics as well as barriers and facilitators to using economic information will inform effective approaches to communicating health economic information. The objective of the present study was to conduct a systematic review of barriers, facilitators and approaches to using, understanding and communicating health economic information for the purposes of informing the optimal presentation of economic evaluations for lay audiences and the development of knowledge translation tools.

METHODS

Review Protocol: A systematic review was conducted following Cochrane methodology (6). All empirical study designs with a defined methodology were included (e.g., systematic reviews, randomized controlled trials (RCTs), observational studies, interviews, focus groups or surveys). The population of interest was lay audiences who were not economists but who may be end-users of health economic information. This included policy makers, practicing clinicians, adult patients or other members of the public. Study populations were classified based on their role, therefore, clinicians or patient/public representatives who play a role in healthcare policy or decision-making were considered policy-makers for the purposes of this review. However, subgroup analyses among different types of policy-makers were considered when available. Practicing clinicians were not limited to physicians but also included other health professions, e.g., pharmacy, nursing, occupational health. The intervention and comparator were methods, interventions or tools used to convey or represent economic information or to identify barriers and facilitators to understanding economic information. Outcomes of interest were: individuals' knowledge or understanding of economic information; barriers to the use of economic

information; facilitators encouraging use of economic information; different approaches or preferences for communicating or presenting economic information.

Search Strategy: Databases searched included Medline (1948 to December 2013), EMBASE (1980-December 2013), and PsychInfo (1987-December 2013) using no language or date restriction. The search concepts were 'communication', 'decision making', 'information dissemination', 'economics', 'health economics' and 'economic evaluation'. Grey literature was searched for unpublished reports using the CADTH Grey Matters checklist as a guide and supplemented by other unpublished sources (7). Studies were also obtained through hand searching of selected journals, reviewing reference lists of potentially relevant studies and suggestions from economic experts.

Study Selection and Data Extraction: Citations were screened for relevance by one review author based on the title and abstract of identified articles. Two review authors independently reviewed the full text of potentially relevant articles to assess exclusion or inclusion. One author extracted data, which was verified by a second author. Any disagreements were resolved by a third reviewer. Study quality was assessed using AMSTAR for systematic reviews (8) and the Cochrane Risk of Bias tool for RCTs (6). Critical appraisal of qualitative studies was informed by recommendations from the Cochrane Collaboration and checklists developed by Greenhalgh et al 2004 (9,10).

Data Analysis: The frequency with which outcomes were reported was tabulated across studies. If studies reported on more than one outcome, they could be counted more than once in subsequent analyses. Due to the heterogeneity of outcome measurement and the qualitative design of most studies, data was synthesized qualitatively (6,11). Descriptive themes were identified following a review of the primary data in order to translate, or categorize, concepts that were reported in various ways in the primary studies. Descriptive themes from the primary studies were then reviewed and an analytic synthesis was conducted to integrate data and

generate additional themes for this systematic review. Trends were identified, where possible, based on key study characteristics including population, geographic region, timeframe, and use of economic information.

RESULTS

Study Characteristics

Of the 5682 citations identified (see Supplementary Figure), 33 studies were included: 9% (n=3) were systematic reviews; 3% (n=1) were RCTs; 70% (n=23) were surveys, interviews and/or focus groups; and 18% (n=6) were mixed methods studies (see Table 1) (12-44).
Seventy percent (n=23/33) of studies reported on policy-makers (11-33), 15% (n=5/33) on practicing clinicians (34-38), 3% (n=1/33) on the general public (39) and 12% (n=4/33) on a heterogeneous mixed population that included both policy-makers and practicing clinicians (40-43). No studies included patients although some studies of policy makers may have incorporated the views of patient representatives who are involved in healthcare policy decision-making.
Studies were published over an approximately 20-year period from 1994 to 2013: 12% (n=4/33) up to 1999, 67% (n=22/33) from 2000-2009 and 21% (n=7/33) from 2010 to present. The studies were conducted in a variety of geographic regions with differing healthcare policy environments: 58% (n=19/33) in Europe, 9% (n=3/33) in the USA, 6% (n=2/33) in each of Canada, Australia and Asia, 3% (n=1/33) in South America and 12% (n=4/33) across multiple regions.

INSERT TABLE 1. Characteristics of Included Studies.

Studies were heterogeneous with respect to study objective (see Supplementary Table 1). However, the majority of studies (88%, n=29/33) focused on formal economic evaluations (e.g., cost-utility analyses, cost-benefit analyses) while the remainder (12%, n=4/33) reported on other types or more specific types of economic information (e.g., willingness to pay, drug costs).
All of the studies that included policy-makers alone or a mixed population focused on using economic information to inform resource allocation decisions. Eighty percent (n=4/5) of studies in

clinicians alone also considered using economic information to inform prescribing practices. The one study conducted in the general public used economic information to inform hospital choice. Other uses that were considered alongside resource allocation included priority setting, healthcare planning, policy development and reimbursement (see Supplementary Table 1). Study quality was assessed and found to be low to moderate when considering the overall body of evidence from the 33 included studies (see Supplementary Table 2). Low response rates and the use of convenience or purposive samples in survey or interview studies may have created response and selection bias. In addition, in the survey studies, the validity of surveys was not always reported. The majority of data were collected through surveys and interviews and may have a self-report bias.

Communication and Presentation

Forty-eight percent (n=16/33) of studies reported on the communication and presentation formats of health economic information. How outcomes were measured and reported varied across studies. However, three areas were frequently commented upon: the format of information, the type of information and preferences for including specific economic concepts when presenting information. Key analytic themes that emerged when synthesizing information from these three areas are summarized in Table 2 while specific details of the outcomes reported in individual studies are provided in Supplementary Table 3. None of the studies in clinicians reported on communication or presentation formats. The one study conducted in the public was an RCT comparing different presentation formats; findings focused on specific formatting approaches that facilitate understanding of general cost information.

INSERT TABLE 2. Themes Related to Communicating and Presenting Health Economic Information.

Knowledge and Understanding

Eight-eight percent (n=29/33) of studies reported on knowledge and understanding of health economic information. How outcomes were measured and reported varied across studies. However, four areas were frequently assessed that addressed knowledge and understanding: attitudes towards or interest in learning or using health economics (n=22/29); experience using economic information (n=14/29); training in health economics (n=11/29) and self-reported or empirically measured knowledge of health economics (n=16/29). Key analytic themes that emerged when synthesizing information from these four areas are summarized in Table 3 while details of the outcomes reported in individual studies are provided in Supplementary Table 4.

INSERT TABLE 3. Themes Related to Knowledge of Health Economic Information

Barriers and Facilitators to Use

There were 79% (n=26/33) studies reporting on barriers to using economic information and 67% (n=22/33) reporting on facilitators. Sixty-five percent (n=17/26) of studies reported poor communication as a type of barrier to using economic information and 45% (n=10/22) reported good communication as a type of facilitator of using economic information. The specific barriers and facilitators to communicating economic information that were reported in individual studies are provided in Table 4.

INSERT TABLE 4. Specific Barriers and Facilitators to Communicating Heath Economic Information as Reported in Individual Studies.

Other types of barriers frequently reported in individual studies (see Figure 1) included lack of knowledge of health economics (73%, n=19/26), perceived lack of methodological credibility (69%, n=18/26), limited relevance (65%, n=17/26), budgetary constraints (65%, n=17/26), perceived industry bias (54%, n=14/26), organizational, political or ethical barriers (50%, n=13/26) and a conflict between clinical and economic values (42%, n=11/26). Two barriers specific to

clinicians' use of economic information to inform prescribing practices were a perceived conflict between clinical and economic values and clinicians' lack of awareness of drug costs. Barriers to use were not reported in the one study conducted in the general public.

Other frequently reported types of facilitators (see Figure 1) were related to having policies in place that support use of economic evaluations (54%, n=12/22), having economic training available (41%, n=9/22), being provided with high-quality (41%, n=9/22) and relevant economic information (36%, n=8/22), strong collaborations with economists (36%, n=8/22), flexible budgets (36%, n=8/22) and having awareness of political or ethical issues (32%, n=7/22). No additional or predominant facilitators specific to clinicians were identified.

INSERT FIGURE 1. Most Frequently Reported Types of Barriers and Facilitators to Using Economic Information.

See Supplementary Table 5 for more details on the individual barriers and facilitators reported in other categories and their distribution across studies (Supplementary Tables 6 and 7).

DISCUSSION

The majority of studies in this systematic review reported on policy makers. Qualitative synthesis found that communication of economic information could be improved by clarifying technical concepts and assumptions, focusing on key principles over methodological details and disaggregating results. It was also suggested that readability could be enhanced and that the included level of detail should depend on the end-user's needs. The synthesis also found that although practical experience with economic evaluations is low, there is strong interest in using or learning about it. There was some training in health economics among the study populations, however, knowledge focused on high-level economic concepts and individuals still had difficulty understanding how to practically apply these concepts to real-world decisions. This review identified several factors that facilitated using economic information: supportive policies, good

communication, economic training, having high-quality and relevant economic information and strong collaborations with economists. Barriers to using economic information included limited knowledge of economics, information lacking methodological credibility, poor communication and limited relevance of the information.

Policy Implications and Relevance to Existing Literature

These findings suggest that opportunities exist to improve the communication and presentation of health economic information to relevant stakeholders. In an environment of limited budgetary resources, where there is strong interest to use economic evaluations and to better understand health economic principles, improved knowledge translation of health economic information can only be a benefit. A key principle of scientific study and decision-making in HTA processes is transparency of reporting (45). However, full transparency requires providing information in an accessible and understandable format (46,47). A number of barriers and facilitators to communicating health economic information were identified in the literature, which could be addressed or leveraged to support better presentation of economic information. Building upon and integrating the different outcomes in this review with existing literature, three key principles to inform knowledge translation tools for economic evaluations are proposed: (1) disaggregated presentation of key principles and results (2) focusing on practical applications and real-world implications of findings (3) tailoring presentation formats to end-users. When comparing these principles to guidelines for conducting economic evaluations, there are some similarities with good reporting practices (48,49). However, while good reporting practices are important and can contribute to clarity, they are usually targeted at health economists and insufficient for good communication to lay audiences.

(1) Disaggregated Presentation of Key Principles and Results

Presenting economic concepts such as an ICER in a disaggregated format that identifies component costs and effects could address these some barriers and better clarify and explain concepts for audiences (16,18,20,30,41). Clearly outlining model assumptions that contribute to

an ICER is also an approach that can facilitate communication of economic information (17). With a clearer breakdown of economic methodology and results, the end-user may be able to better understand the applicability and relevance of the findings to their specific decision-making context. Frequently cited barriers were the inability to transfer or generalize economic evaluations to local decision-making contexts; (15-16,18,22-23,25,42-44) however, one must first be able to understand the components of the economic analysis to determine if results are generalizable. The end-users' ability to deconstruct the economic analysis (30) can also increase credibility in the quality of the analysis and mitigate any perceived sponsorship biases, which are other barriers that have been cited (13-16,19,22,33-34,37,39,41-44). Consistent efforts are required to make detailed economic evaluations as simple as possible, without altering the meaning of complex modeling and techniques, so that audiences can understand it accurately and efficiently. This is challenging, however, clearly breaking down and outlining key components of the evaluation is one possible approach that would maintain the integrity of the technical concepts.

(2) Practical Applications of Results and Implications for Real-World Decision-Making

The second proposed approach is to focus the presentation of economic evaluations on the practical application of results. Study populations had some training in health economics but little practical experience with economic evaluations (12,16-17,20,22,29,34,38,43). Perhaps as a consequence, in some studies individuals reported they understood concepts such as 'cost-effectiveness analysis' and 'quality-adjusted life year', but were unable to apply these concepts in practice (15,27). Policy makers were particularly challenged in determining how numbers established by economic modeling studies and theoretical thresholds are best applied to real-world decisions on resource allocation.

Individual studies often indicated that cost savings identified in economic evaluations were only theoretical and would not result in actual cost savings based on how healthcare decisions are implemented (14,19,22,33,39,41,44). Being able to explain the practical relevance of results and any implications for cost savings was considered a facilitator to communicating economic

evaluations (22). Cost-effectiveness analyses, while distinct from budget impact analyses, are closely linked; however, this relationship is not always clearly presented. Clarifying the difference between whether or not a technology is cost-effective and whether or not it is affordable, may assist audiences trying to understand or to make reimbursement decisions. Opportunity costs are not frequently discussed in economic evaluations and may be challenging to explicitly determine due to local variations in budgets. However, opportunity costs are a concept with which lay audiences are familiar, as they must make decisions in their daily lives on what they can afford and what they cannot afford if money is spent on other items. Providing scenarios describing the opportunity costs of decisions or discussion points alerting audiences to implications for affordability could increase the relevance of economic analyses.

Implementing resource allocation decisions may also require consideration of economic factors that are not explicitly considered in the cost-effectiveness analysis but that could be noted to enhance their relevance (e.g., molecular testing, patent expiries, product listing agreements) (50-52). The ethical, political or organizational context can also influence use of economic information in HTA decision-making (14-15,20,24-25,29,33-34,37,39,41,43-44). Acknowledging potential political or ethical factors and highlighting related components of the economic evaluations (e.g. subgroup analyses, equity considerations, availability and accessibility of comparators) could increase relevance for policy-making. Tools that present different scenarios of how results could be applied to real-world decisions may further assist policymakers in understanding how to transfer theoretical cost-effectiveness results to specific decisions and how to address issues of affordability.

(3) Tailoring Presentation Formats to Different End-Users

The majority of included studies reported on policy makers, however, there are other important end-users each of whom has a unique perspective and informational needs. Studies of clinicians indicated they experienced barriers such as a perceived conflict with clinical values and a lack of awareness of drug costs when using economic information (36-37,42-44). Thus, presenting

economic information to clinicians in a way that acknowledges these barriers could be important. One study found that, in the general population, presentation formats that enhanced clarity and readability were particularly important for individuals with low numeracy skills (40). For example, placing important findings first, removing unnecessary detail and using formatting approaches such as headings and section breaks enhances understanding (23). A study of policy makers using economic abstracts found that format preferences differed, depending on how they were using the information, e.g., for screening, for detailed review (31). Therefore, even within types of end-users, there is a need to tailor the presentation of information. These findings align with knowledge translation principles that advocate tailoring information to end-users (53). Approaches such as layering detailed information and using multiple formats (e.g. text, graphics) to present the same information can help address differences in end-users' needs. In the context of current literature, considerable research has explored how to tailor clinical evidence to policymakers, clinicians and patients when creating briefing notes, practice guidelines and decision-aids, respectively (54-56). Similar concepts could be applied to tools conveying health economic information. This current systematic review relies predominantly on studies in policy makers; further primary research may be required to better elicit specific needs of practicing clinicians, patients and the public. The small number of studies in these latter populations suggests they are not traditionally viewed as end-users of economic information. However, this could also be because economic information is not presented in an accessible manner; tailored presentation approaches could enhance its use in these relevant stakeholder populations.

Strengths and Limitations

Our study's findings are strengthened by its rigorous systematic review methodology and approach of considering multiple related outcomes with the intent of enhancing communication of economic evaluations. Although some systematic reviews have reported on the individual outcomes (26, 34, 41-42), none simultaneously reported on and integrated findings across outcomes relating to communication and presentation preferences, knowledge of economic information and barriers and facilitators to using economic information. This is also one of the

first reviews that has attempted to look at the use of economic information in less traditional audiences such as patients and the public, who have a growing interest in how public funds are allocated for drugs and health technologies (3).

Our work has limitations that merit mentioning. Data synthesis was challenging due to the broad scope of the review and diverse nature of the included studies and outcomes. However, this approach also created a relevant snapshot of the current context in which economic evaluations are understood and communicated, which has not been done previously. Although the included studies were heterogeneous and only one RCT was identified, the consistency of themes across different health care funding systems and requirements for health policy decisions, suggests the main findings are likely generalizable across different settings that use economic information. Given the predominance of studies that included key informants, purposive samples or that had low response rates, these results likely reflect informed and engaged audiences that represent 'best-case' scenarios. However, the findings suggested that even these audiences could benefit from enhanced presentations of economic evaluations. Thus, we feel that on balance, the limitations do not invalidate the main findings. Nonetheless, further study in RCTs is required to rigorously assess how health economic information is best understood and presented to lay audiences.

Conclusions

Effective communication of economic evidence to a lay audience requires the development of novel approaches and tools. Improving presentation formats will make health economic evaluations more accessible. This could ultimately result in a better understanding of the role economics plays in decision making regarding health technologies and improve transparency in decision-making processes. It is important that society is able to use economic evidence to better guide resources to where they are most needed. Policy makers must, at the same time, be able to communicate to the general tax-paying public, why and where these resources have been allocated within the health care system, when decisions have been based, in part, on economic

evidence. Our review outlines components of a knowledge translation strategy for enhancing the presentation of economic evaluations. They are proposed based on a comprehensive body of evidence exploring the knowledge of, barriers and facilitators to and communication of health economic information. Tools that disaggregate results, focus on practical applications and are tailored to end-users could make economic evaluations more accessible and increase confidence in and legitimacy of HTA decision-making processes.

REFERENCES

1. Cookson R, Maynard A. Health technology assessment in Europe: Improving clarity and performance. Int J Technol Assess Health Care 2000;16: 639-650.
2. Simoens S. Use of economic evaluation in decision making: evidence and recommendations for improvement. Drugs 2000;70: 1917-1926
3. Abelson J, Bombard Y, Gauvin FP, Simeonov D, Boesveld S. Assessing the impacts of citizen deliberations on the health technology process. Int J Technol Assess Health Care. 2013; 29: 282-289.
4. Drummond M. Twenty years of using economic evaluations for drug reimbursement decisions: what has been achieved? J Health Polit Policy Law. 2013;38(6): 1081-102.
5. Lehoux P, Denis JL, Tailliez S et al. Dissemination of health technology assessments: identifying the visions guiding an evolving policy innovation in Canada. J Health Polit Policy Law 2005;30: 603-641.
6. Cochrane Collaboration: Cochrane Handbook for Systematic Reviews. Available from: www.cochrane.org [Last accessed March 2, 2014].
7. CADTH. Grey Matters: a practical search tool for evidence-based medicine [2nd edition]. Ottawa: Canadian Agency for Drugs and Technologies in Health; 2013.
8. Shea BJ, Hamel C, Wells GA, Bouter LM, Kristjansson E, Grimshaw J, Henry DA, Boers M. AMSTAR is a reliable and valid measurement tool to assess the methodological quality of systematic reviews. J Clin Epidemiol 2009;62: 1013-1020.
9. Hannes K. Chapter 4: Critical appraisal of qualitative research. In: Noyes J, Booth A, Hannes K, Harden A, Harris J, Lewin S, Lockwood C (editors), Supplementary Guidance for Inclusion of Qualitative Research in Cochrane Systematic Reviews of Interventions. Version 1 (updated August 2011). http://cqrmg cochrane org/supplemental-handbook-guidance 2011. Available from: URL: http://cqrmg.cochrane.org/supplemental-handbook-guidance [Last accessed March 2, 2014]

10. Greenhalgh T, Robert G, Bate P, Kyriakidou O, Macfarlane F, Peacock R. How to spread good ideas A systematic review of the literature on diffusion, dissemination and sustainability of innovations in health service delivery and organisation. London: Report for the National Co- ordinating Centre for NHS Service Delivery and Organisation R&D (NCCSDO); 2004.
11. Thomas J, Harden A. Methods for the thematic synthesis of qualitative research in systematic reviews BMC Medical Research Methodology 2008;8: 45 doi:10.1186/1471-2288-8-45
12. Ammerman AS, Farrelly MA, Cavallo DN, et al. Health economics in public health. Am J Prev Med 2009;36: 273-275.
13. Anell A, Svarvar P. Pharmacoeconomics and clinical practice guidelines. A survey of attitudes in Swedish formulary committees. Pharmacoeconomics 2000;17: 175-185.
14. Baghbanian A., Hughes I., Khavarpour F.A. Resource allocation and economic evaluation in Australia's healthcare system. Australian Health Review 2011;35: 278-283.
15. Chaikledkaew U, Lepitakpong C, Teerawattananon Y, et al. The current capacity and future development of economic evaluation for policy decision making: a survey among researchers and decision makers in Thailand. Value in Health 2009; 12: S31-S35.
16. Chen LC, Ashcroft DM, Elliott RA. Do economic evaluations have a role in decision-making in medicine management committees? A qualitative study. Pharm World Sci 2007; 29: 661-670.
17. Davies L, Coyle D, Drummond M. Current status of economic appraisal of health technology in the European community: report of the network. Soc Sci Med 1994;38: 1601-1607.
18. Drummond M, Fendrick AM, Fullerton P, et al. Use of pharmacoeconomics information – report of the ISPOR task force on us of pharmacoeconomic/health economic information in health-care decision making. Value in Health 2003;6: 407-416.
19. Duthie T, Trueman P, Chancellor J, Diez L. Research into the use of health economics in decision making in the United Kingdom-- Phase II. Is health economics 'for good or evil'? Health Policy 1999;46: 143-157.
20. Eddama O, Coast J. Use of economic evaluation in local health care decision making in England: a qualitative investigation. Health Policy 2009; 89: 261-270.

21. Gunther OH, Konig HH. Decision makers' and scientists' opinion about contingent valuation and choice experiments for measuring willingness to pay in health care: results from a survey in Germany. Int J Technol Assess Health Care 2006; 22: 351-361.

22. Hoffman C, Graf von der Schulenburg JM. The influence of economic evaluation studies on decision making: a European survey. Health Policy 2000; 52: 179-192.
23. Hoffman C, Stoykova BA, Nixon J, et al. Do health care decision makes find economic evaluations useful? The findings of focus group research in UK health authorities. Value in Health 2002;5: 71-79.
24. Lafi R. Robinson S. Williams I. Economic evaluation and the Jordan Rational Drug List: an exploratory study of national-level priority setting. Value in Health 2012; 15: 771-776.
25. Neumann PJ, Jacobson PD, Palmer JA. Measuring the value of public health systems: the disconnect between health economists and public health practitioners. Am J Public Health 2007; 98: 2173-2180.
26. Niessen LW, Bridges J, Lau BD, Wilson RF, Sharma R, Walker DG, Frick KD, Bass EB. Assessing the Impact of Economic Evidence on Policymakers in Health Care—A Systematic Review**.** Methods Research Report (Prepared by the Johns Hopkins University Evidence-based Practice Center under contract No. 290-2007-10061-I). AHRQ Publication No. 12(13)-EHC133-EF. Rockville, MD: Agency for Healthcare Research and Quality. October 2012. www.effectivehealthcare.ahrq.gov/reports/final.cfm. [Last accessed March 2, 2014]
27. Phillips CJ, Fordham R, March K, Bertranou E, et al. Exploring the role of economics in prioritization in public health: what do stakeholders think? Eur J Public Health 2010;1-7.
28. Rocchi A, Menon D, Verma S, et al. The role of economic evidence in canadian oncology reimbursement decision making: to lamda and beyond. Value in Health 2008;11: 771-783.
29. Ross J. The use of economic evaluation in health care: Australian decision makers' perceptions. Health Policy 1995; 31:103–10.

30. Taylor-Robinson D, Milton B, Lloyd Williams F, et al. Policy makers' attitudes to decision support models for coronary heart disease: a qualitative study. J Health Serv Res Policy 2008;13: 209-214.
31. Thurston SJ, Craig D, Wilson P et al. Increasing decision makers' access to economic evaluations: alternative methods of communicating the information. Intl J Technol Assess Health Care 2008;24: 151-157.
32. van Dongen JM. Tompa E. Clune L. Sarnocinska-Hart A. Bongers PM. van Tulder MW et al. Bridging the gap between the economic evaluation literature and daily practice in occupational health: a qualitative study among decision-makers in the healthcare sector. Implementation Science 2013 8: 57.
33. Williams I, Bryan S, McIver S. How should cost-effectiveness analysis be used in health technology coverage decisions? Evidence from the National Institute for Health and Clinical Excellence approach. J. Health Serv Res Policy 2007; 12: 73-9.
34. Williams I, McIver S, Moore D, Bryan S. The use of economic evaluations in NHS decision-making: a review and empirical investigation. Health Technol Assess 2008; 12: 1-175.
35. Fattore G, Torbica A. Economic evaluation in health care: the point of view of informed physicians. Value Health 2006; 9: 157-167.
36. Jansson S, Anell A. The impact of decentralised drug-budgets in Sweden -- a survey of physicians' attitudes towards costs and cost-effectiveness. Health Policy 2006; 76: 299-311.
37. Prosser H, Walley T. A qualitative study of GPs' and PCO stakeholders' views on the importance and influence of cost on prescribing. Soc Sci Med 2005; 60: 1335-1346.
38. Torbica A, Fattore G. Understanding the impact of economic evidence on clinical decision making: a discrete choice experiment. Soc Sci Med 2010;70: 1536-1543.
39. Walley T, Baron S, Cooke J, Drummond M. Economic evaluations of drug therapy: attitudes of primary care prescribing advisors in Great Britain. Health Policy 1997;41: 961-972.
40. Peters E, DieckmanN, Dixon A, et al. Less is more in presenting quality information to consumers. Med Care Res Rev 2007;64: 169-190.

41. Erntoft S. Pharmaceutical priority setting and the use of health economic evaluations: a systematic literature review. Value in Health 2011;14: 587-599.
42. Galani C, Rutten F. Self-reported health care decision makers' attitudes towards economic evaluations of medical technologies Cur Med Res Op 2008;24: 3049-3058.
43. Rubinstein A, Belizán M, Discacciati V. Are economic evaluations and health technology assessments increasingly demanded in times of rationing health services? The case of the Argentine financial crisis. Int J Technol Assess Health Care 2007;23:169-176.
44. Zwart-van Rijkom J, Leufkens G, Busschbach J, et al. Differences in attitudes, knowledge and use of economic evaluations in decision-making in the Netherlands. Pharmacoeconomics 2000;18: 149-160.
45. Dhalla I, Laupacis A. Moving from opacity to transparency in pharmaceutical policy. CMAJ 2008;178: 428-31.
46. Leonard DT. Translating and communicating patient centered research to the public. Am J Pharm Benefits 2010;2: 408-410.
47. McGrergor M. What decision-makers want and what they have been getting. Value in Health 2006;9: 181-184.
48. CADTH. Guidelines for the economic evaluation of health technologies: Canada [3rd Edition]. Ottawa: Canadian Agency for Drugs and Technologies in Health; 2006.
49. Husereau D, Drummond M, Petrou S, Carswell C, Moher D, Greenberg D, et al. Consolidated Health Economic Evaluation Reporting Standards (CHEERS) statement. BMJ 2013; 346: f1049.
50. Merlin T, Farah C, Schubert C, Mitchell A, Hiller JE, Ryan P. Assessing personalized medicines in Australia: a national framework for reviewing codependent technologies. Med Decis Making 2013; 33: 333-342.
51. Clarke PM, Fitzgerald EM. Expiry of patent protection on statins: effects on pharmaceutical expenditure in Australia. Med J Aust 2010;192: 633-636.

52. Morgan S, Thomson P, Daw J, Friesen M. Inter-jurisdictional cooperation on pharmaceutical product listing agreements: views from Canadian provinces. BMC Health Services Research 2013; 13:34
53. Straus S, Tetroe J & Graham ID (Eds.), Knowledge translation in health care: Moving from evidence to practice, 2nd edition. Oxford: Blackwell Publishing, 2013.
54. Chambers D, Wilson P. A framework for production of systematic review based briefings to support evidence-informed decision-making. Syst Rev 2012; 1: 32.
55. Akl EA, Guyatt GH, Irani J, Feldstein D, Wasi P, Shaw E, Shaneyfelt T, Levine M, Schünemann HJ. "Might" or "suggest"? No wording approach was clearly superior in conveying the strength of recommendation. J Clin Epidemiol 2012; 65: 268-275.
56. Trevena L, Davey H, Barratt A, Butow P, Caldwell P. A systematic review on communicating with patients about evidence. J Evalu Clin Practice 2006;12: 13-23

TABLES AND FIGURES

Table 1: Characteristics of Included Studies.

STUDY CHARACTERISTICS			
Author and year	**Study Design**	**Number of Participants**	**Location**
POLICY MAKERS			
Ammerman 2009*	Survey (email)	294	USA
Anell 2000	Survey (mail)	210	Sweden
Baghbanian 2011	Survey (email)	91	Australia
Chaikledkaw 2009**	Survey (mail)	626	Thailand
Chen 2007	Interviews	10	England
Davies 1994 ***	Survey (not reported)	8	Europe
Drummond 2003	Focus group, Literature Review, Meeting Observations	10 participants 16 studies	USA, UK
Duthie 1999	Interviews (semi-structured)	34	UK
Eddama 2008	Interviews, Meeting Observations, Document Analysis	20	UK
Gunther 2006	Survey (mail)	77	Germany

STUDY CHARACTERISTICS			
Author and year	**Study Design**	**Number of Participants**	**Location**
Hoffman 2000	Survey, Interviews and/or Focus Group	1041	Europe
Hoffman 2002	Focus Group	12	UK
Lafi 2012 ****	Interviews	7	Jordan
Neumann 2008	Interviews (semi-structured)	46	USA
Niessen 2012	Systematic Review	43 studies	UK, Australia, North America
Phillips 2010	Focus Group	8	Wales
Rocchi 2008	Focus Group, Literature Review, Document Analysis	13 participants 13 studies	Canada
Ross 1995	Interviews (structured)	34	Australia
Taylor Robinson 2008	Interviews (semi-structured)	33	UK
Thurston 2008	Survey	50	Europe
van Dongen 2013	Interviews (in-depth, structured)	47	Canada
Williams 2007	Interviews, Document Analysis, Meeting Observation	31	England
Williams 2008	Interviews, Systematic Review, Workshops	28 participants, 1 study	UK

STUDY CHARACTERISTICS			
Author and year	**Study Design**	**Number of Participants**	**Location**
CLINICIANS			
Fattore 2006	Survey (mail)	175	Italy
Jansson 2006	Survey (mail)	738	Sweden
Prosser 2005	Interviews (semi-structured)	45	UK
Torbica 2010	Survey (self-administered)	129	Italy
Walley 1997	Survey (mail)	178	Great Britain
PATIENTS OR PUBLIC			
Peters 2007	Randomized Controlled Trial	303	USA
HETEROGENEOUS POPULATION (POLICY MAKERS AND PRACTICING CLINICIANS)			
Erntoft 2011	Systematic Review	31 studies	UK, North America
Galani 2008	Systematic Review	55 studies	North America, Europe
Rubenstein 2007	Interviews (semi-structured), Focus Groups	20	Argentina
Zwart van Rijkom 2000	Interviews, Survey	34	Netherlands
Abbreviations: UK=United Kingdom; USA=United States of America			

*Ammerman 2009 study included researchers but no differences by population were found upon analysis and results were reported in aggregate

** Chaikledkaw 2009 also included 132 researchers; these results were reported separately and not included in this review (total N=758)

*** Davies 1994 surveyed 8 health services researchers who reported on the impact of economic evaluations on decision-makers and policy-makers based on a review of 66 studies in the literature

**** Lafi 2012 also included 3 health economists; these results were reported separately and not included in this review (total N=10)

Table 2. Themes Related to Communication and Presentation of Health Economic Information

Themes Related to Communication and Presentation of Economic Evaluations
• Graphics would be useful but current visuals used by health economists are still too complicated to understand, e.g. cost effectiveness acceptability curve, cost-effectiveness plane • Present results in disaggregate forms as much as possible so the user can deconstruct the information and determine applicability to their context • Level of detail required depends on the end-user and how they will use the information • 'Less is more' and most important information should be presented first • Formatting approaches that allow for rapid scanning are helpful (e.g. headings, separate sections) • Key details such as objectives, assumptions and practical implications should be clear • Technical language and jargon is difficult to understand • QALY has limitations but nothing better exists and allows comparison with other drugs or technologies • ICERs not well-understood but still useful because they can be can be compared against external benchmarks, i.e. thresholds • Different views on the appropriateness of using a cost-effectiveness threshold

Table 3. Themes Related to Knowledge of Health Economic Information

Themes Related to Knowledge of Health Economic Information
• Some training in health economics usually exists but extent and type of training was variable across studies (e.g. workshops, short courses, distance packages) • Practical experience with economic evaluations is limited but strong interest to learn more about health economics • Focus is on high-level concepts, e.g. types of economic evaluations, ICER, QALY • Although there is general familiarity with health economic concepts, difficulties in applying concepts to real-world decisions • Limited understanding of economic modeling and how different factors impact the economic model • Numeracy skills in the general population are low • No clear differences in knowledge between policy makers and clinicians but attitudes towards economic evaluation may be less favourable among clinicians

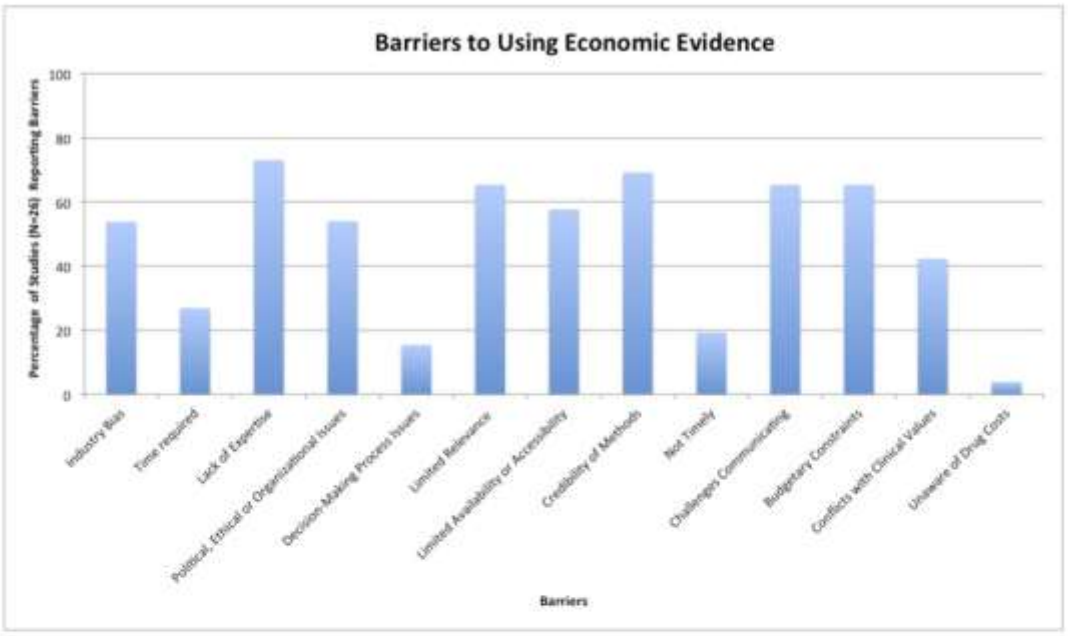
Table 4. Specific Barriers and Facilitators Related to Communicating Health Economic Information as Reported in Individual Studies

Barriers	**Facilitators**
• Graphical presentation of models lacking or needs improved format • Lack of clarity in presentation of economic information • Are better ways to present information than those currently used (e.g. QALY, ICER) • Economists focus more on methodological rigor than communicating the principles involved • Concerns re: how results disseminated • Everyday language takes on different meaning in health economics, which can cause confusion • Technical language or jargon not understood • Oversimplification of complexity • Not able to communicate results to public/policymakers to help them understand value • Cannot explain when queried on results • Complicated and hard to understand	• Reporting of results is easy to understand • Clarity and good communication of economic evaluation • Assumptions in model clearer • Improved communication by economists • Clear standardized presentation format • Providing disaggregate list of costs and benefits • Explaining practical relevance of results (e.g. actual cost savings) • Dissemination by media • If more decision makers aware of results

FIGURE CAPTIONS

Figure 1. Most Frequently Reported Types of Barriers and Facilitators to Using Economic Information.

Supplementary Figure 1. Results of the Literature Search.



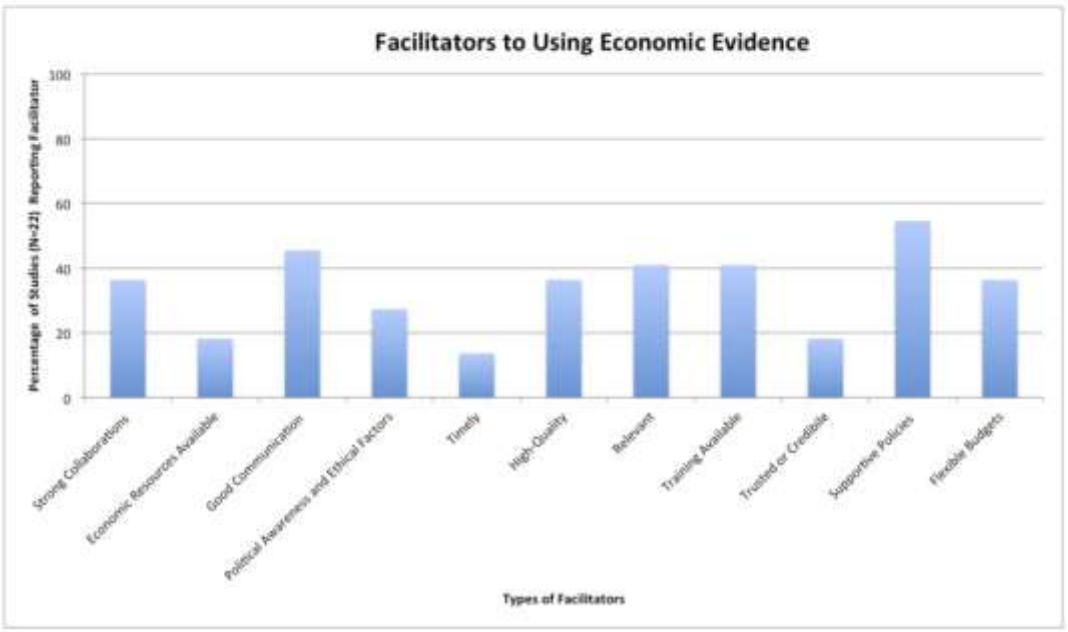
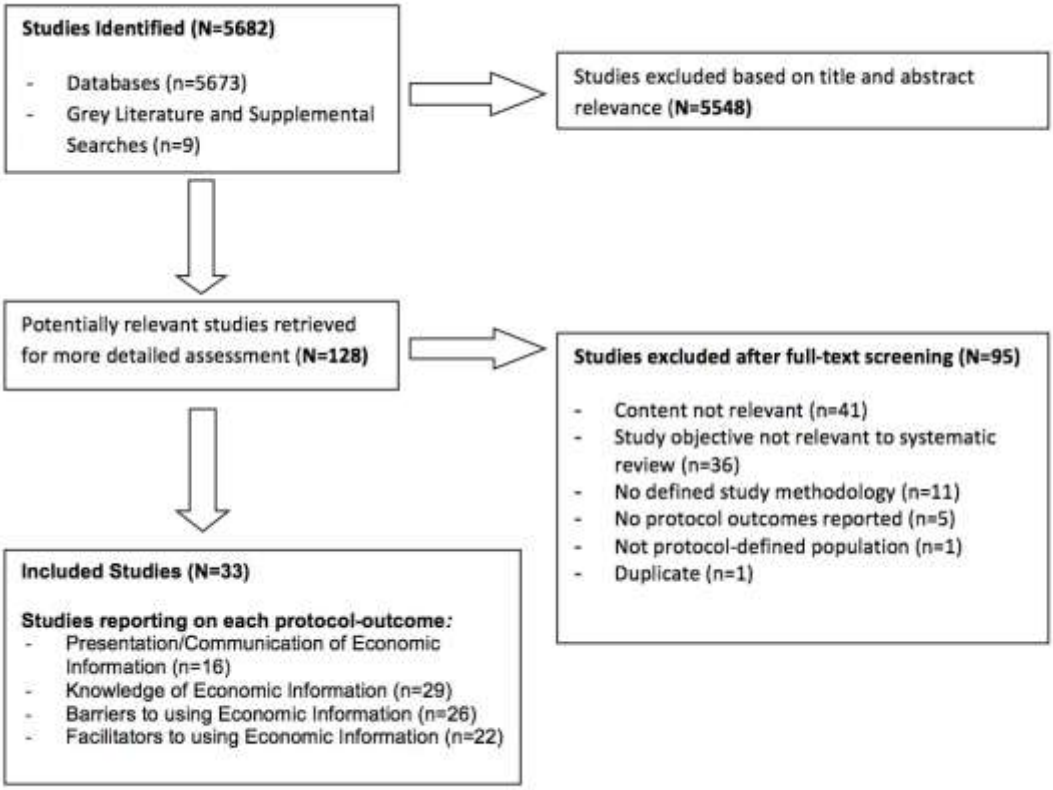


Figure 1. Most Frequently Reported Types of Barriers and Facilitators to Using Economic Information.

SUPPLEMENTAL MATERIALS



Supplementary Figure 1. Results of the Literature Search.

Supplementary Table 1. Detailed Study Characteristics

Author and year	Study Objective	Type of Economic Information	How Economic Information Used
POLICY MAKERS			
Ammerman 2009	To assess current use of health economic tools by public health researchers and practitioners, to identify barriers to their use and potential solutions to overcoming these barriers	Economic Evaluation CEA, CBA ICER	• Determine cost of public health problems • Evaluate economic incentives • Determine treatment cost • Return on investment
Anell 2000	To identify obstacles to the increased use of pharmacoeconomic evaluations in decision-making To identify the information sources and decision criteria that individual committee members perceive as important in establishing clinical practice guidelines.	Economic Evaluation	• Resource allocation decisions • Formulary committee decision-making
Baghbanian 2011	To explore dimensions and varieties of economic evaluations that healthcare decision-makers do or do not use.	Economic Evaluation CEA, CUA, CBA	• Resource allocation decisions
Chaikledkaw 2009	To explore the knowledge, experience, and attitudes towards economic evaluation among decision makers and researchers	Economic Evaluation	• Resource allocation decisions • Drug formulary decisions • Clinical practice guidelines development • Communicating with prescribers
Chen 2007	To explore the perceptions of pharmacists on the use of economic evaluation in formulary decision making at local hospitals To identify factors which influence the uptake of economic evidence and examine the usefulness of different presentations of economic evidence.	Economic Evaluation	• Influence use of new medicines in hospitals • Formulary meetings (desired use)
Davies 1994	To examine the methodology used and to identify the impact the results of economic evaluations have had on decision and policy making in health care based on respondents identification and review of economic evaluations	Economic Evaluation Cost analysis, CEA, CUA, CBA	• Influence health care planning and policy development • Introduce cost effective practices • Promote commercial interests such as drug reimbursement
Drummond 2003	To use the ISPOR task force in order to develop a “toolbox for the health care decision maker wanting to interpret and use health economic studies.	Economic Evaluation Cost effectiveness	• Resource allocation decisions • Formulary decision-making
Duthie 1999	to determine the relevance and appeal of diverse health economic measures to different decision makers.	Economic Evaluation	• Resource allocation decisions
Eddama 2008	To explore decision making and the use of economic evaluation at the local health care decision making level	Economic Evaluation	• To make resource allocation and budget decisions about drugs and other health technologies

Author and year	Study Objective	Type of Economic Information	How Economic Information Used
Gunther 2006	To understand the opinions held by decision makers and scientists regarding the methodology used to assess willingness to pay	Willingness to Pay	• Resource allocation decisions
Hoffman 2000	To report the findings of the EUROMET survey of decision-makers knowledge of economic evaluations, use of study results and barriers and incentives in the use of economic studies.	Economic Evaluation	• Inform healthcare decision-making on health care technologies
Hoffman 2002	To explore if economic evaluations are widely used by healthcare professionals in the UK NHS and if greater access to results of economic evaluations would be beneficial	Economic Evaluation Cost-effectiveness	• Cost effectiveness to inform health care decision making
Lafi 2012	To explore the extent of and barriers to the use of economic evaluation in compiling the Jordan Rational Drug List in the health care system of Jordan.	Economic Evaluation	• Resource allocation decisions • Drug selection for national formularies
Neumann 2008	To investigate ways of defining and measuring the value of services provided by government public health systems.	Economic Evaluation CBA, CEA	• To value public health interventions
Niessen 2012	To assess whether and how economic evidence has an impact on health care decision-making in the United States and in other countries and to identify antecedents or obstacles for use in health policy.	Economic evaluation CEA, CUA Budget impact Cost information	• Resource allocation decisions: formulary reimbursement and coverage of health programs and services • Priority setting in program development • Decisions related to practice guidelines and clinical management of patients • General health policymaking decisions
Phillips 2010	To explore the feasibility and validity of using health economic evaluation techniques in developing priorities within public health programs and consider the extent to which different presentational approaches are likely to be incorporated into decision making, from the perspective of relevant stakeholders based on sharing scenario analyses with a mock advisory board	Economic Evaluation Cost-consequence, CBA, CUA ICER, QALY	• To inform funding decisions • Priority setting for public health programs • Compare with other thresholds and benchmarks
Rocchi 2008	To identify the impact of economic evidence on Canadian drug reimbursement decisions, to determine if there is an economic threshold for decision-making and if the impact or threshold is different for oncology drugs	Economic Evaluation	• Resource allocation decisions • Drug reimbursement and formulary considerations • Identifying populations that benefit most

Author and year	Study Objective	Type of Economic Information	How Economic Information Used
Ross 1995	To identify how economic evaluation is being used by key health care decision makers in different areas of Government; To identify barriers to use of economic evaluation perceived by these decision makers; To determine what, if anything, could be done to enhance the use of economic evaluation in health care policy making.	Economic Evaluation	• Resource allocation decisions (e.g. screening or medical devices, hospital services or programs) • Inform pricing of new drugs,
Taylor-Robinson 2008	To explore attitudes towards the use of models for coronary heart disease to support decision making for policy and service planning.	Economic Modeling	• Economic models used to inform cost-effectiveness • Resource allocation decisions • Predict and plan future needs, service delivery, costs • Convince politicians about value for money
Thurston 2008	To seek decision makers views on different summary formats of economic evaluations including a score, short summary, and structured abstracts of different degrees of detail.	Economic Evaluation	• Users of NHS EED database • Determine cost effectiveness of interventions
van Dongen 2013	To provide insight into the occupational health decision-making process and information needs of decision-makers as a first step in bridging the gap between the economic evaluation literature and daily practice in occupational health	Economic Evaluation	• Occupational health decisions on resource use • Inform business cases for occupational health
Williams 2007	To explore the use of economic evaluations in technology coverage decisions in England.	Economic Evaluation	• Resource allocation decisions for drugs and technologies
Williams 2008	To determine the extent to which health economic information is used in health policy decision-making in the UK, and to consider factors associated with the utilisation of such research findings	Economic Evaluation Budget impact	• Resource allocation decisions • Decision-making framework
CLINICIANS			
Fattore 2006	To investigate the opinion and attitude of health professionals with a solid background in health care management/ economics towards the use of economic evaluation at the policy, organizational, and professional levels of decision making.	Economic Evaluation	• Resource allocation of drugs, preventive services and diagnostic services • National level and guideline development • Organizational, managerial and clinical decisions • Useful at a professional level and as a cultural framework

Author and year	Study Objective	Type of Economic Information	How Economic Information Used
Jansson 2006	To investigate whether decentralised drug-budgets according to a so-called primary-care based model in Swedish health care have led to increased cost awareness and changed attitudes towards cost-minimisation and cost-effectiveness as decision-making criteria among physicians	Economic Evaluation	• Cost-effective prescribing among physicians
Prosser 2005	To explore attitudes of various stakeholders on the relative importance and influence of cost on general practice prescribing.	Drug Costs	• General GP prescribing
Torbica 2010	To evaluate the impact of cost effectiveness information on clinical decision making in cardiology through a survey using discrete choice experiment methodology	ICER	• Resource allocation decisions
Walley 1997	To assesses attitudes of "clinical" advisors to economic evaluations of drug therapy and their perceptions of barriers to achieving cost-effective prescribing by use of these evaluations	Economic Evaluation Cost-effectiveness	• Discussions on prescribing
PATIENTS OR PUBLIC			
Peters 2007	To examine alternative approaches for displaying and conveying comparative hospital performance information (quality and cost indicators) that may improve the understanding of those with lower numeracy skill. To explore the importance of numeracy skill in understanding and using health information to make hospital choices.	Cost information	• Inform hospital performance quality indicators • Choice of hospital and quality health care
HETEROGENEOUS POPULATIONS (POLICY MAKERS AND PRACTICING CLINICIANS)			
Erntoft 2011	To investigate which factors and criteria are used in priority setting of pharmaceuticals, in what contexts health economic evaluations are used, and barriers to the use of health economic evaluations at micro, meso, and macro health-care levels	Economic Evaluation	• Pharmaceutical priority setting • Resource allocation decisions • Formulary and reimbursement decisions • To help manage potentially large budget impact
Galani 2008	To summarize self-reported attitudes of healthcare decision-makers towards economic evaluations, to determine the extent to which they are used in health policy decision-making and to consider factors associated with their use. Central, local and physician level decision-makers were included.	Economic evaluations Cost consequences, CEA, CMA, CBA, CUA	• Health-care policy and resource allocation decisions (pharmaceutical policy, other healthcare intervention policies) • Clinical-decision making on choice of treatment for patients

Author and year	Study Objective	Type of Economic Information	How Economic Information Used
Rubinstein 2007	To explore micro, meso and macro-level health decision-makers' knowledge and attitudes about economic evaluations and whether health technology assessments were increasingly used for decision-making. This included clinicians and policy makers (managers and public health officer)	Economic Evaluation	• Resource allocation decisions
Zwart van Rijkom 2000	To investigate differences in attitudes, knowledge and actual use of economic evaluations in different groups of decision makers and to compare the results from the Netherlands with the overall European results, as part of the EUROMET project. This included physician opinion leaders, hospital pharmacists, politicians, and regulators.	Economic Evaluation	• Resource allocation decisions

CBA=cost-benefit analysis; CEA=cost-effectiveness analysis; CMA=cost-minimization analysis; CUA=cost-utility analysis; EUROMET=European Network on Methodology and Application of Economic Evaluation Techniques; ICER=incremental cost-effectiveness ratio; NHS EED=National Health Services Economic Evaluation Database; QALY=quality-adjusted life year

Supplemental Table 2A. Quality assessment of included systematic reviews based on AMSTAR criteria.

SYSTEMATIC REVIEWS	Research Protocol	Search and Inclusion	Data Extraction	Included Study Details	Analysis and Conclusions	Publication Bias	COI	Other Sources and Notes
Erntoft 2011		x		x	x	x	x	Search for unpublished reports limited; excluded studies not identified (n=60); no critical appraisal and impact of study quality on conclusions not assessed; publication bias not assessed; COI not reported
Galani 2008		x	x	x	x	x	x	No supplemental search strategy or grey literature search; excluded studies not identified (n=338); no duplicate data extraction; no critical appraisal and impact of study quality on conclusions not assessed; publication bias not assessed; COI not reported
Niessen 2012		x				x	x	No supplemental search strategy or grey literature search; publication bias not assessed; COI not reported for individual studies
Abbreviations: COI=conflict of interest								

'x' – indicates an issue was identified during critical appraisal of the study

Supplemental Table 2B. Quality assessment of included randomized controlled trials (RCTs) based on Cochrane Risk of Bias Tool

RCT	Sequence Generation	Allocation Concealment	Blinding	Incomplete Outcome Data	Selective Outcome Reporting	Other Sources and Notes
Peters 2007	x	x	x	x	x	No details on sequence generation or allocation concealment; blinding not reported, e.g. outcome assessors and study administrators to numeracy skills; completeness of data or methods of handling missing data not reported; potential for selective outcome reporting not assessed; convenience sample; funding sources and conflict of interest not reported

'x' – indicates an issue was identified during critical appraisal of the study

Supplemental Table 2C. Quality assessment of qualitative or mixed methods studies using checklists from Greenhalgh et al. 2004

QUALITATIVE STUDIES	Research Question/ Design	Sampling/ Response	Data Collection/ Instrument	Analysis	Results/ Interpretation	Other Sources and Notes
SURVEYS						
Ammerman 2009		x	x	x	x	**Strengths:** Large sample size; survey developed based on focus group data; closed- and open-ended questions; broad distribution through opinion leaders in community and research networks **Limitations:** Piloting, validity or reliability of survey not reported; responder bias; no information on non-response due to recruitment methods; self-report bias; analysis methods unclear; potential for researcher bias on interpretation unclear
Anell 2000		x	x			**Strengths:** Sample identified through national register; large sample size; 69% response rate; survey piloted; appropriate statistical analysis (closed-ended questions) **Limitations:** Responder bias; self-report bias
Baghbanian 2011		x	x			**Strengths:** Purposive sampling; validated questionnaire (closed-ended questions) **Limitations:** 24% response rate (26% expected); responder bias; self-report bias
Chailkledkaew 2009		x	x			**Strengths:** Purposive and random sampling; large sample size; appropriate statistical analysis (closed-ended questions) **Limitations:** Development, piloting, validity or reliability of survey not reported; low response rate (29%); self-report bias;
Davies 1994		x	x	x	x	**Strengths:** Purposive sample; 80% response rate; standardized questionnaire **Limitations:** Small population targeted; questionnaire development methods not reported; self-report bias; analysis methods unclear; potential for researcher bias on interpretation unclear

QUALITATIVE STUDIES	Research Question/ Design	Sampling/ Response	Data Collection/ Instrument	Analysis	Results/ Interpretation	Other Sources and Notes
Fattore 2006		x	x			**Strengths:** Large sample size; 47% response rate; sample identified through institutional database; appropriate statistical analysis (closed-ended questions) **Limitations:** Development, piloting, validity or reliability of survey not reported; convenience sample; responder bias; self-report bias
Gunther 2006		x	x			**Strengths:** Sample identified through government database; large sample size; survey development by expert group; 62% response rate; appropriate statistical analysis (closed-ended questions) **Limitations:** Responder bias; self-report bias
Jansson 2006		x	x			**Strengths:** Sample identified through government database; large sample size; survey piloted; 49% response rate; appropriate statistical analysis (closed-ended questions) **Limitations:** Responder bias; self-report bias
Thurston 2008	x	x	x	x	x	**Strengths:** Large sampling frame; sample identified through databases and user lists **Limitations:** Study objective better suited to other designs (e.g. randomized comparisons); development, piloting, validity or reliability of survey not reported; responder bias; low response rate (3.5%); self-report bias; analysis methods unclear; potential for researcher bias on interpretation unclear

QUALITATIVE STUDIES	Research Question/ Design	Sampling/ Response	Data Collection/ Instrument	Analysis	Results/ Interpretation	Other Sources and Notes
Torbica 2010		x	x			**Strengths:** Survey developed based on focus group data and piloted; 61% response rate; objective measure/analysis of using economic information to inform decisions; appropriate statistical analysis (closed-ended questions) **Limitations:** Convenience sample (conference attendees); responder bias; self-report bias
Walley 1997			x			**Strengths:** Survey developed based on focus group data; sample identified through database; 65% response rate; awareness of economic information tested objectively and answers validated (e.g. have you read this article); appropriate statistical analysis (closed-ended questions) **Limitations:** Self-report bias
INTERVIEWS OR FOCUS GROUPS						
Chen 2007					x	**Strengths:** purposive sample; development interview guide described; data saturation achieved; themes analyzed for consensus; analysis themes linked to economic framework/theory **Limitations:** researcher biases on interpretations not considered
Duthie 1999		x		x	x	**Strengths:** interview guide piloted; participants interviewed in pairs to promote quality discussion; consensus on analysis themes reached among researchers **Limitations:** industry funded; unclear how sample identified; analysis methods not well-described; researcher biases on interpretations not considered; ethics approval not reported

QUALITATIVE STUDIES	Research Question/ Design	Sampling/ Response	Data Collection/ Instrument	Analysis	Results/ Interpretation	Other Sources and Notes
Hoffman 2002	x	x	x	x	x	**Strengths:** key informant sample **Limitations:** Convenience sample; focus group guides not well-described; analysis methods not well-described; few illustrative quotes; researcher biases on data collection or interpretations not considered; ethics approval not reported
Lafi 2012		x			x	**Strengths:** Data collection and analysis well-described; illustrative quotes provided; interpretations linked to results and theory **Limitations:** Key informant sample (29% participation rate); inadequate sample size for conceptual generalizations; researcher biases on data collection or interpretations not considered; ethics approval not reported
Neumann 2008			x	x	x	**Strengths:** Context supported by literature review; key informant sample (100% participation) **Limitations:** Development or piloting interview guides not reported; analysis methods not well-described; few illustrative quotes; researcher biases on data collection or interpretations not considered; ethics approval not reported
Phillips 2010	x	x	x	x	x	**Strengths:** Purposive sample **Limitations:** Study design not clearly defined but study objective better suited to other designs; Inadequate sample size for conceptual generalizations; data collection and analysis methods not well-described; exploratory process and focus group guides not well-described; few illustrative quotes; researcher biases on data collection or interpretations not considered; ethics approval not reported

QUALITATIVE STUDIES	Research Question/ Design	Sampling/ Response	Data Collection/ Instrument	Analysis	Results/ Interpretation	Other Sources and Notes
Prosser 2005					x	**Strengths:** Development interview guides described and piloted; both focus groups and interviews conducted; purposive sampling; grounded-theory based approach; data saturation achieved; illustrative quotes provided **Limitations:** Researcher biases on interpretation not considered; ethics approval not reported
Ross 1995			x	x	x	**Strengths:** Purposive sampling; adequate sample size **Limitations**: Development or piloting interview guides not reported; analysis methods not well-described; few illustrative quotes; researcher biases on data collection or interpretations not considered
Rubenstein 2007			x		x	**Strengths:** Both focus groups and interviews conducted; key informant sample; adequate sample size; grounded-theory based approach; analysis methods well-described; illustrative quotes provided **Limitations:** Development or piloting interview guides not reported; researcher biases on data collection or interpretations not considered
Taylor Robinson 2008			x			**Strengths:** Purposive sampling; adequate sample size; 'framework' theory approach; independent coding and verification by third researcher for inter-researcher consistency; analysis methods well-described; illustrative quotes provided **Limitations:** Development or piloting interview guides not reported

QUALITATIVE STUDIES	Research Question/ Design	Sampling/ Response	Data Collection/ Instrument	Analysis	Results/ Interpretation	Other Sources and Notes
Van Dongen 2013			x			**Strengths:** Purposive sampling; interview protocol used; data saturation achieved; analysis methods well-described; illustrative quotes provided; data interpretations verified with interviewers to enhance robustness **Limitations:** Development or piloting interview guides not reported
MIXED METHODS (QUANTITATIVE AND QUALITATIVE)						
Drummond 2003		x	x	x	x	**Strengths:** triangulation of multiple data sources **Limitations:** literature review not well-reported or independently conducted - based on subset of studies from a recent systematic review; focus group methodology not reported; biases of researchers on interpretations not considered; industry funded; ethics approval not reported
Eddama 2008			x		x	**Strengths:** purposive interview sample; sufficient sample size for conceptual generalizations; analysis methods well-described; analysis based on constant comparative approach; negative, or disconfirming cases, were explored; illustrative examples used; triangulation of interview data with document analysis and meeting observations; grant-funded **Limitations:** Development or piloting of interview guides not described; researcher biases on interpretations not considered

QUALITATIVE STUDIES	Research Question/ Design	Sampling/ Response	Data Collection/ Instrument	Analysis	Results/ Interpretation	Other Sources and Notes
Hoffman 2000		x		x	x	**Strengths:** triangulation of data from surveys, interviews and focus groups; validated questionnaire adapted for all methods; sufficient sample size for conceptual generalizations; grant-funded **Limitations:** variable response rates (12%-100%); high-level themes reported but not detailed results or illustrative examples- refers to additional publication for detailed methods and results; researcher biases not considered
Rocchi 2008		x	x	x	x	**Strengths:** key informant sample; sufficient sample size for conceptual generalizations; triangulation of focus group data with document analysis and literature review **Limitations:** literature review not well-reported; focus group methodology not reported; analysis methods not well-described; few illustrative examples; researcher biases not considered; industry funded; ethics approval not reported
Williams 2007			x		x	**Strengths:** Purposive sample; sufficient sample size for conceptual generalizations; analysis methods described; triangulation of interview data with document analysis and meeting observations **Limitations:** Development or piloting of interview guides not described; biases of researchers on interpretations not considered; funding sources and ethics approval not reported

QUALITATIVE STUDIES	Research Question/ Design	Sampling/ Response	Data Collection/ Instrument	Analysis	Results/ Interpretation	Other Sources and Notes
Williams 2008			x		x	**Strengths:** Purposive sample; systematic review well-reported; analysis methods described; analysis based on grounded theory; illustrative examples provided; triangulation of multiple data sources; independent institutional funding **Limitations:** Development or piloting of interview guides not described; biases of researchers on interpretations not considered;
Zwart-van-Rijkom 2000				x	x	**Strengths:** Key informant sample; sufficient sample size for conceptual generalizations; triangulation of data from surveys, interviews and focus groups; validated questionnaire adapted for all methods; grant-funded **Limitations:** analysis methods not well-described; few illustrative examples provided; biases of researchers on interpretations not considered; ethics approval not reported

'x' – indicates an issue was identified during critical appraisal of the study

Supplemental Table 3: Included studies highlighting preferences for the communication and presentation of health economic information among different target audiences.

Study	Communication and Presentation of EE
POLICY MAKERS	
Ammerman 2009	NR
Anell 2000	NR
Baghbanian 2011	NR
Chaikledkaew 2009	NR
Chen 2007	• Disaggregate estimates with ranges most favoured (more than disaggregate without range) • ICER too difficult to understand • Like idea of visuals but CE plane and CEAC hard to understand even when explained • CE plane impressive but difficult to understand and could be deceiving • CEAC most difficult to understand, especially when don't know WTP threshold
Davies 1994	NR
Drummond 2003	• Outcome measurements and objective of study should be clearly stated • Description of relevant patient populations and size at local level • Identify relevant budgets and impact on each of adopting new therapy • Disaggregated presentation of costs and outcomes with ICER • Results by relevant subgroup • Practical implications of adopting a new therapy • Listing of key assumptions and data sources for transparency • Sensitivity analyses presented using local data • Assumptions identified to assist transferability
Duthie 1999	• Per quality adjusted life year mistrusted • Cost benefit ratio meaningless • Mixed feelings re: willingness to pay • NNT very interesting/relevant

Eddama 2009	• Present like an impact assessment:
	  ○ Look at costs
	  ○ Look at how to assess benefits (i.e. not utilities), e.g. access, distress, capacity
	  ○ Create potential scoring criteria for benefits

Gunther 2006	NR
Hoffman 2000	NR

Study	Communication and Presentation of EE
Hoffman 2002	• If study summaries too brief, can't base decisions upon • Satisfied with readability of abstracts • Problems understanding economic terminology • Commentary and conclusions should be at the beginning of abstract so findings/limitations clear at first glance • Want more explicit critiques of original paper (e.g. quality of clinical data) • Details on study interventions too brief • Want abstracts with clear headings • Want more distinct separate sections • Interactive interface offering links to relevant sites (e.g. definitions, glossaries) • Include quality scoring system
Lafi 2012	• Confusion between EE and concepts like cost analysis • Better understand simple language like "what's the add-on of this drug that can justify the cost difference"
Neumann 2008	• CBA, CEA potential ways to measure value and present • Want to demonstrate 'return on investment' not just lives saved
Niessen 2012	NR

Phillips 2010	• Scenarios with more information preferred
	• Preferred comparing the results of the analysis against an external benchmark (cost per QALY against threshold by NICE), i.e. provides an external definition of 'good'
	• CBA or CUA preferred over CEA or CCA (costs and effects not comparable in these analyses) despite added complexity and less transparency
	• Details required for CCA included:
	  ○ if setting representative
	  ○ more details on intervention for quality assessment and implementation
	  ○ baseline level of condition to determine if improvements in use could produce health gains
	  ○ value of effects of intervention so they could be directly compared against cost of intervention

Rocchi 2008	• To focus on a single outcome and a single measure (ICER) is detrimental for reimbursement decisions – aggregate outcomes a "necessary evil" • Simplistic use of threshold value rejected – use of only the ICER is dangerous • QALY not ideal but necessary to have common denominator (allows comparison to other drugs)
Ross 1995	• Everyday language takes on specific economic meaning, which can lead to misunderstanding • Need more emphasis on principles involved than rigour of methods • More user friendly
Taylor-Robinson 2008	• Provides confidence limits • Provides realistic time horizons • Gives general direction of results rather than precise numbers • Simple enough to understand - explaining complexity or jargon can create doubt, e.g. Markov model) • Able to break down the results and understand where it was coming from/how it was derived

Study	Communication and Presentation of EE
Thurston 2008	• Very short summary plus a more detailed structured abstract preferred over summary sentences and score. Short abstract preferred over long abstract (both with score/summary) • No extra benefit of long abstract • Short format population b/c allows to find quickly the info relevant to them without losing essential detail • Similar results, regardless of specific use of econ info • Long abstract may be too complex for majority of decision makers but health economic decision-makers preferred detail offered in long abstract. Short abstract lacked detail on costs used in studies and would benefit from a more comprehensive analysis of costs and benefits • Quality score useful as a screening tool (82%) but not enough info for decision-makers on content of study • Quality score with summary useful by 92%
Van Dongen 2013	NR
Walley 1997	NR
Williams 2007	NR
Williams 2008	• QALYs and cost per event – there are better ways to present CE info • Clearer breakdown re: which budgets get savings • Address implementation issues, e.g. increase in demand when approved • Consistent format would assist understanding • Use of consistent methodologies • Sensitivity analyses very useful but need to be more accessible • User-friendly overviews that do more than just repeat what is in reports • Should address specific questions committee members likely to have • Graphical representations of model helpful • Limitations of QALY
CLINICIANS	
Fattore 2006	NR
Jansson 2006	NR
Prosser 2005	NR
Torbica 2010	NR
PATIENTS AND PUBLIC	
Peters 2007	• Delete less important information • Order information from more to less important • Highlight the meaning of important information • Using symbols (rather than numbers) can help but must be chosen carefully and tested in population to not increase cognitive burden • Presentation format matters less in individuals with high numeracy skills but more in individuals with low numeracy skills • Reporting metrics where higher scores represent better performance and lower scores represent poor performance (vs. higher scores representing poor performance, etc.)
MIXED POPULATIONS	

Study	Communication and Presentation of EE
Erntoft 2011	• Preference for disaggregated measures instead of QALYs • Need graphic presentations of models • Methodological concepts challenging to understand: uncertainty, consistency, assessment of benefits in very severe conditions, irreversibility
Galani 2008	NR
Rubenstein 2007	NR
Zwart van Rijkom 2000	• Majority disagreed with use of an explicit cut-off point for cost-effectiveness o large uncertainty around ICER o doubts on value of QALY o discriminates older/disabled o for individuals a small health gain could mean a lot (e.g. musician's hand) o can't assess care-giving, only cure o ahistorical approach – no evidence of used previously • But pharmacists generally agreed with use of threshold– thought it would promote objective/transparent decision-making • Politicians very resistant, clinicians varied • Pharmacists preferred CMA, CBA, CEA. CUA least popular.

Supplemental Table 4. Included studies reporting outcomes related to knowledge and understanding of health economics among different target audiences.

Study*	Attitudes or Interest in Learning about Health Economics	Experience or Actual Use of Health Economics	Training in Health Economics	Knowledge of Health Economics (Self-Report or Measured)
POLICYMAKERS				
Ammerman 2009	• 88% interested or very interested to learn • 66% interested in learning more about CBA, CE, determining the costs of public health problems • >50% interested in learning about resources and specific examples of applying HE methods	• >50% use very little or not at all in current work	NR	NR
Anell 2000	• 81% disagree there is no interest in HE	NR	• 52% had some HE training • 31% had 1-3 day seminar • 13% had >3 day program • 8% other form of training	NR
Baghbanian 2011	• 82.5% think only marginally or to some extent influences resource allocation • 65.9% think HE should or will play larger role in future	NR	• 50% had some HE training (but thought could be useful) • 81% no plan for or unsure about future HE training (b/c already some training, no time or no interest)	• 86% moderately to very familiar with CEA (less with CBA and CUA)

Study*	Attitudes or Interest in Learning about Health Economics	Experience or Actual Use of Health Economics	Training in Health Economics	Knowledge of Health Economics (Self-Report or Measured)
Chaikledkaew 2009	• 80% interested in short-course and on-the-job HE training		• 71% no HE training	• >70% not able to use or don't have confidence using HE • Challenging concepts included: unit cost, direct cost, indirect cost, CBA, CUA, marginal cost, CMA, QALY, discounting, ICER, sensitivity analysis
Chen 2007	• Interest to use at formulary meetings but not used currently	• 20% had ever used EE for decision-making	• 20% had HE training	NR
Davies 1994		• Cost utility analysis less frequently used but more likely to influence decision making	NR	NR
Drummond 2003	NR	NR	NR	NR
Duthie 1999	• Considered HE had value • Interested in HE	NR	NR	NR
Eddama 2009	NR	• Not used much locally (only at NICE) • Published EE not used	NR	• Don't understand balance b/w costs and effects – no reflection on if less effective and cheaper drugs were efficient use of resources (e.g. vs. more effective and more expensive drugs) • Don't understand what HE is, e.g. think only associated with money or requires epidemiologic expertise • Understand scarcity • Familiar with concept of 'value for money' but don't apply it, unclear how well it is understood
Gunther 2006	• ~50% intend to learn more about methods for measuring WTP	• Low level of practical experience with WTP	NR	• 45.5% had knowledge of methods for measuring WTP (scored ≥3 on a 5 point Likert scale where 1 is low and 5 is high level of knowledge)

Study*	Attitudes or Interest in Learning about Health Economics	Experience or Actual Use of Health Economics	Training in Health Economics	Knowledge of Health Economics (Self-Report or Measured)
Hoffman 2000	• Less willing to use CE info than EE • HE expertise considered desirable when hiring staff	• 33% used EE in decision making	• 33% staff with HE training	• Poor knowledge of CBA, CEA or CUA
Hoffman 2002	• Consensus that they are useful	• Most had previously used	NR	NR
Lafi 2012	• All had interest in HE training but no specifics on what's most helpful	NR	NR	• Varied in estimation of own knowledge. • If no training, considered EE easy to understand. • If some training, concerned about poor understanding of EE and level of expertise in EE.
Neumann 2008	• critical to measure value	NR	NR	NR
Niessen 2012	• 30 studies HE had an impact • 11 studies HE impact limited • 2 studies HE had no impact	NR	NR	NR
Phillips 2010	NR	NR	NR	• Unable to decide if intervention represented good use of public funds in CCA and CBA scenarios
Rocchi 2008	• majority consider HE contributes to decision-making • clinical evidence more important but HE is supportive	NR	NR	NR
Ross 1995	NR	• 38% had used in decision-making	NR	• High level of awareness of HE • 90% able to define economic evaluation
Taylor-Robinson 2008	NR	NR	NR	NR
Thurston 2008	NR	NR	NR	NR

Study*	Attitudes or Interest in Learning about Health Economics	Experience or Actual Use of Health Economics	Training in Health Economics	Knowledge of Health Economics (Self-Report or Measured)
Van Dongen 2013	• 79% interested in HE training • Specifically want more knowledge on CBA than CEA and CUA	NR	• 36% with HE training	• Confusion among differences in types of EE – thought to be synonymous • 93% familiar with CBA and 72% provided correct definition, • 71% familiar with CEA and 11% provided correct definition, • Most not familiar with CUA.
Walley 1997	• 91% HE important • 98% wanted to use HE more	NR	• 40% some HE training • Most had distance learning package	NR
Williams 2007	• Cost impact more valued than formal CEA	• formal CEA used less than other forms of cost info	• Some HE expertise on committees	• Confusion in HE terminology, e.g. economic evaluation vs. cost-effectiveness
Williams 2008	NR	• 19-23% seeking HE information in primary and secondary care	NR	• Variable knowledge across committee members • Have basic understanding of QALY, cost per QALY, threshold, CEAC, ICER • Knowledge of non-economists increased over time on committee • Appreciate what a QALY is but don't understand how modeled • Don't understand modeling, e.g. Markov model, Monte Carlo analysis, bootstrapping • Don't understand how factors in model influence QALY • Don't always understand the limitations/uncertainties in the analysis
CLINICIANS				
Fattore 2006	• Positive attitude to HE • Useful for professional activities (3.83 of 5) and as a cultural or logical framework (4.19 of 5)	NR	• All have some training in HE	• Perceived knowledge of health economics was 2.79 of 5; score similar between clinicians and managers

Study*	Attitudes or Interest in Learning about Health Economics	Experience or Actual Use of Health Economics	Training in Health Economics	Knowledge of Health Economics (Self-Report or Measured)
Jansson 2006	• Perceived responsibility to use HE higher in GPs than specialists	NR	• 13-16% had HE training	NR
Prosser 2005	• Most agree costs important • No interest in more formal HE approaches	• Appear to be conducting informal CE analyses on their own (e.g. step-up approaches in prescribing, anticipating long-term benefits that would justify high costs) • tend to be cost-minimization approaches	NR	NR
Torbica 2010	• Agreed HE should be used more, especially younger clinicians	• Currently only cost of drug used • 84.1% had read at least 1 HE study during past year, 29.0% read >3 studies • HE influences decision-making based on discrete choice experiment	NR	• Self-assessed knowledge of HE, 69.1% ≥ 3 on 5 point scale • Cost-effectiveness level more important than an extremely cost-effective profile implying a threshold is being applied in decision-making
PATIENTS or PUBLIC				
Peters 2007	NR	NR	NR	• % correct responses to each numeracy skill questions ranged from 11% to 92% across 14 questions
HETEROGENOUS POPULATIONS (POLICY MAKERS AND PRACTICING CLINICIANS)				

Study*	Attitudes or Interest in Learning about Health Economics	Experience or Actual Use of Health Economics	Training in Health Economics	Knowledge of Health Economics (Self-Report or Measured)
Erntoft 2011	• Physicians view clinical issues more important than economic • Pharmacists more interested in costs and budget impact than CE	NR	NR	NR
Galani 2008	NR	NR	NR	NR
Rubenstein 2007	• Clinicians don't see they have role in resource allocation • Policy makers think clinicians have a role in resource allocation	• Not used frequently b/c not well-known • HTA reports used sometimes	NR	• HE not well known • Don't know taxonomy of different analysis types • Difficult to appraise studies/ analyses • What's interpreted as EE are partial evaluations of cost based on utilization measures
Zwart van Rijkom 2000	NR	• Some had used HE • For most, HE not a formal requirement	• Some training in EE (self-taught, 1-day to 3-week course) • HE training standard among pharmacists	• Physicians didn't know differences b/w CEA and CUA

Abbreviations: CBA=cost benefit analysis; CE=cost-effective; CEA=cost-effectiveness analysis; CEAC=cost effectiveness acceptability curve; CMA=cost minimization analysis; CUA=cost-utility analysis; EE=economic evaluation; GP=general practitioner; HE=health economic; HTA=health technology assessment; ICER=incremental cost-effectiveness ratio; NR=not reported; QALY=quality adjusted life year

Supplementary Table 5. Specific Barriers and Facilitators Reported in Individual Studies and Grouped by Theme

BARRIERS	FACILITATORS
Industry Bias • Industry bias • Direct-to-consumer advertising • Promotion by pharmaceutical industry • Few peer-review publications **Time required** • Time and resources required to conduct • Limited time available **Lack of Training/Expertise** • No access to economists • Not enough trained economists • Lack of knowledge or experience with HE • Don't understand how models works • Dependent upon economists • Difficult to understand EE • Unable to critique EE • Not part of professional education and training • No self-confidence in using EE • Decision-makers lack knowledge of EE • Willingness to pay, CBA, CEA beyond analytic capacity **Political, Ethical or Organizational Barriers** • Political and ethical issues • No clear government policy on using EE • No formal requirement/legislation/policy to use EE • Lack of organizational support • No incentive to use • Political imperatives take precedence • No personal power to act on EE results • Political challenges of disinvestment decisions • EEs don't consider of opportunity costs • Political decision making process that doesn't use EE • No societal understanding that EE and rationing/prioritization are necessary • No political will to use EE • Patient and political pressure overturns decisions • Contradictory messages from government	**Economic Collaborations** • Decision-makers involved in EE conduct • EE commissioned or funded by decision makers • When EE purpose is to inform government policy • Increased collaboration between economists and decision makers • Mentorship by academics • Economists more available to local organizations • Development of academic-practitioner partnerships • Pooling resources across jurisdictions to conduct analyses • User support provided for economic model **Economic Resources Available** • Books, journals, lectures, internet materials available • EE studies easily accessible (e.g. widely read journals) **Political Awareness and Ethical Factors** • Increased political/public awareness of politically sensitive topics can lead to EE • Political pressure to use results of EE • Patient groups lacking power/legitimacy to protest against priority setting decisions • Political will to use EE • Ethical acceptance of EE • Seminars for politicians on EE • Encouraging public debate on costs and benefits of specific interventions • Patient demand for care may lead to EE • Corporate pressures/institutional image to provide care may lead to EE • To sell EE to clinicians, must be build on structure/beliefs/values already in health care system • Financial crisis was trigger for greater focus on resource allocation and need for EE • Greater use of ethics to resolve conflicts arising from different perspectives could support use of EE **Timeliness** • Results available to make timely decision **High-Quality Economic Evaluations** • Convinced of reliability of EE results: accurate and unbiased

- Political pressures when life-threatening conditions, orphan drugs or no other therapies available
- Considered unethical in some cases, e.g. terminally ill patients, when there are other ways of saving money
- Hospital-initiated prescribing, not individual
- Financial support and resources for EE not provided in organization
- Costs of conducting HE analysis
- Variability among healthcare organizations prevent EE being used
- Perceived as incompatible with ethical decisions because doesn't incorporate equity considerations

Decision-Making Process Issues

- Poor quality and transparency of decision-making process
- Poor reputation of research/decision-making agency
- EE not integrated into decision-making process
- Policy makers don't understand the process
- Function of formulary committees not understood
- Decision-making criteria not clearly specified
- Relationship of committee with decision-makers/resource allocators unclear
- EE inappropriately used to delay a decision

Lack of Relevance

- Outcomes not relevant for clinical practice, wider scope of resource use, longer-term outcomes
- EEs don't account for real-world implementation issues, e.g. implications in spending, increase in demand when technologies are approved
- Available information not relevant
- Results not transferable to local context (e.g. international results, hospital settings, health promotion programs)
- Not relevant for type of decisions made locally (e.g. funding for staff vs. CT scanner)
- EEs answer narrow questions that don't allow for complexity of decision-making processes in real-world (e.g. technologies vs programs)
- EE inappropriate for scope of policy questions to be addressed
- Not making decisions about interventions most important to health care providers (e.g. mental health/heart disease vs. cancer/expensive drugs)
- CE in practice different from what's predicted in clinical trial (efficacy vs. effectiveness)
- Results can't be used intuitively
- Results not generalizable or comparable

Availability/Awareness/Accessibility of Economic Information

- Lack of available EE studies, tools, data or databases
- Understanding uncertainties and limitations inherent in analyses
- Availability of good quality data and EE results
- Collecting better cost and outcome data to inform EEs
- Economic guidelines that increase comparability of studies and standardization in performing EE
- Improved methodological quality of EE studies
- Transparency and clarity of EE studies
- Ease of understanding assumptions, limitations and how model works
- Intuitive
- Accessible to practitioners outside academia
- Validated in different settings or prospective validation
- Sufficient sensitivity analyses are conducted

Relevant Economic Evaluations

- Can apply results to decision-makers setting
- More comparability of EE
- Outcome measures are clinically relevant and widely accepted
- Economists are more responsive to needs of decision-makers
- Decision-makers have an opportunity to identify and clarify relevant policy issues in the question and analysis, e.g. Two-stage process
- Efforts to adapt studies to own country or carry out local EE
- Analysis reflects constraints of local decision-making setting
- Subgroups in which treatment has acceptable CE are defined

Training in Health Economics

- More EE training opportunities for decision makers
- More education of decision-makers to make better use of EE as a tool
- Decision makers educated on HE techniques and relevance to their organization/practice
- More general awareness of economics
- Promotion of economics by professional bodies
- Dissemination of information on the benefits of EE tools
- National resource centre or archive of EE tools and models
- More experts trained in EE
- Integration of economics training into public health
- Educating clinicians to be comfortable with EE
- Adding economics to medical school curriculum so clinicians more sensitive to economic concepts
- Include in formal undergrad training
- Short easily assimilated courses developed in consultation with practitioners
- Economic training for formulary committee members
- Economist is on formulary committees

- Initiative not taken to obtain economic information
- Missing data (real life costs, hospitalizations) or unavailable/fragmented data required for EE
- EEs don't account for extra costs of adverse drug reactions
- Data necessary for EE doesn't exist at time when new drugs first marketed
- Never any future follow-up (when data becomes available) to confirm original analyses
- Clinical trialists not aware of needs of HE data requirements

Credibility of Methods

- Low credibility in quality of methods
- If information sources conflict
- Assumptions and extrapolation not justified so results not reliable
- Too many assumptions
- Skeptical of modeling
- Structure of model wrong
- Models not used appropriately
- Models are inflexible
- Poor transparency of studies
- Poor quality of data (clinical, cost, resource use, quality of life) used in EE
- Estimation methods not reliable
- Models not transparent re: what variables are included
- HE studies not comparable b/c costing methods vary
- Excessive variation in HE methodologies and need for more consistency
- No good guidelines for conducting EE
- Too few costs included
- Difficulties handling indirect costs
- Provides false sense of precision
- Will never be perfect but sometimes that's what people want
- Challenges with thresholds (disagree with threshold, don't know the threshold, no threshold exists, don't know how to implement a threshold)
- Challenges with QALYs (relevance, equity, lack of transparency, methodological limitations, concerns aggregating health benefits in a single index)

Not Timely

- Not available in time for decision
- Decisions have to be made quickly
- No time/way to assess whether previous decisions were still best in light of new decisions

Budget Barriers

- Broader expertise on decision-making committees
- Joint economic training so similar organizations/committees have same EE knowledge
- Funding and encouragement of economic training by decision-makers
- Protected funding for economic training

Trusted Sources or Credibility of Information

- Assessment/appraisal by trusted sources
- Independent commissioning of studies vs. relying on industry
- Impartial and objective EE that is not promoting a specific agenda
- Peer reviewed
- Increase credibility of EE research
- EE has been tested and piloted, therefore trusted
- Compares well with other models therefore trusted

Supportive Policies

- Official requirement to use EE
- Use of EE is integrated into decision-making processes
- Clear framework for using EE and direct link to legislation
- A national reimbursement agency with an economic orientation exists
- Clear identification of problems and how outputs of decision-making committees contribute to finance and resource allocation decisions
- Improved quality and transparency of decisions and decision-making processes
- Commitment from senior management to use EE
- Able to use local/national guidance to turn down patient requests
- Policies provide external motivation to use EE through incentive schemes (vs. internal motivation required to use EE)
- Management guidelines to encourage and assist use of EE
- Policies that promote integration of services and alignment of decision-making interests at different levels
- Creation of national agency to define priorities
- Existence of explicit priority setting processes
- Development of a weighting system for evidence and values
- Dynamic formularies that can adapt with evolving treatment practices and clinical data
- Integration of CE into clinical practice guidelines
- Economic research is funded by the government

Flexible Budgets

- More flexibility in health care budgets

• More interested in budget impact and cost containment than CEA • Budgetary constraints (i.e. limited funds) don't allow EE results to be applied • Even if CE, can't afford because of size of patient population • Know can't afford drug so don't do EE analysis • Fragmentation of health care services, lack of global budgets and inability to transfer between budgets • Decisions made with short-term perspective (budget containment) not long-term (e.g., CE prescribing) • Need clearer breakdown of which budgets will get savings • Only concerned with primary care drug costs • Economists don't understand healthcare budgets • Savings/costs are anticipated or theoretical, not real • Difficult to translate findings into concrete recommendations • Practical issues implementing EE recommendations **Conflict with Clinical Values** • Efficacy more important than EE for clinicians • Clinicians consider factors other than economics • Clinicians focus on individual patient vs. population • Undermines quality prescribing and patient-centered care • Perception it is about money/costs not about patients • Clinicians focus on own patients and specialty (silo mentality) • Not willing to ration care or consider opportunity costs - focus on own clinical area • Fear of losing credibility among patients • Clinician fear of malpractice and lawsuits • Difficulty denying patients drugs they ask for • Patients unwilling to accept health care resources are limited • Perceived difficulty in replacing a drug a patient is already using • Medical professionals dominate over healthcare policy-making and don't want to use EE • Some clinicians don't perceive responsibility/role in resource allocation (should be decision-makers) • Some clinicians don't accept non-clinician decision-makers can make resource allocation decisions (because they can't evaluate clinical outcomes appropriately) **Unaware of Drug Costs** • Unaware of actual drug costs (although more aware of relative drug costs) • Changing drug prices so difficult to know current price • Would need to know magnitude of cost savings (to outweigh patient	• Individuals receive a direct benefit to me/my department • If potentially large budget impact can promote use of EE • Requesting additional budgets to finance use of expensive drugs • Actual costs savings b/c comes out of same budget (vs. Theoretical savings) • Decentralization of drug budgets

switching issues)	
Abbreviations: CE=cost-effective; DM=decision-maker; EE=economic evaluation; GP=general practitioner; HE=health economic; ICER=incremental cost-effectiveness ratio; QALY=quality adjusted life year	

Supplementary Table 6: Types of barriers to using economic information and distribution across individual studies.

Author and Year	Type of Barrier Discussed in Study Publication												
	Industry Bias	Time required	Lack of Expertise	Political, Ethical or Institutional	Decision Making Process	Limited Relevance	Limited Availability or Accessibility	Credibility of Methods	Not in Time for Decisions	Challenges Communicating	Budgetary Constraints	Conflict b/w Clinical and Economic Values	Clinician Awareness of Drug Costs
POLICYMAKERS													
Ammerman 2009		x	x				x			x			
Anell 2000	x		x				x	x			x		
Baghbanian 2011	x		x	x		x		x	x		x		
Chaikledkaw 2009	x	x	x	x		x	x	x		x			
Chen 2007	x		x			x	x	x		x	x		
Davies 1994													
Drummond 2003	x		x			x		x		x	x		
Duthie 1999								x		x	x	x	
Eddama 2008		x	x	x	x	x					x	x	
Gunther 2006													
Hoffman 2000	x					x				x	x		
Hoffman 2002						x		x		x			
Lafi 2012			x	x			x					x	
Neumann 2008			x	x		x	x			x	x		
Niessen 2012					x	x	x	x		x	x		
Phillips 2010													
Rocchi 2008								x					
Ross 1995			x	x		x	x	x	x	x		x	
Taylor Robinson 2008			x			x		x		x			
Thurston 2008													
Van Dongen 2013													
Williams 2007	x		x	x			x	x			x	x	
Williams 2008	x	x	x	x	x	x	x	x	x	x	x	x	
CLINICIANS													
Fattore 2006			x	x		x				x	x		
Jansson 2006												x	
Prosser 2005	x	x	x	x			x	x				x	x

Author and Year	Type of Barrier Discussed in Study Publication												
	Industry Bias	Time required	Lack of Expertise	Political, Ethical or Institutional	Decision Making Process	Limited Relevance	Limited Availability or Accessibility	Credibility of Methods	Not in Time for Decisions	Challenges Communicating	Budgetary Constraints	Conflict b/w Clinical and Economic Values	Clinician Awareness of Drug Costs
Torbica 2010													
Walley 1997	x		x	x	x		x	x			x		
PATIENTS OR PUBLIC													
Peters 2007													
HETEROGENOUS POPULATIONS (POLICY MAKERS AND PRACTICING CLINICIANS)													
Erntoft 2011	x	x	x	x		x	x	x	x	x	x		
Galani 2008	x		x			x		x	x	x	x	x	
Rubenstein 2007	x		x	x		x	x			x	x	x	
Zwarat van Rijkom 2000	x	x		x		x	x	x		x	x	x	
N of 26 Studies	14	7	19	14	4	17	15	18	5	17	17	11	1
% Studies	54	27	73	54	15	65	58	69	19	65	65	42	4

'X' identifies the barrier being present.

Supplementary Table 7: Types of facilitators to using economic information and distribution across individual studies.

Author and Year	Type of Facilitator Discussed in Study Publication										
	Strong Collaborations	Economic Resources Available	Good Communication	Political and/or Ethical Awareness	Timely	High-Quality	Relevant	Training Available	Trusted or Credible	Supportive Policies	Flexible Budgets
Ammerman 2009	x	x									
Anell 2000											
Baghbanian 2011		x	x				x	x	x	x	x
Chaikledkaw 2009											
Chen 2007											
Davies 1994	x		x	x	x					x	
Drummond 2003						x	x				
Duthie 1999	x										x
Eddama 2008	x		x				x				
Gunther 2006											
Hoffman 2000		x	x				x	x			x
Hoffman 2002											
Lafi 2012									x	x	
Neumann 2008	x	x				x					
Niessen 2012			x	x		x				x	
Phillips 2010											
Rocchi 2008				x				x		x	
Ross 1995			x	x	x	x	x	x		x	
Taylor Robinson 2008	x		x			x			x		
Thurston 2008											
Van Dongen 2013											
Williams 2007			x				x	x		x	
Williams 2008	x		x				x	x		x	
CLINICIANS											
Fattore 2006						x		x			x
Jansson 2006											x
Prosser 2005										x	
Torbica 2010											
Walley 1997											
PUBLIC OR PATIENTS											
Peters 2007											
HETEROGENOUS POPULATIONS (POLICY MAKERS AND PRACTICING CLINICIANS)											

Author and Year	Type of Facilitator Discussed in Study Publication										
	Strong Collaborations	Economic Resources Available	Good Communication	Political and/or Ethical Awareness	Timely	High-Quality	Relevant	Training Available	Trusted or Credible	Supportive Policies	Flexible Budgets
Erntoft 2011				x		x				x	x
Galani 2008	x		x		x	x		x	x	x	x
Rubenstein 2007				x			x	x		x	x
Zwarat van Rijkom 2000							x				
N of 22 Studies	8	4	10	7	3	9	8	9	4	12	8
% Studies	36	18	45	32	13	41	36	41	18	54	36

'X' identifies the facilitator being present.

CHAPTER 4.

GUIDELINE STANDARDS FOR PRESENTING ECONOMIC EVALUATIONS: A SYSTEMATIC REVIEW

CHAPTER OVERVIEW

Purpose and Relationship to Other Manuscripts:

One of the goals of this thesis is to understand researchers' approaches to communicating and presenting economic evaluations. Therefore, the objective of this second manuscript is to systematically review guidelines for economic evaluation in health care for information on how to present economic evaluations. In addition it will identify and describe concepts commonly required for transparent reporting, but which may need further explanation for non-technical audiences. This work will be used to identify what guidance is available for researchers, i.e., knowledge producers, on how to communicate and present economic evaluations and to determine how their perspective differs from the informational needs of end-users of economic evaluations, i.e. knowledge users. These informational needs were identified in Chapter 3, a systematic review of literature on communicating and understanding economic information.

Supporting Material:

No supporting material

Manuscript Status:

Submitted for review and awaiting a decision. All information in this chapter, including the supplementary materials, has been submitted to the journal.

Authors:

Shannon M. Sullivan (SMS), George Wells (GW), Doug Coyle (DC)

Author Roles (SMS, GW, DC) and Contributions:

SMS, GW, DC were responsible for the study conception.
SMS, DC were responsible for the search for trials.
SMS, DC were responsible for the selection of trials.
SMS, DC were responsible for the data extraction.
SMS, DC were responsible for the data analysis.
SMS, GW, DC were responsible for the interpretation of results.
SMS, DC were responsible for drafting the manuscript.

All authors (SMS, GW, DC) provided comments on drafts and revisions to the manuscript, agree with the manuscript's results and conclusions, and read and approved the final manuscript.

Title: What Guidance are Economists Given on How to Present Economic Evaluations? A Systematic Review.

Authors: Shannon M. Sullivan MSc [1,2] George Wells PhD [1,3] Doug Coyle PhD [3]

1 University of Ottawa Heart Institute, Ottawa, Ontario

2 pan-Canadian Oncology Drug Review, Toronto, Ontario

3 University of Ottawa, Department of Epidemiology and Community Medicine, Ottawa, Ontario

Corresponding Author:

Shannon M. Sullivan, MSc

University of Ottawa Heart Institute;

40 Ruskin Street H1281, Ottawa, ON, Canada; K1Y 4W7

Running Title: Guidelines on Presenting Economic Evaluations

Article Type: Systematic Review

Key Words: economic evaluation; presentation formats; guidelines; systematic review

Acknowledgements: None

Financial Support: Canadian Institutes of Health Research

Conflict of Interest: None

KEY POINTS

(i) What is already known about the topic?

- Policy makers (and those impacted by policy decisions) can have difficulty understanding economic evaluations and creating simple but accurate explanations is challenging.
- There is little research on how economists should present work to non-technical audiences.
- Knowing what is important to economists when reporting on and presenting their work, could lead to understandable but accurate explanations of economic evaluations for non-technical audiences.

(ii) What does the paper add to existing knowledge?

- Current economic guidelines do not provide advice on how to tailor the presentation of economic evaluations to non-technical audiences
- Economists have consistent presentation formats and reporting elements that should be considered when developing user-friendly explanations and tools for general audiences.

(iii) (optional) What insights does the paper provide for informing health care-related decision making?

- Knowledge translation approaches taken for clinical evidence could be used to develop technically accurate but more simplified explanations of economic evaluations for general audiences
- Current approaches used by economists could inform the development of educational resources and initiatives to enhance health economic capacity among policy makers.

ABSTRACT

Objectives: To systematically review economic guidelines for information on how to present economic evaluations and describe concepts commonly required for transparent reporting, which may need further explanation for non-technical audiences.

Methods: Electronic databases and supplementary sources were searched for economic evaluation guidelines. Presentation format details and common reporting recommendations were extracted and frequencies tabulated

Results: Thirty guidelines were included. Many guidelines include a standard reporting format with some examples of tables and graphs to be included. Common presentation formats include well-cited tables of data sources; transparent model diagrams and descriptions; disaggregated results; and tabular and graphical displays of sensitivity analyses. Despite most guidelines being funded by policy makers for the purpose of developing economic evaluations for policy makers, few guidelines provided advice on presenting economic evaluations to non-economists. However, some guidelines included a glossary of economic terminology for non-technical readers. Common concepts that may require further explanation include differences in economic perspectives; appropriateness of time horizons; how economic outcomes such as quality-adjusted life-years relate to their component clinical outcomes; and choice of sensitivity analyses.

Conclusions: Economists have consistent presentation formats and common reporting elements that should be considered when developing user-friendly explanations for general audiences. These overlap with policy-makers informational needs but may not be sufficient for understanding by non-technical audiences. Developing presentation formats and tools that incorporate viewpoints of both economists and non-economists will allow for better application of the results of economic evaluations and enhance the transparency and legitimacy of decision-making processes that are informed by economic evaluations.

Introduction

Using cost-effectiveness analyses in health policy and reimbursement decisions has sometimes been questioned and the policy impact of economic evaluations is still uncertain after twenty years (1,2). Enhancing the transparency and understanding of economic evaluations could increase confidence in and perceived legitimacy of health technology assessment (HTA) decision-making processes (3,4). Economic evaluations are not always provided in an accessible format and more effective presentation of economic evaluations could increase understanding of how economic evaluations inform policy decisions. This would also enable policy makers to better defend decisions that are based, in part, on economic evaluations.

Health policy decision makers, and those impacted by policy decisions such as the tax-paying public, patients and health care providers, do not have the same technical expertise as health economists (5). However, they are important end-users of economic evaluations. Policy makers sometimes have limited training in health economics and may be unclear in how to practically apply the results of economic evaluations to actual decisions if results are not clearly presented (6,7).

Providing simple but robust explanations of economic evaluations to a non-technical audience is challenging (8). An appropriate balance of making information understandable while maintaining the completeness and technical accuracy of the economic evaluation must be achieved. Oversimplifying technical information may lose important nuances that are relevant for practical and consistent decision-making. Therefore, considering how economists present their work to policymakers and other lay audiences merits further exploration.

A number of studies have reported on policy makers' use of economic evaluations. In general, policy makers want to be able to deconstruct analyses, identify key principles and understand how results practically apply to specific decisions (9-11). It is unclear how this aligns with what economists consider important to present in economic evaluations.

Elements that economists report on are described in guidelines for conducting economic evaluations (12,13). The general purpose of these guidelines is to increase both the methodological quality of economic evaluations and the transparency of how they are conducted. While good reporting practices outlined in economic guidelines are important to follow and can contribute to clarity, they may be insufficient for good communication to non-technical audiences (5,14). It is unclear how much guidance economists are given with respect to presentation formats and tailoring presentation of their work to non-economists.

A systematic approach to knowledge translation (15) and the development of economic tools that would enhance accessibility of health economics and legitimacy of and confidence in health policy decision-making would be of value. Incorporating guidance from economic guidelines on reporting and presentation standards would lead to knowledge translation tools that can robustly represent complex economic information.

The primary objective of this study is to systematically review economic guidelines to determine what guidance economists are provided on how to present economic evaluations. A secondary objective is to identify fundamental elements and technical concepts in guidelines that are commonly required for transparent reporting, and which may need to be frequently explained to a general audience.

Methods

Systematic Review Protocol: A systematic review was conducted following Cochrane methodology (16). Guidelines for conducting or reporting on economic evaluations of pharmaceuticals were included that targeted policy makers, economists, researchers or others end-users of economic evaluations. Guidelines were excluded if they only addressed specific

components of economic evaluations (e.g. modeling, willingness to pay, etc.) and not economic evaluations as a whole; if they addressed affordability (i.e. budget impact) rather than cost-effectiveness; if they advocated a disease-specific or technology-specific approach to economic evaluation; editorials or opinion papers providing recommendations; if the guidelines were not directly linked to a decision-making or reimbursement process; if more recent versions of the guidelines were available or if the guideline was developed prior to 2000. The latter was necessary given the methodological advances in the discipline. Outcomes of interest were the type and frequency of information on presentation formats and the type and frequency of key reporting recommendations.

Search Strategy: Databases searched included Medline (1996 to November 2013), EMBASE (1980-November 2013) and the Cochrane Database of Systematic Reviews, using a date limit of 2000 but no language restriction. The search concepts were 'guidelines', 'health economics' and 'economic evaluation'. Grey literature was searched for unpublished reports using the CADTH Grey Matters checklist as a guide, in addition to other relevant economic resources, including the ISPOR website (17,18). Studies were also obtained through hand searching of selected journals, reviewing reference lists of potentially relevant studies and suggestions from economic experts.

Study Selection: Citations were screened for relevance by one review author based on the title and abstract of identified articles. Two review authors independently reviewed the full text of potentially relevant guidelines to assess exclusion or inclusion.

Critical Appraisal: Guideline quality was assessed using the AGREE II instrument (19). AGREEII was designed, in part, to assess the quality of clinical practice guidelines. No instruments currently exist to assess the quality of methodological guidelines or guidelines for economic evaluations. Therefore, minor modifications to the AGREEII instrument were made (see Supplementary Materials Table 1 for more details).

Data extraction: The following data were extracted from guidelines (see Supplementary Materials Table 2 for detailed variable definitions).

(1) Guideline characteristics: author affiliations, guideline purpose, type of guideline, target audience, if a standard reporting format, sample tables, sample figures or a glossary was provided, if advice was provided on presenting to non-economists.
(2) Presentation format: details on how to present data sources, model, base-case results and sensitivity analyses (deterministic and probabilistic) were extracted.
(3) Common reporting recommendations on: perspective, target population, subgroup analyses, time horizon, assumptions required, modeling, preferred outcome measures, sensitivity analysis parameters, sensitivity analysis ranges and methods for conducting sensitivity analyses

Detailed guidance on how to conduct an economic evaluation was not extracted (e.g. analytic approaches, distributions to apply) and is not the focus of this systematic review. When available, guideline characteristics from the ISPOR database (18,20) were used, and were verified with source documents by one author. Disagreement were verified by a second author. For all remaining data, one author extracted data, which was verified by a second author and disagreements were resolved by a third reviewer.

Data Analysis: The frequency of recommendations on reporting and presenting economic evaluations was tabulated and common trends were noted.

Results

Of the 6098 citations identified, 31 reports, representing 30 guidelines, were included (see Supplementary Figure) (13, 21-50). Key characteristics of the guidelines are described in Table 1 with additional details provided in Supplementary Table 3. Guideline quality was generally low

when critically appraised. Common limitations were not conducting a systematic search for evidence and an unclear process for formulating recommendations. Although recommendations were clearly identified in most guidelines, differences in terminology when comparing across guidelines could create confusion and lead to lack of clarity. More details on the critical appraisal are provided in Supplementary Table 4.

INSERT TABLE 1. Characteristics of Economic Evaluation Guidelines

Standard Reporting Templates

The majority of guidelines (70%, n=21/30) provided a standard reporting format for an economic evaluation. The specific content of the report format varied across guidelines but generally addressed major components of the analysis including the design, methods, evidence summary, results, and interpretation. Some guidelines provided sample tables (30%, n=9/30) and figures (17%, n=5/30) for how to present economic evaluations. However, when all guidelines were reviewed in detail, each provided some guidance on presenting various elements of economic evaluations and is included in the subsequent synthesis of presentation formats.

Presentation of Economic Evaluations to Non-Economists

Guidelines were most frequently developed by or in collaboration with decision-makers (87%, n=26/30) and targeted towards industry (90%, n=27/30), decision-makers (90%, n=27/30) or researchers (77%, n=23/30). Advice on presenting economic evaluations to non-technical audiences was noted in 17% guidelines (n= 5/30) One guideline suggested that model calculations should be clearly explained for non-economists and results disaggregated for formulary committees (30). One guideline suggested graphical presentations of results would assist general audiences (40). Three guidelines recommended including an executive summary accessible to non-economists (13,34,39) (See Supplementary Table 5 for more details). In addition, 37% guidelines (n=11/30) included a glossary of technical terms for those less familiar with health economics.

Recommendations on Presentation Formats

The most frequently recommended presentation formats (≥ 20%) are provided in Table 2.

INSERT TABLE 2. Most Frequently (≥20%) Recommended Presentation Formats Across Economic Guidelines

Presenting Data Sources

Sixty-six percent (n=20/30) of guidelines provided direction on how to present data sources. The most common suggestions on presentation formats were providing appropriate cross-referencing and citations of sources (46%) and providing detailed tables or a list of included variables and data (37%). Guidance on which variables to include in the tables varied with some requesting all variables, others requesting only the most important variables and some indicating only that there should be sufficient transparency to reproduce the analysis. Presenting a list or table including resource use and costs was also frequently recommended (23%).

Less frequently recommended formats for presenting data sources included: having copies of original data sources or ensuring their availability (13%), summarizing clinical trial data in a tabular format (10%) and including an assessment of the quality of the data sources (7%). Three guidelines (10%) also suggested providing more detailed information such as data collection forms or detailed data tables in appendices.

Presenting Economic Models

Seventy percent (n=21/30) of guidelines provided suggestions on presenting the economic model. Many guidelines (40%) emphasized the need for transparency and clarity in presenting the model description so that the analysis could be replicated and individuals can understand the justifications or why it is different from other existing models. It was noted that the presentation should allow individuals to follow the steps in the modeling and analysis and understand how the

data moves through the model. Similarly, a number of guidelines emphasized the need for clarity in presenting model assumptions (33%). Figures commonly suggested included decision tree diagrams (30%) or model structure diagrams (30%). For the model structure diagrams, guidelines either requested traditional transition-state Markov model diagrams or did not specify details of the model diagram. Only one guideline specified that a less-detailed schematic diagram of the model be provided for non-economists (30). Presenting transition probabilities in a matrix was less frequently recommended (10%). Two guidelines (7%) suggested model diagrams could be presented in appendices.

Presenting Base-Case Results

Ninety-three percent (n=28/30) of guidelines provided suggestions on how to present the base-case results. Although presenting incremental changes in results was predominantly recommended (60%), many guidelines recommended that total costs and benefits also be presented (47%) to aid in interpretation of the results. One guideline (3%) recommended presenting results using an efficiency frontier approach rather than incremental cost-effectiveness ratios. Some guidelines also specifically requested that uncertainty around the ICER be presented either using confidence intervals or a range for the ICER (13%). Some guidelines (9%) recommend presenting results using natural clinical units before converting to QALYS (e.g. heart attacks, cirrhosis, adverse events avoided). A few guidelines also suggested presenting costs and benefits at a per patient level, not just a population level (9%).

Many guidelines recommended reporting results in disaggregate as much as possible, leading up to aggregate responses (63%). Examples of disaggregating results included presenting costs and effects separately for each intervention and presenting results for different subpopulations, disease states, model cycles or healthcare settings. In addition, some guidelines recommended separately reporting QALYs and LYs when there was a survival benefit (13%). Some guidelines (13%) indicated that results should be presented in a way to allow verification or reproducibility.

Two guidelines (6%) also indicated that the limitations and/or quality of the results should be described for transparency purposes.

Presenting results in tables was frequently recommended (37%). Specific suggestions for tables included summarizing costs, effects and cost effectiveness for each therapy or comparing short-term trial-based results with long-term extrapolations. Presenting results using general graphical or visual displays was encouraged (23%). Specific types of graphics included the cost-effectiveness plane (13%), net benefit (6%), changes in outcomes over time when permitted by the model (3%) and willingness to pay (3%).

Some guidelines also recommended discussing practical issues when presenting results. For example, presenting results that allow one to determine relevance or generalizability to the population was recommended in two guidelines (6%). This could include consideration of disease epidemiology, population characteristics, practice patterns, or patterns of resource use. In addition, two guidelines (6%) requested that the most influential factors impacting the ICER be clearly presented. One guideline (3%) also requested that the magnitude and direction of this impact also be clearly identified. Some guidelines (13%) requested presentation of cost savings and/or opportunity costs when relevant. Additional suggestions related to the relevance and applicability of results included: noting implementation factors and the feasibility of adopting based on financial constraints (3%) or the ability to relocate freed resources for other purposes (3%).

Two guidelines (6%) provided suggestions on how to present conclusions. For example, one guideline indicated that conclusions should always present the uncertainty around the estimate and that dominance should not be concluded unless based on a wide range of sensitivity analyses (38). One guideline provided standard wording suggestions for presenting cost-effectiveness results (43): (a) “compared to another product, this product costs $X, or the range from $Y to $Z, for an additional quality-adjusted life-year (QALY)”; (b) “compared to the

alternative therapy, this product increases cost by $X (the marginal cost), but reduces Y major side effects and Z other side effects."

Presenting Sensitivity Analyses: Deterministic and Probabilistic

Deterministic Sensitivity Analyses

Sixty percent (n=18/30) of guidelines provided suggestions on how to present the deterministic sensitivity analyses results. A summary of what parameters sensitivity analyses should be conducted on is described later, in Table 3, as part of guideline reporting requirements.

The most common recommendation was presenting sensitivity analyses in tables (27%). Specific types of tables included: a table comparing sensitivity analyses of broader drug use with the intended drug use, a matrix of variable effects and outcomes that differ between two comparators, tabulating the sensitivity analyses beside the base case results.

Presenting sensitivity analyses using tornado diagrams to highlight key drivers (23%) was also frequently recommended. Other graphical recommendations included presenting the last gradient of efficiency frontier as function of varying parameter (31), line graph showing parameter change vs. ICER change (40) or a graph that includes threshold values (35). Two guidelines recommended graphs but did not specify what type.

Identifying influential parameters that impact on results was also frequently recommended (23%). Two guidelines also recommended specifying the magnitude and direction of those effects. The elasticity of the parameters or the percent change from the reference case were also recommended formats for presenting sensitivity analyses (13,42). One guideline recommended identifying the parameter values above or below which the technology would be cost savings (31).

Probabilistic Sensitivity Analyses

Forty percent (12/30) of guidelines provided suggestions on how to present probabilistic sensitivity analyses. The most frequently recommended formats were the cost effectiveness acceptability curve (33%) and the scatter plot for the cost-effectiveness plane (30%).

Less frequent recommendations included presenting in a tabular format (10%), presenting expected value of information analyses (3%) and identifying the percent of simulations with cost savings (3%). Other graphical presentations that were recommended in only one guideline included CEAF (3%), histogram of incremental net benefit results (3%), the cumulative distribution of the parameter over the efficiency frontier (3%) and general diagrams (3%).

Frequent reporting recommendations that may require additional explanation for non-economists

The most frequent reporting recommendations (≥ 20%) are provided in Table 3. Less frequent recommendations (<20%) are outlined in Supplementary Table 6.

INSERT TABLE 3. Most Frequent Recommendations (≥20%) for Reporting and Conducting Economic Evaluation Guidelines

The perspectives most frequently recommended were the healthcare payer and the societal. Some guidelines suggested presenting multiple perspectives and, in some guidelines, even though one perspective was recommended as the base case, others perspectives were identified as optional.

The approved indication (60%) or a clearly specified target population (40%) were the most frequently recommended populations. Subgroup analyses were requested or permitted (67%) but common guidance across guidelines on which subgroups should be analyzed was not provided.

It was most frequently recommended to use a time horizon that could account for all of the outcomes and consequences (60%). In addition, considering the disease characteristics when selecting the time horizon was also frequently recommended, e.g. natural history, chronic vs. acute, long-term effects (27%)

Assumptions were permitted in most guidelines (67%) and modeling details were also required (60%). Common guidance was not provided for either aspect except to indicate that choices should be justified and transparently described.

Outcome measures most frequently preferred were QALYs (83%), natural clinical measures (73%) (e.g. adverse events avoided, heart attacks) that could be converted to QALYs, life years or mortality (67%) and quality of life (37%).

The most frequent methods of conducting sensitivity analyses were one-way (67%), multi-way (53%) and probabilistic (50%). Uncertain parameters were most frequently explored in sensitivity analyses (60%). The parameter values used in sensitivity analyses were most frequently based on confidence intervals (27%), extreme best and worst case scenarios (23%) and those representing a full or credible range of variability (27%).

Discussion

Key Findings

Most guidelines offer some advice on how to present economic evaluations and many include a standard reporting format with examples of tables and graphs to be included. Common ideas across guidelines include presenting well-cited tables of data sources; transparent model descriptions and diagram; disaggregated results; and tabular and graphical displays of sensitivity analyses. However, despite most guidelines being funded by policy makers for the purpose of developing economic evaluations for policy makers, few guidelines provided specifics on how to present economic evaluations to non-economists. Common elements that are required for transparent reporting of economic evaluations and that may require further explanation for non-economists include differences between the societal and healthcare payer perspective, adequacy of time horizons that cover the period of costs and effects while appropriately considering aspects of the disease, how economic outcomes such as QALYs and LYs relate to their component clinical outcomes such as heart attacks or adverse events and how sensitivity analyses (methods, parameters, ranges) are chosen to be conducted. Best practices and educational resources for communicating these concepts to non-technical audiences could be further explored. Some guidelines included a glossary of economic terminology for non-technical readers.

Comparison with Other Literature

These findings should be considered in light of (a) what policy makers want from economic evaluations and (b) frequent limitations in economic models, to determine if this guidance is sufficient for presenting economic evaluations to broader audiences. Bridging this gap and finding common ground between how economists present economic evaluations and what decision-makers need to understand from an economic evaluation is essential for creating economic

evaluations that adequately balance both rigor and accessibility for those who must apply their results and understand their implications in the real-world. A recent ISPOR task force partially attempted to address this gap by developing a questionnaire for decision-makers to assess the relevance and credibility of economic modeling studies (51).

The recommendations on presentation formats and reporting elements identified in these 30 guidelines sometimes, but not always, overlap with issues raised by policy makers when applying the results of economic evaluations to inform decisions. Some similarities include ensuring transparency of methods and disaggregation of results. For example, policy makers have indicated that being able to deconstruct the economic analysis (9) and clearly outlining model assumptions that contribute to an ICER (10) can facilitate their understanding. However, policy makers have also focused on understanding the relevance of economic evaluation results to practical decisions and implementation issues. For example, appreciating the difference between theoretical cost savings and savings that can be realized in practice or recognizing opportunity costs are important considerations for policy makers. However, only a few guidelines emphasized the importance of presenting cost savings or other factors related to implementation and adoption feasibility. Even when specific suggestions were provided on how to present information to target audiences, these centered on enhancing clarity and transparency, not pointing out relationships to real-world decisions.

Across guidelines, there were no common recommendations on details to include when reporting on modeling and assumptions. However, these are important concepts when understanding the limitations of economic evaluations. Common limitations identified in submissions of economic evaluations for reimbursement frequently included uncertain clinical estimates due to clinical and cost assumptions or construction of the economic model (52-54). Ensuring that sufficient detail on the methods is provided to allow for consistent decision-making, without overwhelming non-technical users requires some consideration. Some guidelines recommended presenting more

detailed tables or technical information on modeling in appendices, which could help audiences focus on the most relevant information in the reports.

Strengths and Limitations

This is the first work to systematically review the guidance economists are given on how to present economic evaluations and consider this in the context of decision-makers and other lay audiences with an interest in understanding the results of economic evaluations. The consistency of recommendations across 30 guidelines suggests that key concepts were captured in this systematic review. Methodological guidelines do not exist to critically appraise the quality of economic evaluations and a tool for critically appraising clinical practice guidelines was adapted. Therefore, concepts unique to methodological guidelines may have been missed, however, the appraisal allowed identification of some key issues in the guidelines. Extracting qualitative data from a wide variety of guideline documents requires judgment. However, this task was facilitated by the use of a pre-existing database of guideline characteristics (18,20).

Policy Implications/Interpretation

Economists have consistent presentation formats and reporting elements that should be considered when providing information to policy makers. However, tailoring information to non-economists is also important to enhance the accessibility of economic evaluations and the legitimacy of their use in health policy decision-making. Although policy makers have influenced how economic evaluations are conducted, less work has been done on how best to present and communicate economic evaluations to non-economists. Some considerations for enhancing presentation formats are outlined below:

(1) Presentations of economic evaluations are currently targeted to technical users. Policy makers have reported challenges understanding technical concepts such as QALYs and

current graphics such as CEACs (55-57). Linking technical concepts to ideas that have practical relevance in a decision-making context and that are already familiar to policy makers could be valuable. In addition, ensuring educational resources and initiatives are in place for policy makers to gain a better understanding of commonly reported technical elements is important.

(2) Decision-makers have indicated that disaggregating results and focusing on the practical relevance of results is important. Economists have also indicated that presenting disaggregated results is important for transparent reporting. However, while disaggregation is important, on its own, it may be insufficient for policy makers. Following frameworks on how to communicate science to the lay public could guide audiences in better understanding these technical concepts, e.g. judicious choice of content, structuring and organizing knowledge, use of analogies to explain novel concepts and use of narratives (58).

(3) Even when reporting guidelines are followed, there will always be complex methodological and modeling decisions that are open to interpretation and their relevance to decision-making context should be considered. For example limitations in the clinical evidence base and uncertain extrapolations of these data frequently influence the results of economic evaluations (52-54). Ensuring these uncertainties and potential limitations are clearly explained to non-economists is important for enhancing the credibility of and consistency in health policy decision-making.

Future Research and Conclusions

Knowledge translation approaches have frequently been applied to enhancing understandability and accessibility of clinical trial evidence for policy makers, healthcare providers and patients (59-61). However, these approaches have not been applied in the field of health economics and

could be used to develop technically accurate but more simplified explanations of economic evaluations for general audiences. For example, in addition to providing technical reports and scientific publications of economic evaluations, decision support tools could be developed for complex decisions. They could also be used to inform other educational initiatives or supplementary resources for policy makers and other non-economists to understand economic evaluations. With further development of presentation formats that are tailored to non-economists, individuals will be able to better apply the results of economic evaluations to policy decisions and enhance the transparency and legitimacy of decision-making processes.

REFERENCES

1. Schlander M. The use of cost-effectiveness by the National Institute for Health and Clinical Excellence (NICE): no(t yet an) exemplar of a deliberative process. J Med Ethics 2008; 34: 534-539.

2. Drummond M. Twenty years of using economic evaluations for drug reimbursement decisions: what has been achieved? J Health Polit Policy Law. 2013;38(6): 1081-102.

3. Daniels N & Sabin J. Limits to health care: Fair procedures, democratic deliberations and the legitimacy problem for insurers. Philosophy & Public Affairs 1997; 303-350.

4. Lehoux P, Denis JL, Tailliez S et al. Dissemination of health technology assessments: identifying the visions guiding an evolving policy innovation in Canada. J Health Polit Policy Law 2005;30: 603-641.

5. Thurston SJ, Craig D, Wilson P, Drummond MF. Increasing decision-makers' access to economic evaluations: alternative methods of communicating the information. Int J Technol Assess Health Care. 2008 Spring;24(2):151-7.

6. Chaikledkaew U, Lepitakpong C, Teerawattananon Y, et al. The current capacity and future development of economic evaluation for policy decision making: a survey among researchers and decision makers in Thailand. Value in Health 2009; 12: S31-S35.

7. Phillips CJ, Fordham R, March K, Bertranou E, et al. Exploring the role of economics in prioritization in public health: what do stakeholders think? Eur J Public Health 2010;1-7.

8. Choi B, Pang T, Lin V, Puska P, Sherman G, Goddard M et al. Can scientists and policy makers work together? J Epidemiol Community Health 2005; 59:632–637.

9. Taylor-Robinson D, Milton B, Lloyd Williams F, et al. Policy makers' attitudes to decision support models for coronary heart disease: a qualitative study. J Health Serv Res Policy 2008;13: 209-214.

10. Davies L, Coyle D, Drummond M. Current status of economic appraisal of health technology in the European community: report of the network. Soc Sci Med 1994;38: 1601-1607.

11. Hoffman C, Graf von der Schulenburg JM. The influence of economic evaluation studies on decision making: a European survey. Health Policy 2000; 52: 179-192.

12. Husereau D, Drummond M, Petrou S, Carswell C, Moher D, Greenberg D, et al. Consolidated Health Economic Evaluation Reporting Standards (CHEERS) statement. BMJ 2013; 346: f1049.

13. CADTH. Guidelines for the economic evaluation of health technologies: Canada [3rd Edition]. Ottawa: Canadian Agency for Drugs and Technologies in Health; 2006.

14. Galani C & Rutten F. Self-reported healthcare decision-makers' attitudes towards economic evaluations of medical technologies. Curr Med Res Opin. 2008; 24(11): 3049-3058.

15. Straus S, Tetroe J & Graham ID (Eds.), Knowledge translation in health care: Moving from evidence to practice, 2nd edition. Oxford: Blackwell Publishing, 2013.

16. Cochrane Collaboration: Cochrane Handbook for Systematic Reviews. Available from: www.cochrane.org [Last accessed May 19, 2014].

17. CADTH. Grey Matters: a practical search tool for evidence-based medicine [2nd edition]. Ottawa: Canadian Agency for Drugs and Technologies in Health; 2013.

18. International Society for Pharmacoeconomics and Outcomes Research. Pharmacoeconomic (ISPOR) guidelines around the world. Available from: http://www.ispor.org/PEguidelines/index.asp [Accessed May 19, 2014]

19. Brouwers MC, Kho ME, Browman GP, Burgers JS, Cluzeau F, Feder G, Fervers B, Graham ID, Grimshaw J, Hanna SE, Littlejohns P, Makarski J, Zitzelsberger L; AGREE Next Steps Consortium. AGREE II: advancing guideline development, reporting and evaluation in health care. J Clin Epidemiol. 2010 Dec;63(12):1308-11. doi: 10.1016/j.jclinepi.2010.07.001.

20. Eldessouki R, Smith MD, Health Care System Information Sharing: A Step Toward Better Health Globally. Value Health Regional Issues 2012; 1:118-129

21. Bae S, Lee S, Bae E, Jang S. Korean Guidelines for Pharmacoeconomic Evaluation (Second and Updated Version). Consensus and Compromise. PharmacoEconomics 2013, 31:257–267.

22. Pharmaceutical Economic Evaluation Guidelines. South Korea: Health Insurance Review Agency, 2006.

23. Guidelines on How to Conduct Pharmacoeconomic Analyses. Norway: The Norwegian Medicines Agency (NOMA), 2012.

24. China Guidelines for Pharmacoeconomic Evaluations. China: Pharmacoconomics China 2011.

25. Guidelines for the Economic Evaluation of Health Technologies in Ireland. Ireland: Health Information and Quality Authority, 2010.

26. Lopez-Bastida J, Oliva J, Antonanzas F, Garcıa-Altes A, Gisbert R, Mar J et al. Spanish recommendations on economic evaluation of health technologies. Eur J Health Econ 2010, 11: 513–520.

27. Avksentieva MV, Antonova NV, Arutyunov GP, et al. Procedure for Clinical and Economic Evaluation of Drug Lists That Are Submitted for Reimbursement Coverage from Public Health Care Budget. Russia: Research Center for Clinical and Economic Evaluation and Pharmacoeconomics of the Russian State Medical University; 2010.

28. Guidelines for the Submission of a Request to Include a Pharmaceutical Product in the National List of Health Services, version 8. Israel: Ministry of Health Pharmaceutical Administration, 2010.

29. Regulations Relating to a Transparent Pricing System for Medicines and Scheduled Substances: (Draft Guidelines for Pharmacoeconomic Submissions). South Africa: Department of Health, 2010.

30. The AMCP Format for Formulary Submissions Version 3.0. United States. Alexandria, VA: Academy of Managed Care Pharmacy, 2009.

31. General Methods for the Assessment of the Relation of Benefits to Costs, version 1.0. Germany: Institute for Quality and Efficiency in Health Care (IQWiG), 2009.

32. Guidelines for Preparing a Health Economic Evaluation [unofficial translation]. Finland: Ministry of Social Affairs and Health, 2009.

33. Guidelines for Conducting Health Technology Assessment, version 2.1. Poland: Agency for Health Technology Assessment, 2009.

34. De Mello Vianna CM, Caetano R. Methodological Guidelines: Economic Evaluation of Health Technologies. Brazil: Ministry of Health; 2009.

35. Longson C, Longworth L, Turner K. Guide to the Methods of Technology Appraisal. England & Wales: National Institute for Health and Clinical Excellence (NICE), 2008.

36. Cleemput I, Van Wilder P , Vrijens F , Huybrechts M, Ramaekers D . Guidelines for Pharmacoeconomic Evaluations in Belgium. Belgium: Federaal Kenniscentrum voor de Gezondheidszorg; 2008;

37. Guidelines for Economic Evaluation of Health Care Interventions. Slovakia: Ministry of Health Slovak Republic, 2008.

38. Guidelines for Preparing Submissions to the Pharmaceutical Benefits Advisory Committee, Version 4.3. Australia: Pharmaceutical Benefits Advisory Committee, 2008.

39. Bertozzi SM, Mendoza AV, Garduno AH, Barajas ER. Guidelines for Economic Appraisal Studies for Updating the Health Sector's National Formulary in Mexico. Mexico: National Institute of Public Health, 2008.

40. Teerawattananon Y, Chaikledkaew U. Thai health technology assessment guideline development. *J Med Assoc Thai*. 2008;91(Suppl.):S11–S15.

41. Guidance to Manufacturers for Completion of New Product Assessment Form, Revised. Scotland: Scottish Medicines Consortium (SMC), 2007.

42. Grocott R, Schoeler R, Priest V, et al. Prescription for Pharmacoeconomic Analysis, version 2. New Zealand: Pharmaceutical Management Agency (PHARMAC), 2007.

43. Guidelines of Methodological Standards for Pharmacoeconomic Evaluations. Taiwan: Taiwan Society for Pharmacoeconomics and Outcomes Research, 2006.

44. Guidelines for Pharmacoeconomic Research. The Netherlands: The Health Care Insurance Board, 2006.

45. Boulenger S, Ulmann P. French Guidelines for the Economic Evaluation of Health Care Technologies. France: Collège des Économistes de la Santé, 2004.

46. González AMG. Methodological Guidelines for Health Economic Evaluation. Cuba: Ministry of Public Health; 2003.

47. Edling A, Stenberg AM. General Guidelines for Economic Evaluations from the Pharmaceutical Benefits Board. Sweden: The Pharmaceutical Benefits Board, 2003.

48. Behmane D, Lambot K, Irs A, et al. Baltic Guideline for Economic Evaluation of Pharmaceuticals (Pharmacoeconomic Analysis). The Baltics: Baltic Health Authorities; 2002.

49. Szende A, Mogyorosy Z, Muszbek N, Nagy J, Palloss G, Dozas C. Methodological guidelines for conducting economic evaluation of healthcare interventions in Hungary: a Hungarian proposal for methodology standards. Eur J Health Econom 2002 3:196-206

50. Capri S, Ceci A, Terranova L, Merlo F, Mantovani L, Members of the Italian Group for Pharmacoeconomic Studies. Guidelines for economic evaluations in Italy: recommendations from the Italian group of pharmacoeconomic studies. Drug Information Journal 2001, 35: 189–201.

51. Caro J, Eddy DM, Kan H et al. Questionnaire to assess relevance and credibility of modeling studies for informing health care decision making: an ISPOR-AMCP-NCP Good Practice Task Force report. Value in Health 2014, 17: 174-182.

52. Drummond M & Sculpher M. Common Methodological Flaws in Economic Evaluations. Med Care 2005, 43 (7 Suppl): 5-14.

53. Kaltenthaler E, Boland A, Carroll C, et al. Evidence Review Group approaches to the critical appraisal of manufacturer submissions for the NICE STA process: a mapping study and thematic analysis. Health Technol Assess 2011, 15: 1-82.

54. Hill SR, Mitchell AS, Henry DA. Problems with the interpretation of pharmacoeconomic analyses: a review of submissions to the Australian Pharmaceutical Benefits Scheme. JAMA. 2000, 283: 2116-21.

55. Chen LC, Ashcroft DM, Elliott RA. Do economic evaluations have a role in decision-making in medicine management committees? A qualitative study. Pharm World Sci 2007; 29: 661-670.

56. Hoffman C, Stoykova BA, Nixon J, et al. Do health care decision makes find economic evaluations useful? The findings of focus group research in UK health authorities. Value in Health 2002;5: 71-79.

57. Williams I, McIver S, Moore D, Bryan S. The use of economic evaluations in NHS decision- making: a review and empirical investigation. Health Technol Assess 2008; 12: 1-175.

58. Baram-Tsabari A and Bruce V. Lewenstein An Instrument for Assessing Scientists' Written Skills in Public Communication of Science Science Communication 35(1) 56–85 2013

59. Chambers D, Wilson P. A framework for production of systematic review based briefings to support evidence-informed decision-making. Syst Rev 2012; 1: 32.

60. Akl EA, Guyatt GH, Irani J, Feldstein D, Wasi P, Shaw E, Shaneyfelt T, Levine M, Schünemann HJ. "Might" or "suggest"? No wording approach was clearly superior in conveying the strength of recommendation. J Clin Epidemiol 2012; 65: 268-275.

61. Trevena L, Davey H, Barratt A, Butow P, Caldwell P. A systematic review on communicating with patients about evidence. J Evalu Clin Practice 2006;12: 13-23.

TABLES AND FIGURES

Table 1. Characteristics of Economic Evaluation Guidelines

Country	Author Affiliation	Target audience	Standard Reporting Format	Sample Tables Included	Sample Figures Included	Glossary Included	Presenting to Non-Economists
S. Africa 2013	Decision-Makers	Industry Decision-Makers Researchers	Yes	-	-	-	-
S. Korea 2013	Decision-Maker	Industry Decision-Makers Researchers	-	Yes	Yes	Yes	-
Norway 2012	Decision-Maker	Industry Decision-Makers Researchers	No	-	-	-	-
China 2011	Decision-Maker Researchers	Industry Decision-Makers Researchers Clinical Providers	Yes	-	-	Yes	-

Country	Author Affiliation	Target audience	Standard Reporting Format	Sample Tables Included	Sample Figures Included	Glossary Included	Presenting to Non-Economists
Russia 2010	Researchers	Industry Decision-Makers Researchers	Yes	Yes	-	-	-
Ireland 2010	Decision-Maker	Industry Decision-Makers Researchers	-	Yes	Yes	Yes	-
Israel 2010	Decision-Maker	Industry Decision-Makers Researchers	Yes	Yes	-	-	-
Brazil 2009	Decision-Makers Researchers	Industry Decision-Makers Researchers	Yes	-	-	Yes	Yes
Finland 2009	Decision-Maker	Industry Researchers	-	-	-	-	-
Germany 2009	Decision-Maker	Decision-Makers	-	-	Yes	Yes	-

Country	Author Affiliation	Target audience	Standard Reporting Format	Sample Tables Included	Sample Figures Included	Glossary Included	Presenting to Non-Economists
Spain 2009	Decision-Makers Researchers	**All:** Industry Decision-Makers Researchers Health Providers Patients Citizens	**-**	**-**	**-**	**-**	-
USA 2009	Decision-Maker	Decision-Makers Researchers	Yes	Yes	-	Yes	Yes
Australia 2008	Decision-Maker	Industry	Yes	Yes	-	-	-
Belgium 2008	Decision-Maker	Industry Decision-Makers Researchers	Yes	Yes	-	-	-
England & Wales 2008 *	Decision-Maker	Industry Researchers	-	-	-	Yes	Yes
Mexico 2008	Researchers Decision-Makers	Industry Decision-Makers Researchers	Yes	-	-	Yes	Yes

Country	Author Affiliation	Target audience	Standard Reporting Format	Sample Tables Included	Sample Figures Included	Glossary Included	Presenting to Non-Economists
Slovak 2008	Decision-Makers Researchers	Industry Decision-Makers Researchers	Yes	-	-	-	-
Thailand 2008	Researchers Decision-Makers	Industry Decision-Makers Researchers	Yes	-	Yes	-	Yes
Scotland 2007	Decision-Makers	Industry Decision-Makers	-	-	-	-	-
New Zealand 2007	Decision-Maker	Industry Decision-Makers Researchers	Yes	Yes	-	Yes	-
Poland 2007	Decision-Maker Researchers	Industry Decision-Makers Researchers	Yes	-	-	-	-
Canada 2006	Decision-Maker	Decision-Makers	Yes	-	Yes	Yes	Yes
Taiwan 2006	Researchers Industry	Industry Decision-Makers Researchers	Yes	-	-	-	-

Country	Author Affiliation	Target audience	Standard Reporting Format	Sample Tables Included	Sample Figures Included	Glossary Included	Presenting to Non-Economists
Netherlands 2004	Decision-Maker Researchers	Industry Decision-Makers Researchers	Yes	-	-	-	-
Cuba 2003	Decision-Makers Researchers	Industry Decision-Makers Researchers	Yes	-	-	Yes	-
France 2003	Researchers	Industry Decision-Makers Researchers	Yes	-	-	-	-
Sweden 2003	Decision-Maker	Industry Decision-Makers	-	-	-	-	-
Baltic States 2002	Decision-Maker	Industry Decision-Makers	-	-	-	-	-
Hungary 2002	Decision-Maker Researchers	Industry Decision-Makers Health Providers	Yes	-	-	-	-

Country	Author Affiliation	Target audience	Standard Reporting Format	Sample Tables Included	Sample Figures Included	Glossary Included	Presenting to Non-Economists
Italy 2001	Decision-Maker Researchers Industry	Industry Decision-Makers Researchers	Yes	-	-	-	-

* At the time of this review, guidance from England & Wales was being updated but had not yet been finalized.

Table 2. Most Frequently (≥20%) Recommended Presentation Formats Across Economic Guidelines

Recommendations on How to Present Information	n (%) Guidelines, N=30
Data Sources	
Cross-referencing and citations of sources	14 (47)
Detailed table of data and variables	11 (37)
Table of resource use and costs	7 (23)
Model	
Model description transparent enough to justify and/or replicate	12 (40)
Assumptions and/or methods clearly presented	10 (33)
Decision tree diagram	9 (30)
Relevant model diagram and structure presented	9 (30)
Base Case Results	
Disaggregated results to increasingly aggregated	19 (63)
Incremental costs and benefits presented	18 (60)
Total costs and benefits presented	14 (47)
Tables of results	11 (37)
Graphical displays requested or encouraged	7 (23)
Sensitivity Analyses	
Deterministic	
Tabular format	8 (27)
Tornado diagrams to highlight drivers	7 (23)
Identify influential parameters impacting results	7 (23)
Probabilistic	
Cost effectiveness acceptability curve	10 (33)
Scatter plot – cost effectiveness plane	9 (30)

Note: multiple recommendations could be made in each guideline

Table 3. Most Frequent Recommendations (≥20%) for Reporting and Conducting Economic Evaluation Guidelines

Guideline Element	Recommendations on Conduct and Reporting	n (%) Guidelines, N=30
Primary Perspective	Healthcare payer	18 (60)
	Societal	14 (47)
Target Population	Approved indication	18 (60)
	Clearly specified target population	12 (40)
Subgroup Analyses	Requested/permitted but no specific guidance	20 (67)
	Define relevant subgroups	6 (20)
Time Horizon	Long enough to cover consequences and benefits	18 (60)
	Disease characteristics considered	8 (27)
Assumptions Required	Yes	20 (67)
Modelling Permitted	Yes, requires details	11 (37)
	Yes, details not specified	7 (23)
Preferred Outcome Measures	Quality-adjusted life years	25 (83)
	Natural, clinically relevant or patient-relevant units	22 (73)
	Life-year gained or mortality	20 (67)
	Quality of life	11 (37)
Sensitivity Analyses - Parameters	Parameters with greatest uncertainty	18 (60)
Sensitivity Analyses - Range	Confidence intervals	8 (27)
	Representing full or credible range of variability	8 (27)
	Extreme or best/worst case values	7 (23)
Sensitivity Analyses - Methods	One way or univariate	20 (67)
	Multi-way or multivariate scenario analyses	16 (53)
	Probabilistic	15 (50)

Note: multiple recommendations could be made in each guideline
Note: For more details, see ISPOR Guideline Tables (18). For less frequent recommendations (<20%), see Supplementary Materials

SUPPLEMENTARY MATERIALS – METHODS

Supplementary Table 1. AGREE II (Unmodified)* for Critical Appraisal of Clinical Practice Guidelines

* To apply AGREE II to the critical appraisal of methodological guidelines for economic evaluations (rather than to clinical practice guidelines), the following items were modified:

- For item 2 rather than 'health' question, the 'methodological question(s) covered by the guideline is (are) specifically described.
- For item 11 rather than 'health benefits, side effects and risks/, the 'advantages and disadvantages have been considered in formulating the recommendations'
- Item 16 was removed: 'the different options for management of the condition or health issue are clearly presented'

Domain	Item
Scope and purpose	1. The overall objective(s) of the guideline is (are) specifically described.
	2. The health question(s) covered by the guideline is (are) specifically described.*
	3. The population (patients, public, etc.) to whom the guideline is meant to apply is specifically described.
Stakeholder involvement	4. The guideline development group includes individuals from all the relevant professional groups.
	5. The views and preferences of the target population (patients, public, etc.) have been sought.
	6. The target users of the guideline are clearly defined.
Rigor of development	7. Systematic methods were used to search for evidence.
	8. The criteria for selecting the evidence are clearly described.
	9. The strengths and limitations of the body of evidence are clearly described.
	10. The methods for formulating the recommendations are clearly described.
	11. The health benefits, side effects and risks have been considered in formulating the recommendations.*
	12. There is an explicit link between the recommendations and the supporting evidence.
	13. The guideline has been externally reviewed by experts prior to its publication.
	14. A procedure for updating the guideline is provided.

Clarity of presentation	15. The recommendations are specific and unambiguous.
	16. The different options for management of the condition or health issue are clearly presented.*
	17. Key recommendations are easily identifiable.
Applicability	18. The guideline describes facilitators and barriers to its application.
	19. The guideline provides advice and/or tools on how the recommendations can be put into practice.
	20. The potential resource implications of applying the recommendations have been considered.
	21. The guideline presents monitoring and/ or auditing criteria.
Editorial independence	22. The views of the funding body have not influenced the content of the guideline.
	23. Competing interests of guideline development group members have been recorded and addressed.

Supplementary Table 2. Definitions of Variables and Data Extracted in Systematic Review

Variables	Categories and/or Definitions
Guideline Type *	• **Published PE Recommendations**: country-specific economic evaluation guidelines or recommendations published by experts in the field but are not "officially" recognized or required by the healthcare decision making bodies/entities in this country/region for reimbursement • **PE Guidelines**: country-specific "official" guidelines or policies concerning economic evaluation that are recognized or required by the healthcare decision making bodies/entities in this country/region for reimbursement. • **Submission Guidelines**: country-specific "official" guidelines or policies concerning drug submission requirements with an economic evaluation part/section and are required by the healthcare decision making bodies/entities in this country/region for reimbursement.
Author Affiliations and Target Audience	• **Decision-Maker**: defined as governments or health authorities making health care and policy decisions or HTA organizations directly funded by governments or health authorities to support decision-making. • **Industry:** pharmaceutical or health technology manufacturers with a commercial interest in products that are the subject of economic evaluations, including those preparing submissions and conducting economic evaluations for the purposes of obtaining funding and reimbursement • **Researchers**: economists or other researchers affiliated with academic or other independent research organizations or professional organizations
Presentation Formats	Information on a specific presentation format or details on how to present information must be provided. For example, formats could have included: diagrams, graphs, tables, lists, detailed descriptions, types of metrics, referencing practices, appendices, sample formats or adjectives describing presentation (e.g. detailed, summary, clearly). Only stating that information should be provided/reported was insufficient for it to be defined and extracted as a presentation format.
Data Sources	Information on how to present information on original sources and clinical or economic data included in the model.
Economic Model	Information on how to present model description, model structure, modeling methods and limitations, transition states and probabilities and model assumptions
Base Case Results	Information on how to present metrics and key concepts to emphasize in results.
Sensitivity Analyses	Information on how to present sensitivity analyses, with specific attention to differences between deterministic and probabilistic sensitivity analyses

***** As defined in the ISPOR guideline database (18, 20)

SUPPLEMENTARY MATERIALS – RESULTS

Supplementary Table 3. Additional Guideline Characteristics

Country	Type of Guideline	Purpose of Guidance
S. Africa 2013	Published PE Recommendations	To describe the process to be followed for PE applications and the criteria for medicines which require submissions; To create a forum which provides independent and objective review of the value of medicines
S. Korea 2013	PE Guidelines	To provide guidance on economic submissions for reimbursement in Korea
Norway 2012	PE Guidelines	Improve the basis for decision making as regards granting of reimbursement for drug expenses in the National Insurance Scheme and determine whether the cost of a medicine is in reasonable proportion to the therapeutic value and the costs associated with alternative treatment(s).
China 2011	Published PE Recommendations	Provide scientific basis for researchers to conduct economic analysis of medicines for pricing and reimbursement policy making in China
Ireland 2010	PE Guidelines	To promote the production of HTAs that are timely, reliable, consistent and relevant to the needs of decision makers and key stakeholders
Israel 2010	Submission Guidelines	For submission for a listing of new drug products
Russia 2010	Published PE Recommendations	Provide methodological and reporting guidelines for clinical and economic evaluation of pharmaceuticals submitted for inclusion in reimbursement
Brazil 2009	PE Guidelines	Contribute to the standardization of the local studies in order to improve transparency, review and reproduction of them.
Finland 2009	Submission Guidelines	Assist those making a submission to the Pharmaceuticals Pricing Board.
Germany 2009	PE Guidelines	Describe a method to be used for the health economic evaluations generated by the Institute for Quality and Efficiency in Health Care (IQWiG). Guidance to prepare reports on health economic evaluations that the contracting bodies have commissioned IQWiG to undertake.
Spain 2009	Published PE Recommendations	Provide methodological and reporting guidelines for PE evaluations and evaluation of health technologies
USA 2009	Published PE Recommendations	Improve the timeliness, scope, quality and relevance of information available to a health system's evaluators and ultimately to its P&T Committees. Streamline the data acquisition and review process for health system staff pharmacists. Support the informed selection of pharmaceuticals, biologics and vaccines by: a) Standardizing and communicating product and supporting program information requirements; b) Requiring projections of product impact on both the organization and its enrolled patient population; c) Requesting information on the value of products; and d) Making evidence and rationale supporting all choice(s) more clear, transparent and evaluable by decision makers.

Country	Type of Guideline	Purpose of Guidance
Australia 2008	Submission Guidelines	Provide manufacturers with guidance to prepare the clinical and economic data for submissions to the PBAC, who make recommendations on the suitability of drug products for subsidy by the Australian Government.
Belgium 2008	PE Guidelines	Provide methodological and reporting guidelines for PE evaluations
England & Wales 2008	Submission Guidelines	Provide an overview of the principles and methods of HTA and appraisal within the context of the NICE appraisal process. Provides guidance for all organisations considering submitting evidence to the NICE technology appraisal programme
México 2008	PE Guidelines	Provide methodological and reporting guidelines for new drug PE appraisals
Slovakia 2008	PE Guidelines	Provide methodological and reporting guidelines for PE evaluations
Thailand 2008	Submission Guidelines	Provide methodological and reporting guidelines for economic evaluations of health technology
New Zealand 2007	PE Guidelines	Secure the best possible health outcomes from within the funding provided.
Poland 2007	Submission Guidelines	Provide methodological and reporting guidelines for HTA evaluations of medical technologies
Scotland 2007	Submission Guidelines	Assist submission to Scottish Medicines Consortium
Canada 2006	PE Guidelines	Inform decision-making regarding the appropriateness and availability of health care interventions including drugs. Direct study design and provide standardized and reliable information to the target audience. Provide a template for final reports.
Taiwan 2006	PE Guidelines	Provide stakeholders a guide to conduct, evaluate and report a PE study
Netherlands 2004	PE Guidelines	Provide guidance for conduct of CE studies for reimbursement issues in order to improve their quality, credibility and comparability
Cuba 2003	PE Guidelines	Provide recommendations for conduct of studies to yield transparent, comparable, valid and credible results and guide researchers on how to complete an economic evaluation.
France 2003	PE Guidelines	Provide methodological and reporting guidelines for PE evaluations.
Sweden 2003	PE Guidelines	Provide criteria to be met for recognition of medical technologies/services as reimbursable
Baltic States 2002	PE Guidelines	For drug reimbursement and other state funding decisions in Latvia, Lithuania and Estonia
Hungary 2002	Published PE Recommendations	To support the decision-making process with sound CE data and good quality economic study.
Italy 2001	Published PE Recommendations	For pricing and reimbursement

Supplementary Table 4. Quality of Economic Guidelines as Assessed by modified AGREEII Instrument*

Guideline	Scope	Stakeholder Involvement	Rigour	Clarity	Applicability	Editorial Independence	Comments
S. Korea 2013			x		x	x	**Strengths:** clear objectives and audience; led by decision-makers with consultations with academics, industry, public; review of literature and international guidelines conducted with focus on new updates and revisions; attempts to reach consensus on recommendations through mtgs; recommendations clear and identifiable; accompanying tools; COI reported **Limits:** process for final agreement on recommendations unclear; no criteria for assessing implementation; funding not reported
S. Africa 2013	x	x	x		x	x	**Strengths:** clear objectives; scientific references; recommendations clear and identifiable; accompanying tools; funded by decision-maker **Limits:** audience unclear; no description of methodological development; stakeholder involvement not described; selection of evidence and process for agreement on final recommendations unclear; no criteria for assessing implementation; author affiliations, funding and COI not reported.
Norway 2012		x	x		x	x	**Strengths:** objective and audience clear; working group led by decision-makers and consultations on guidelines occurred; recommendations clear and identifiable. **Limits:** stakeholders involved in working group and consultations not reported; no systematic search for evidence; process for agreement on final recommendations unclear; no accompanying tools; no process for updating; no criteria for assessing implementation; funding and COI not reported
China 2011		x	x		x	x	**Strengths:** objectives and audience clear; scientific references provided; clear recommendations; recommendations clear and identifiable; mechanisms to dynamically update and revise; accompanying tools in appendix **Limits:** stakeholder involvement unknown; no description of methodological development; process for agreement on final recommendations unclear; no criteria for assessing implementation; funding and COI not reported.

Ireland 2010			x		x	x	**Strengths:** objective and audience clear; developed by decision-makers with methods experts; public consultation with stakeholders; scientific references provided; clear recommendations; tools in appendices **Limits:** no systematic search for evidence; process for agreement on final recommendations unclear; no process for updating; no criteria for assessing implementation; funding and COI unclear
Israel 2010	x	x	x		x	x	**Strengths:** specific recommendations identified **Limits:** objective and audience not specifically described; no indication of stakeholder involvement; no description of methodological development or references to evidence; no accompanying tools; no process for updating; no criteria for assessing implementation; funding and COI not reported.
Russia 2010		x	x		x	x	**Strengths:** objective and audience clear; authors and affiliations identified; recommendations clear and identifiable; tools in appendices. **Limits:** no indication of stakeholder involvement; no description of methodological development or scientific references; process for agreement on final recommendations unclear; no process for updating; no criteria for assessing implementation; funding and COI not reported.
Brazil 2009			x			x	**Strengths:** clear objectives and audience; identifies sources searched (journals, texts HTA guidelines); working group, workshops and public consultation during development; recommendations clear and identifiable; to be evaluated in 2 years after more economic studies apply them; future updates planned; funded by decision-maker; author affiliations reported. **Limits:** criteria for selecting evidence and process for agreement on final recommendations unclear; COI not reported.
Finland 2009	x	x	x		x	x	**Strengths:** specific recommendations identified **Limits:** objective and audience not specifically described; no indication of stakeholder involvement; no description of methodological development or references to scientific evidence; no accompanying tools; no process for updating; no criteria for assessing implementation; funding and COI not reported.
Germany 2009		x	x		x	x	**Strengths:** objectives and audience clear; developed by decision-makers through consultations with advisory boards and international experts; scientific references provided; clear recommendations. **Limits:** industry stakeholders not engaged; no systematic search for evidence; process for agreement on final recommendations unclear; no process for updating; no criteria for assessing implementation; no accompanying tools; funding and COI unclear

Spain 2009			x		x	x	**Strengths:** objective and audience clear; multi-stakeholder working group and review of drafts; review of international guidelines; scientific references provided; suggestions for inclusion debated by full working group; recommendations clear; additional details on development available on request or through other Spanish sources; funded by decision-maker **Limits:** no accompanying tools; no process for updating; no criteria for assessing implementation; COI not reported
USA 2009			x		x	x	**Strengths:** clear objectives and audience; multi-stakeholder development groups; opportunities for public comment provided; general scientific references provided; recommendations clear and identifiable; accompanying tools in appendices; **Limits:** no description of methodological development; process for agreement on final recommendations unclear; no process for updating; no criteria for assessing implementation; funding and COI not reported.
Belgium 2008			x		x	x	**Strengths:** evidence on which it was based is clear (select guidelines) and no new knowledge from other guidelines identified. Pilot study to assess implementation. COI clearly reported. Funded by decision-makers. **Limits:** No systematic search for evidence, clear criteria for evidence selection or strengths/limitations of evidence. No clear process for formulating recommendations or updating guidelines. No ongoing process to monitor implementation.
Australia 2008			x		x	x	**Strengths:** clear audience and objectives; stakeholders involved in development clear recommendations; a living documented to be updated regularly; funded by decision-makers. **Limits:** No systematic search for evidence, unclear process for formulating recommendations or criteria for monitoring implementation. COI not reported.
England & Wales 2008			x		x	x	**Strengths:** objective and audience clear; multi-stakeholder working party; accompanying tools; recommendations clear and identifiable; methods guidance regularly reviewed; accompanying tools and guides **Limits:** no description of methodological development or scientific references; process for agreement on final recommendations unclear; no criteria for assessing implementation; funding and COI not reported.
México 2008			x			x	**Strengths:** objective and audience clear; multi-stakeholder working group; drafts reviewed by stakeholders; based on literature review and international guidelines; scientific references provided; recommendations clear and identifiable; tools available; to be updated periodically; funded by decision-maker; author affiliations reported **Limits:** process for agreement on final recommendations unclear; COI not reported

Slovak 2008	x	x	x		x	x	**Strengths:** clear objective; identified basis of recommendations (expertise); recommendations clear and identifiable **Limits:** no description of methodological development or scientific references; process for agreement on final recommendations unclear; no criteria for assessing implementation; no accompanying tools; author affiliations, funding and COI not reported.
Thailand 2008			x			x	**Strengths:** clear objectives and audience; consultations with stakeholders; based on review of literature and existing guidelines; scientific references provided; selection criteria identified; clear recommendations; identify implementation barriers; accompanying tools; author affiliations reported; decision-maker funding **Limits:** subjective selection criteria; process for agreement on final recommendations unclear; no criteria for assessing implementation; COI not reported
New Zealand 2007					x	x	**Strengths:** objective and audience clear; led by decision-makers with wide multi-disciplinary stakeholder consultation; review of literature and international guidelines; consensus document outline decisions; clear recommendations; tools in appendices; living document to be updated over time. **Limits:** process for agreement on final recommendations unclear; no criteria for assessing implementation; funding and COI unclear
Poland 2007		x	x		x	x	**Strengths:** guideline task force identified; scientific references provided; dissent with and responsibility for final document and recommendations identified; recommendations clear and identifiable; funded by decision-makers **Limits:** no indication of stakeholder involvement; no description of methodological development; no accompanying tools; no process for updating; no criteria for assessing implementation; COI not reported
Scotland 2007		x	x		x	x	**Strengths:** objective and audience clear; references to other guidance (e.g. NICE); recommendations clear and identifiable; **Limits:** no indication of stakeholder involvement; no description of methodological development or scientific references; process for agreement on final recommendations unclear; no process for updating; no criteria for assessing implementation; no accompanying tools; funding and COI not reported.
Canada 2006		x	x		x	x	**Strengths:** clear recommendations were made by consensus. Funded by decision-makers. **Limits:** Manufacturers key stakeholders but limited involvement in guideline development. No systematic search for evidence, evidence selection or strengths/limitations of evidence. No process for updating guideline. No criteria for assessing guideline implementation

Taiwan 2006		x	x		x	x	**Strengths:** clear objective and audience; scientific references provided; clear recommendations; accompanying tools. **Limits:** stakeholder involvement unknown; no description of methodological development; process for agreement on final recommendations unclear; implementation issues not identified; no criteria for assessing implementation; no process for updating; funding, authors and COI not reported.
Netherlands 2004		x	x		x	x	**Strengths:** objective and audience clear; recommendations clear and identifiable; accompanying tools; steering committee established; scientific references provided but not exhaustive; refers to additional report elaborating on methods for updating guidelines. **Limits:** additional report on guideline methods inaccessible; affiliations of steering committee not reported; no indication of stakeholder involvement; no description of methodological development or scientific references; process for agreement on final recommendations unclear; no process for updating; no criteria for assessing implementation; no accompanying tools; funding and COI not reported.
Cuba 2003			x			x	**Strengths:** clear objectives and audience; based on literature search and global trends; stakeholders involved through working group with economists and expert review; recommendations clear and identifiable; 2 year evaluation based on experiences with guidelines; future update planned; independent funding; author affiliations reported. **Limits:** criteria for selecting evidence and process for agreement on final recommendations unclear; COI not reported.
France 2003			x		x	x	**Strengths:** clear scope; specific recommendations identified; references provided; developed by economists with key stakeholders engaged; 'spirit of consensus' applied but not achieved; tools provided in appendices **Limits:** no systematic search for evidence; no process for updating; no criteria for assessing implementation; funding from industry; COI not clearly reported.
Sweden 2003		x	x		x	x	**Strengths:** clear objective and audience; recommendations clear **Limits:** no description of methodological development or scientific references; stakeholder involvement unclear; process for selecting evidence or agreement on final recommendations unclear; no process for updating; no criteria for assessing implementation; no accompanying tools; funding and COI not reported.
Baltic 2002	x	x	x		x	x	**Strengths:** Clear recommendations; funded by decision-makers. **Limits:** Methodological question not clearly specified. Involvement of decision-makers, manufacturers and economists unclear. No systematic search for evidence, unclear process for formulating recommendations or no process for updating guideline. No criteria for assessing guideline implementation. COI unclear.

Hungary 2002			x		x	x	**Strengths:** clear objective and audience; stakeholders involved in development and consultation; based on search of international guidelines; scientific references provided; clear recommendations; method of making recommendations reported (consensus); plan to update as continuous improvement; accompanying tools; funded by decision-maker **Limits:** systematic search for and selection of evidence unclear; recommendations made by consensus; no criteria for assessing implementation; COI unclear
Italy 2001	x		x		x	x	**Strengths:** objective and audience clear; large multi-stakeholder development group initiated by economists; considered contributions from literature; scientific references provided; recommendations clear and identifiable. **Limits:** no systematic search for evidence; process for agreement on final recommendations unclear; no accompanying tools; no process for updating; no criteria for assessing implementation; funding and COI unclear.

'x' – indicates an issue was identified during critical appraisal of the guideline

* AGREEII was modified to assess methodological guidelines rather than clinical practice guidelines, as described in the Methods section of the Supplementary Materials.

Supplementary Table 5. Advice on Presenting Economic Evaluations to Non-Economists

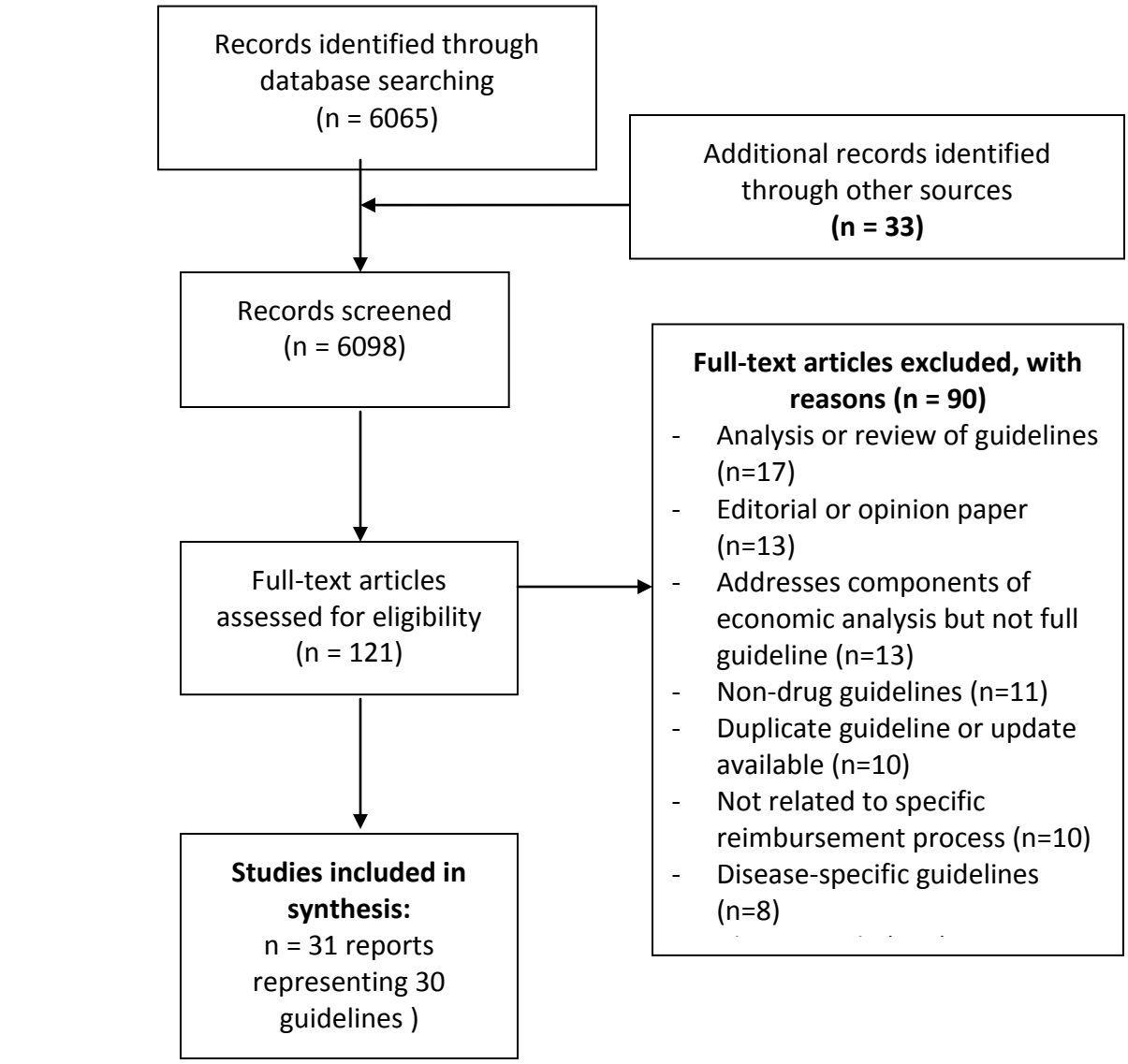
Guideline	Advice on Presenting Economic Evaluations to Non-Economists
USA 2009	• "All calculations should be explained in a simple straightforward manner to allow a non-health economist to comprehend the analysis." • Regarding disaggregated results: "Data presented in this format are more easily understood and interpreted by healthcare system formulary committees"
Brazil 2009	• "Presentation of the analysis in a transparent manner is a prerequisite for those who take the decisions about the technologies and interventions can judge the generalizability of the results." • "A 3-page executive summary must be written in a context accessible to a non-technical reader."
Mexico 2008	• "The executive summary and conclusions should be written so that they can be understood by a non-specialist reader using non-technical terms as far as possible." • "The executive summary can occupy a maximum of two pages and should be written in non-technical language"
Thailand 2008	• "Graphical presentations of the study results would be helpful for general readers"
Canada 2006	• "The Executive Summary and Conclusions should be written so that they can be understood by a non-technical reader"

Supplementary Table 6. Less Frequent (<20%) Reporting Recommendations in Economic Guidelines

Reporting Element	Reporting Recommendations	n (%) Guidelines, N=30
Perspective	End-user of analysis	1 (3)
	Patient	1 (3)
Target Population	No specific guidance	4 (13)
	May include off-label uses in some circumstances	3 (10)
	Population to be treated	2 (7)
	Consistent with clinical data	1 (3)
	Age, sex, comorbidities defined	1 (3)
	Insured and uninsured populations	1 (3)
	Could be subgroup within indication	1 (3)
	As commissioned	1 (3)
Subgroup Analyses	When defined a priori based on biologic plausibility and clinical rationale	1 (3)
	If defined a priori in clinical trial and statistical power adequate	1 (3)
	Subgroups with different costs or efficacy	1 (3)
	Gender, age, severity, risk level	1 (3)
	If clinical, diagnostic or genetic markers to identify subpopulations	1 (3)
	Treatment regimens, disease groups, symptoms, disease severity, patient characteristics	1 (3)
	When treatment options available and clinical practice varies	1 (3)
	As necessary	1 (3)
	Unknown	2 (7)
Time Horizon	Trial duration considered as primary or sensitivity analysis	4 (13)
	As related to study purpose or relevance to decision-maker	3 (10)
	Possibility of lifetime horizon considered	3 (10)
	Specific time horizon suggested as part of analysis (1-year, 5-years)	1 (3)
	Unknown	3 (10)
Assumptions Required	Assumptions Clearly Stated or Justified	4 (13)
	Try to Avoid Assumptions	1 (3)
	Unknown	5 (17)
Modelling Permitted	Yes, based on observational data	5 (17)
	In some situations (e.g., trial length insufficient for disease duration)	3 (10)
	Yes, technical modeling approaches suggested	2 (7)
	Yes, transparent and well-described	2 (7)
	Yes, validity carefully evaluated	1 (3)
Preferred Outcome Measures	Willingness to Pay	4 (13)
	Toxicity or adverse events	3 (10)
	Unknown	1 (3)

Reporting Element	Reporting Recommendations	n (%) Guidelines, N=30
Sensitivity Analyses - Parameters	Most influential	5 (17)
	All variables	4 (13)
	Main assumptions	4 (13)
	Variables sensitive in one-way	2 (7)
	Identified by experts	1 (3)
	Unknown	1 (3)
Sensitivity Analyses - Range	Range should be justified	4 (13)
	Expert opinion	1 (3)
	Values from literature	1 (3)
	Author decisions	1 (3)
	Unknown	1 (3)
Sensitivity Analyses - Methods	Two way	5 (17)
	Unknown	5 (17)

Supplementary Figure 1. Results of Systematic Review Search.



CHAPTER 5.

KNOWLEDGE, BARRIERS, FACILITATORS AND APPROACHES TO COMMUNICATING NETWORK META-ANALYSES: A SYSTEMATIC REVIEW

CHAPTER OVERVIEW

Purpose and Relationship to Other Manuscripts:

One of the goals of this thesis is to understand barriers and facilitators to communicating NMAs from the perspective of end-users. Therefore, the objective of the third manuscript is to systematically review knowledge, barriers, facilitators and approaches to communicating and understanding network meta-analyses (NMAs). This work will be used to identify the informational needs and optimal presentation of NMAs for non-technical audiences, i.e. knowledge users. Important knowledge users include policymakers who use NMAs to inform policy decisions and clinicians who use NMAs to inform treatment decisions. This manuscript is similar to the systematic review related to communicating economic information that is presented in Chapter 3. However, unlike the economic systematic review, in which the majority of studies focused on policymakers, the majority of the studies identified in the NMA systematic review provide perspectives from NMA researchers and not NMA end-users. Therefore, there is a gap in the literature with respect to end-users' preferences (e.g. clinicians, policymakers) for how to present and communicate NMAs.

Supporting Material:

None

Manuscript Status:

Submitted for review and awaiting a decision. All information in this chapter, including the supplementary materials, has been submitted to the journal.

Authors: Shannon M Sullivan (SMS), Doug Coyle (DC), George Wells (GW)

Author Roles (SMS, DC, GW) and Contributions:

SMS, DC, GW were responsible for the study conception.
SMS, GW were responsible for the search for trials.
SMS, GW were responsible for the selection of trials.
SMS, GW were responsible for the data extraction.
SMS, GW were responsible for the data analysis.

SMS, DC, GW were responsible for in the interpretation of results.
SMS, GW were responsible for drafting the manuscript.

All authors (SMS, DC, GW) provided comments on drafts and revisions to the manuscript, agree with the manuscript's results and conclusions, and read and approved the final manuscript.

TITLE:

Approaches to Communicating and Presenting Network Meta-Analyses: A Systematic Review

SHORT TITLE:

Communicating NMAs

AUTHORS:

Sullivan SM,[1] Coyle D,[2] Wells GA[1,2]

AFFILIATIONS:

1 University of Ottawa Heart Institute, Cardiovascular Research Methods Centre, Ottawa, ON

2 Department of Epidemiology and Community Medicine, University of Ottawa, Ottawa, ON

ABSTRACT

Objective: To conduct a systematic review of knowledge, barriers, facilitators and approaches to understanding and using network meta-analyses (NMAs).

Methods: Electronic databases and supplementary sources were searched for studies that reported approaches to presenting or communicating NMAs, understanding and knowledge of NMAs and barriers and facilitators to using NMAs. Frequencies were tabulated and qualitative data synthesis conducted as needed.

Results: Sixteen included studies identified many graphical or tabular options for communicating NMA data sources, assumptions and results. Only one study developed approaches targeted to different end-users. NMA knowledge was variable among researchers; inadequate conduct and misinterpretations of published NMAs occurred. Barriers to using NMAs included the large volume of data to synthesize, challenges interpreting and presenting data, evolving methodologies and limited expertise. Facilitators included knowing the validity of NMA assumptions; leveraging the methodological strengths of NMAs; and having consistent reporting standards.

Conclusions: Identified presentation formats may assist researchers but could be insufficient for non-technical audiences to understand what NMA results mean and how to use them. Tailored knowledge translation approaches could enhance understandability and accessibility of NMAs for policymakers and healthcare providers while increasing the transparency and legitimacy of policy and treatment decisions informed by NMAs.

INTRODUCTION

Transparency is a key principle underlying fair and legitimate health technology assessment (HTA) processes and related policy decisions (Daniels & Sabin, 1997). Transparency requires not only providing sufficient information, but also providing information in an accessible and understandable format for end-users. This is especially relevant when complex methods such as network meta-analyses (NMA) form the basis of a HTA.

In the absence of head-to-head trials of relevant comparators, NMAs frequently inform cost-effectiveness evaluations and therapeutic or drug class reviews, which are used in HTAs (Cooper et al., 2011). One frequently cited concern with NMAs is that the complex statistical methods do not permit end-users to understand how the results were obtained or if they are valid (Salanti et al. 2011; Mills et al., 2013). Providing simple but accurate explanations of NMAs for audiences such as policymakers and clinicians can be challenging. This is further complicated by the variable expertise among researchers in conducting and interpreting NMAs (Abdelhamid et al., 2012, Coleman et al., 2012) and the rapidly evolving developments of NMA methods. (Chaimani et al., 2013; Salanti et al., 2014; Veroniki et al., 2013)

In order to develop more understandable and relevant formats for presenting NMAs to non-technical end-users such as policymakers and clinicians, a review of current approaches being taken to communicate NMAs to different audiences is required. This could help identify best practices for effectively communicating NMAs to various audiences and gain insight into end-users' preferences for receiving NMA information. Understanding end-users' knowledge of NMA and existing barriers and facilitators to using and communicating NMAs is supportive information that can be used to inform the development of knowledge translation tools and tailor the presentation of NMAs for different audiences (Straus et al., 2013). A systematic approach to knowledge translation of NMAs could also improve the accessibility of HTAs that are based on NMAs, thereby enhancing the legitimacy of and confidence in health policy decision-making processes (Daniels & Sabin, 1997; Dhalla & Laupaucis 2008)

The objective of this study was to systematically review knowledge and understanding of different NMA end-users; as well as barriers, facilitators and approaches to communicating NMAs. These findings can inform the optimal presentation of NMAs for non-technical end-users and the development of NMA-specific knowledge translation tools.

METHODS

Review Protocol: A systematic review was conducted following Cochrane methodology (Cochrane Collaboration, 2014). All empirical study designs with a defined methodology were included (e.g., systematic reviews, randomized controlled trials (RCTs), observational studies, interviews, focus groups or surveys, case studies). The population of interest was any adult population conducting or using NMAs; this included researchers, policymakers, practicing clinicians, adult patients or other members of the public. Researchers were included in the population of interest as NMAs are a novel approach and sometimes considered complex even within the research community (Abdelhamid et al., 2012). The intervention and comparator were approaches, interventions and tools used to convey NMAs or used to identify barriers and facilitators of NMAs. Outcomes of interest were: approaches or preferences for communicating or presenting NMAs; individuals' knowledge or understanding of NMAs; barriers to the use or understanding of NMAs; facilitators encouraging use or understanding of NMAs.

Search Strategy: Databases searched included Medline (1996 to June 2014), EMBASE (1980 to June 2014), and the Cochrane Database of Systematic Reviews, using the earlier data limit of 2000 but no language restriction. The search concepts were 'indirect comparisons', 'network meta-analyses and 'multiple or mixed treatment comparisons'. Grey literature was searched for unpublished reports using the CADTH Grey Matters checklist as a guide and supplemented by other unpublished sources (CADTH, 2013). Studies were also obtained through hand searching selected journals, reviewing reference lists of potentially relevant studies and suggestions from NMA experts.

Study Selection: Citations were screened for relevance by one review author based on the title and abstract of identified articles. Two review authors independently reviewed the full text of potentially relevant articles to assess exclusion or inclusion.

Data Extraction: The following data were extracted from guidelines:

1) **Approaches to Communicating NMAs**: information on NMA formatting approaches and recommended options for presenting NMA data sources (evidence networks, individual trial characteristics, critical appraisals), assumptions (heterogeneity, inconsistency, other assumptions) and results (effect sizes, rankings, uncertainty).
2) **Knowledge**: self-reported or measured knowledge; or frequency of appropriate conduct and interpretation of NMAs.
3) **Barriers and Facilitators**: self-reported barriers and facilitators; or author perspectives on NMA barriers and facilitators if other protocol-outcomes were reported in the study.

To be extracted, information was required on the approach or format for presenting and communicating information, not just that the information be provided or reported (i.e., focusing on 'how' to present information not just 'what' to report). Similarly, methodological approaches for obtaining results were not extracted; to be extracted information was required on 'how' to present the results, not what approach was taken to conduct the analysis and obtain the results.

One author extracted data, which was verified by a second author. Disagreements were resolved by a third reviewer. Study quality was assessed using AMSTAR for systematic reviews (Shea et al., 2009). Critical appraisal of case reports and qualitative studies was informed by recommendations from the Cochrane Collaboration (Hannes et al., 2011) and checklists developed by Greenhalgh et al. (2004).

Data Analysis: The frequency with which outcomes and communication approaches were reported was tabulated across studies. When study-level frequencies were reported, they were based on the number of unique studies, not total number of reports. Due to the heterogeneity of data, descriptions of findings from individual studies are frequently provided. Data on barriers and facilitators were synthesized qualitatively (Cochrane Collaboration, 2014; Thomas et al., 2008). Descriptive themes were identified following a review of the primary data in order to categorize concepts that were reported in various ways in the primary studies.

RESULTS

Study Characteristics

Of the 3498 citations identified (see Figure 1), 16 unique studies (based on 17 reports) were included: 7 were case studies (43%), 7 were reviews (43%), 1 was an online survey (6%) and 1 was a mixed methods study (6%) that included a review and focus group. (Abdelhamid et al., 2012; Bafeta et al., 2013; Bafeta et al., 2014; Chaimani et al., 2013; Coleman et al., 2012; Cope et al., 2013; Donegan et al., 2010; GRADE Working Group, 2013; Hutton et al., 2014; Nikolakopoulou et al., 2014; Ohlssen et al., 2014; Salanti et al., 2011; Salanti et al., 2014; Sobieraj et al., 2013; Song et al., 2009; Tan et al., 2013; Tan et al., 2014) The mixed methods study was available in the grey literature with some components of the review published. (Coleman et al., 2012; Sobieraj et al., 2013)

INSERT FIGURE 1. Results of the Literature Search

Among the reviews, different selection criteria were used, however, the number of included NMAs or IDCs ranged from 19 (Tan et al., 2013) to 186 (Nikolakopoulou et al., 2014). Studies reporting on case studies sometimes used one, two or three examples and the NMAs used were often selected for their complexity and ability to demonstrate different NMA concepts. The size of the networks ranged from a minimum of 4 therapies based on 10 studies (Cope et al., 2013) to a maximum of 11 therapies based on 40 studies (GRADE Working Group, 2013).

The two qualitative studies included participants who were researchers, either Cochrane review authors (Abdelhamid et al., 2012) or authors of Bayesian mixed treatment comparisons (Coleman et al., 2012). Among the Cochrane review authors, 54% considered themselves researchers, 27% clinicians and 19% both. In the focus group of 9 authors of Bayesian NMAs, 67% were from academic settings.

Few studies specified the target audience, but given their publication in academic journals, the intended audience is likely individuals with an academic or research background. However, a number of studies (n=14, 82%) acknowledged that policymakers (n=10, 59%), clinicians (n=9, 53%) and patients or the public (n=3, 18%) could be end-users of NMA as NMAs are used to inform policy or treatment decisions.

The first study was published in 2009 (Song et al.) and publications continued over the six-year period to the present day. Study authors were primarily from Europe (n=6, 35%), the UK (n=5, 29%) and North America (n=5, 29%), with one report authored by an international working group (GRADE Working Group, 2013).

INSERT TABLE 1. Characteristics of Included Studies

Study quality was assessed based on study design and was generally found to be moderate when considering the overall body of evidence from the 17 included studies (see Supplemental Tables 1A, 1B, 1C). Common limitations of reviews were limitations on search strategies and assessment of biases. Common limitations of qualitative studies were not considering researcher bias when interpreting results. Common limitations of case reports were insufficient rationale for selection of the case and the development of only one case which limited generalizations.

Studies were heterogeneous with respect to study objective (see Supplementary Table 2) although most were conducted with the intent of informing improvements in NMA conduct, reporting or interpretation. When considering the type of NMA information provided, most studies provided details on NMA reporting characteristics (n=7/16, 41%), NMA methodological

characteristics (n=6/16, 35%) or NMA presentation formats (n=7/16, 41%). Two studies (12%) provided information on approaches to NMA quality assessment, two studies (12%) provided information on NMA training needs and one study (6%) provided information on NMA network characteristics.

Based on the protocol-outcomes defined in this systematic review, almost all studies provided information on approaches to communicating NMAs (n=15/16, 94%). All studies provided information on barriers and facilitators either through formal surveys (n=2/16, 12%) or that could be extracted from authors perspectives described in the included studies. Seven studies provided information on NMA knowledge either through self-reported knowledge (n=2/16, 12%) or indirect measures such as the appropriate conduct and interpretation NMAs.

Formatting Options for Presenting NMAs

Eleven studies (n=11/16, 69%) explored different formatting options for presenting NMAs. Some provided details on how NMAs have been presented in the literature (Bafeta et al., 2014; Coleman et al., 2012; Tan et al., 2013) while others were case studies discussing advantages and disadvantages of different figures and tables to illustrate NMA principles (Chaimani et al., 2013; Cope et al., 2013; Salanti et al., 2011; Salanti et al., 2014; Tan et al., 2014). Most reviews noted that a mix of tables, figures and text were used to present NMAs (Bafeta et al., 2013; Hutton et al., 2014; Tan et al., 2013). The printed length of an NMA report was approximately 16 pages and the majority of articles also had a supplementary appendix where more details could be provided (Coleman et al., 2012). Five studies presented reasons for using different figures and tables, which were illustrated through an NMA case study. Different graphical formats were used support different NMA aspects including assessing assumptions, conducting quality assessments, reporting results and considering uncertainty. One study also emphasized the importance of incorporating graphical design features when creating tables and figures to facilitate interpretations (Tan et al., 2014). Key findings on formatting options are summarized in Table 2 with individual study details and specific approaches described in Supplementary Table 3.

INSERT TABLE 2. Key Findings Related to Formatting Options for Communicating NMAs

Approaches to Communicating NMA Data Sources

Eight studies (n=8/16, 50%) provided information on approaches to communicating data sources (See Table 3). More study-level details are reported in supplementary materials (see Supplementary Table 4)

INSERT TABLE 3. Key Findings on Approaches to Communicating NMA Data Sources

Evidence Networks

Three approaches to presenting evidence networks were identified: network diagrams (Chaimani et al., 2013; Hutton et al., 2014; Tan et al., 2013; Salanti et al., 2014) , evidence network tables (Tan et al., 2013) and contribution plots (Chaimani et al., 2013; Salanti et al., 2014). Network diagrams are graphical schematics representing the available evidence and connections between the pieces of evidence. Possible features that could be included in the network diagram included sample sizes, the number of studies per comparison, the presence of multi-arm studies, effect modifiers, risk of bias and key trial characteristics. Evidence network tables layout the available evidence, similar to network diagrams but do this in a tabular format rather than a figure. However, they were considered feasible only if three-arm trials were not part of the network. Contribution plots graphically represent how much direct evidence and how much indirect evidence contributes to each of the NMA estimates. Contribution plots were considered valuable as they can identify how much direct and indirect evidence contributes to each NMA estimate as well as the overall network. These approaches were supported by one literature review noting that 74% provided a network diagram (Bafeta et al., 2014) and another review noting that 21% provided a network diagram while 58% provided evidence network tables and 11% provided both (Tan et al., 2013).

Trial Characteristics

Two approaches to presenting individual trial characteristics were identified: a tabular format (Hutton et al., 2014) or highlighting key characteristics in the network diagram (Chaimani et al., 2013). Specific recommendations on which characteristics to include were not made but the importance of highlighting effect modifiers was emphasized. The inclusion of information on individual trial characteristics in NMAs has varied with three reviews presenting a range of estimates (21% in Bafeta et al., 2013; 62% in Coleman et al., 2012; and 88% in Donegan et al., 2010).

Critical Appraisal

Three approaches to critical appraisal, risk of bias (RoB) assessments, publication bias assessment and GRADE-related assessments (GRADE Working Group 2013; Salanti et al., 2014), were identified with different formats identified to support each approach. Four approaches to presenting RoB in NMAs were identified: a tabular format with other trial-level characteristics (Hutton et al., 2014), as part of a network diagram with coloured lines representing RoB assessments for individual trials (Chaimani et al., 2013), through a pie chart showing the RoB contributions to each treatment comparison (Salanti et al., 2014) or with a bar graph showing

RoB contributions to network and direct estimates (Salanti et al., 2014). When conducting a GRADE-related quality assessment, two different table formats were suggested for presenting final conclusions, depending on how the GRADE framework was implemented (GRADE et al., 2013; Salanti et al., 2014). A number of figures were also identified as supportive of the quality assessments including contribution plots, RoB pie charts and RoB bar graphs and rankograms (Salanti et al., 2014). A comparison-adjusted funnel plot or a contour enhanced funnel plot were the two approaches suggested for assessing potential publication bias (Chaimani et al., 2014 Salanti et al., 2014). One review reported on the frequency of critical appraisal in NMAs, with 58% NMAs not reporting RoB assessments for individual studies and 84% not reporting methods for assessing publication bias (Bafeta et al., 2013).

Approaches to Communicating NMA Assumptions

Nine studies (n=9/16, 55%) provided information on approaches to communicating NMA assumptions, focusing primarily on heterogeneity and inconsistency (see Table 4). More study-level details are reported in supplementary materials (see Supplementary Table 4). No communication approaches were identified that were associated with other key NMA assumptions or methodological approaches. However, examples of different ways of reporting on the similarity assumption were summarized in one review (Donegan et al., 2014).

INSERT TABLE 4. Key Findings on Approaches to Communicating NMA Assumptions

Heterogeneity

Although many studies noted the importance of evaluating heterogeneity assumptions and suggested methods or measures for doing this (e.g. between-study variance, I^2 , Tau^2), only three general approaches to communicating heterogeneity were identified: tables, forest plots or network diagrams that include heterogeneity measures (GRADE Working Group, 2013; Salanti et al., 2014; Tan et al., 2014). In addition, two figures supportive of the assessment of heterogeneity and potential impact of effect modifiers were identified: forest plots including predictive intervals and network diagrams incorporating potential effect modifiers. Because heterogeneity is often assessed by looking at differences in individual studies and considering the potential impact on results, presenting information on trial characteristics (see previous section on Communicating NMA Data Sources) and trial results (see subsequent section on Communicating NMA Results) can provide important supportive tables and figures for this assessment. One review reporting on the assessment of heterogeneity in published NMAs found that only 32% evaluated heterogeneity in the NMA while 61% evaluated heterogeneity in the accompanying traditional pairwise meta-analysis estimates (Coleman et al., 2012).

Inconsistency

Two approaches to communicating inconsistency were identified: inconsistency plots (Chaimani et al., 2013) and conclusions as a component of quality assessment tables (Salanti et al., 2014). Chaimani et al., (2013) suggested an inconsistency plot, which would present inconsistency factors as a forest plot with confidence intervals displayed with an additional column reporting loop-specific heterogeneity estimates. Salanti et al., (2014) suggested that assessment of inconsistency should be presented in a table format that summarizes the quality of the NMA, along with other GRADE-related assessments. Plots displaying predictive intervals were also noted to be supportive of inconsistency assessments (Salanti et al., 2014). Across six reviews (Bafeta et al., 2013; Coleman et al., 2012; Donegan et al., 2010; Nikolakopoulou et al., 2014, Song et al., 2009; Tan et al., 2013), estimates of how frequently inconsistency was assessed and reported ranged from 11% (Tan et al., 2013) to 71% (Coleman et al., 2012).

Approaches to Communicating NMA Results

Twelve studies (n=12/16, 75%) provided information on communicating NMA results (see Table 5). More study-level details are reported in supplementary materials (see Supplementary Table 4)

INSERT TABLE 5. Key Findings on Approaches to Communicating NMA Results

Effect Estimates, Including Comparisons Between Direct and Indirect/Mixed Evidence

Eight approaches to presenting effect estimates, or comparing direct and indirect evidence, were identified: forest plots, (Bafeta et al., 2014; Chaimani et al., 2013; GRADE Working Group, 2013; Ohlssen et al., 2014; Salanti et al., 2011; Salanti et al., 2014; Tan et al., 2013; Tan et al., 2014) caterpillar plots, (Tan et al., 2013) matrix tables, (Tan et al., 2013, Bafeta et al., 2014) tables of relative or absolute effects, (Bafeta et al., 2014, Ohlssen et al., 2014; Tan et al., 2013) a summary forest plot matrix or summary forest plot table (Tan et al., 2014) or assessment of 'indirectness' incorporated into GRADE-related quality assessment tables (GRADE Working Group, 2013; Salanti et al., 2014). The summary forest plot table and matrix incorporated features from tables, forest plots and rankings and applied graphical design features to facilitate comparisons (e.g. shading, juxtapositioning of information) along with flexibility in what effect estimates are presented in order to tailor formats to end-users preferences (e.g. direct, indirect or mixed evidence; which treatment comparisons; absolute or relative measures, number of outcomes).

Hutton et al (2014) also noted the variety of effect estimates that were presented in NMAs and considered this could be standardized in reporting guidelines. For example, effect estimates can be presented for all comparisons in a network or only for some comparisons; estimates can be

presented from individual trials, traditional pairwise meta-analyses and network meta-analyses and estimates can be presented based on absolute or relative measures. In Tan et al., (2013), for those NMAs that reported comparative effectiveness, 61% reported all combinations of treatment comparisons, while 38% focused on select comparisons, either versus placebo or other active treatments. One review noted that 56% of NMAs reported direct results from individual trials (Bafeta et al., 2014). Two reviews found that 52% or 89% (Bafeta et al., 2014; Coleman et al., 2012) of NMAs reported direct results from traditional pairwise meta-analyses. One review reported that 64% NMAs explicitly compared direct and indirect evidence (Song et al., 2009) and another noted that 58% made a distinction between direct and indirect evidence (Donegan et al., 2010).

Rankings

Twelve approaches to communicating treatment rankings were identified. One approach was to present tables with the estimates probabilities and/or related metrics (e.g. SUCRA%) (Salanti et al., 2011; Salanti et al., 2014; Tan et al., 2013;). A number of figures presented rankings in different ways, including rankograms (Coleman et al., 2012; Cope et al., 2013; Salanti et al., 2011; Salanti et al., 2014, Tan et al., 2013; Tan et al., 2014), bar charts with the probability each treatment has a specific rank (Salanti et al., 20122), cumulative probability ranking plots (Bafeta et al., 2014; Chaimani et al., 20122; Salanti et al., 2011), plots of posterior distributions of rankings (Coleman et al., 2012; Ohlssen et al., 2014; Salanti et al., 2011) a plot of treatment ranking versus SUCRA % or MDS estimate (Chaimani et al., 2013), Other unique approaches to presenting rankings include a threshold plot based on an acceptable tolerance threshold and where each treatment ranking falls in relationship to it (Chaimani et al., 2011; Salanti et al., 2011) and a clustered ranking plot displaying multiple outcomes (Chaimani et al., 2011) .Three approaches to visually summarizing rankings alongside other results were found: summary forest plot matrix or summary forest plot table incorporating rankings and associated probabilities (Tan et al., 2014), a median ranking chart presenting simplified rankings for large networks (Tan et al., 2014). One study focused on developing an approach to communicating survival rankings and presented options based on time-independent or time-varying measures (bar charts vs. line graphs) with rankings represented through rankograms, SUCRA% or the probability a treatment is best (Cope et al., 2013).

One review noted that 62% NMAs reported treatment rankings and associated probabilities (Coleman et al., 2012). One review reported that 26% NMAs reported the probability each treatment was best (Tan et al., 2013), noting that most studies had provided information in tables, not a graphical format. Another review noted that 57% of Bayesian NMAs provided treatment rankings, 44% provided the probability of each treatment being best and 8% presented absolute

or cumulative rankograms (Bafeta et al., 2014). It was noted that of the studies in Bafeta et al. (2014) that provided rankings, only 17% provided a measure of uncertainty in ranking; this was usually done with absolute or cumulative rankograms (78%) and the remainder used confidence or credible intervals (22%). The overview by Hutton et al., (2014) also discussed how frequently rankings were reported in the literature and proposed there should be reporting recommendations around treatment rankings and corresponding probabilities.

One study suggested that rankograms could be used to support a quality assessment of the treatment ranking estimates from an NMA (Salanti et al., 2014). In addition, graphics allowing assessment of risk of bias, inconsistency and publication bias (described in their respective sections of this paper) would support the quality assessment of rankings.

Uncertainty

Four approaches to communicating uncertainty were identified: confidence or credible intervals on forest plots or in tables (Bafeta et al., 2014; Chaimani et al., 2013; Coleman et al., 2012; GRADE Working Group, 2013; Hutton et al., 2014; Ohlssen et al., 2014, Salanti et al., 2011; Salanti et al., 2014; Tan et al., 2013; Tan et al., 2014), predictive intervals on forest plots or in tables (Chaimani et al., 2013; Salanti et al., 2011; Salanti et al., 2014; Tan et al., 2014). The frequency of using these approaches was reported in two reviews, finding that all included NMAs had presented either confidence or credible intervals (Coleman et al., 2012; Tan et al., 2013). The importance of presenting the uncertainty associated with rankings was also discussed in four studies and different approaches suggested (Bafeta et al., 2014; Chaimani et al., 2013; Salanti et al., 2011; Salanti et al., 2014). These approaches were the use of confidence intervals, cumulative probability ranking plots, plots of the posterior distributions of rankings or comparing rankings with the magnitude of effect estimates. One study also discussed uncertainty associated with ranking survival outcomes over time and discussed various methods of presenting survival rankings over time (Cope et al., 2013).

Individuals' Knowledge of NMAs

Seven studies (n=7/16, 41%) reported on knowledge and understanding of NMAs. Two studies formally surveyed researchers to gain insight on their knowledge of NMAs (Abdelhamid et al., 2012; Coleman et al., 2012). Five studies reported details on appropriate conduct, changes in conduct over time (Bafeta et al., 2013; Donegan et al., 2012; Nikolakopoulou et al., 2014, Song et al., 2009) and interpretation (Salanti et al., 2011) of NMAs, which provide an indication of individuals' NMA knowledge. Key findings are summarized below and reported in Table 6 with individual study level details provided in Supplementary Table 5.

INSERT TABLE 6. Key Findings Related to NMA Knowledge

In an on-line survey of 48 Cochrane reviewers (Abdelhamid et al., 2012) knowledge and acceptance of NMAs varied but 87% had at least heard about indirect comparisons and 76% would consider indirect evidence at least some, if not all of the time. However, 20% were not aware of NMA methodology and 23% did not have the expertise to conduct an NMA. In a focus group of 9 authors of Bayesian NMAs (Coleman et al., 2012), 89% were trained in NMA methods and 78% had conducted at least two NMAs. All participants agreed that combining indirect and direct evidence provides value and 67% considered this provides more precise estimates. When considering terminology, 55% participants thought 'NMA' was ambiguous but 44% thought 'MTC' was clearer, i.e., it was used unambiguously and consistently in the medical literature. Approximately 56% of respondents thought current NMA guidance is sufficient and 78% thought NMAs should include network diagrams. No clear trends were identified for questions about whether or not direct evidence takes precedence over indirect; if direct and indirect estimates should both always be considered and if comparisons not of direct relevance (e.g. placebo, older treatments) should be considered.

Four reviews considered if NMAs were adequately conducted, although different definitions of 'adequate' were applied across reviews. In one review, 24% NMAs used inappropriate statistical methods for assessing inconsistency (Nikolakopoulou et al., 2014). This review also that reporting of inconsistency increased from 58% in 2011 to 74% in 2012. Another review (Bafeta et al., 2013) found that when comparing NMA reporting with Cochrane meta-analysis standards, 99% were inadequate and methods related to searching and data collection were inadequate: 11% only searched one database, 20% only looked at published literature, 43% didn't duplicate study selection and 21% didn't duplicate data extraction. A review by Donegan et al., (2010) found that 5% of indirect comparisons (IDCs) used inadequate analysis methods that did not preserve randomization, 30% had inadequate methods for assessing heterogeneity, 59% had inadequate methods for investigating clinical or methodological homogeneity, if heterogeneity existed, 71% incorrectly analyzed 3-arm trials and 51% did not have a reasonable approach to assessing similarity assumptions. This review also noted the increase in the IDC publications over time, ranging from one in 1992 to seven in 2007, for a total of 41 IDCs overall. Similarly, a review by Song et al. (2009) found that there was inadequate handling of 3-arm trials and that of 88 IDCs, 13 used inadequate informal IDC approaches and 6 used naïve IDC approaches.

One study (Salanti et al., 2011) also discussed the incorrect interpretation of two different NMAs in the published literature and considered different presentation options that could have

prevented this; for example, use of 95% predictive intervals demonstrating uncertainty associated with effect estimates or graphics identifying uncertainty associated with treatment ranking claims.

Barriers and Facilitators to Using NMAs

All 16 studies provided some information on NMA barriers and facilitators. Two studies (Abdelhamid et al., 2012; Coleman et al., 2012) formally surveyed participants to obtain this information. In the remaining 14 studies, authors' perspectives on NMA barriers and/or facilitators were extracted from published papers and common themes noted.

When barriers were considered from all 16 studies, common themes emerged (See Figure 2A). Barriers identified in >25% studies, included: the large volume of data and studies to synthesize and interpret (50%), challenges interpreting NMAs (44%), challenges presenting NMAs (38%), evolving methods (38%) and lack of training or expertise (31%). Other themes that were identified included an inadequate evidence base, a lack of transparency in NMA reporting, low confidence in methods or acceptance of NMAs, general complexity of NMAs, lack of reporting guidance, time and resources required and inconsistent language used to describe NMAs.

When facilitators were considered from all 16 studies, common themes emerged (see Figure 2B). Facilitators identified in >25% studies, included: being able to asses assumptions (38%), methodological strengths of NMAs (38%) and having reporting standards (31%). Other themes that were identified included being able to understand the available evidence, knowing treatment rankings and their uncertainty, having critical appraisal standards and guidance for presentation of NMAs, increased transparency of NMAs, knowing the uncertainty in treatment rankings, acceptance of NMA methods, more NMA training, having a better evidence base and standard terminology to describe NMAs.

INSERT FIGURE 2. Most Frequently Reported Types of Barriers and Facilitators to Using and Communicating NMAs

Supplementary Table 6 provides more details on the individual barriers and facilitators reported in each category, including those barriers and facilitators identified in formal surveys of Cochrane reviews and authors of Bayesian NMAs.

DISCUSSION

Summary of Findings

Numerous graphical or tabular options for communicating data sources, assumptions and results were identified. The largest number of options was provided for communicating treatment rankings. However, approaches were also identified for evidence networks, individual study characteristics, critical appraisal approaches, heterogeneity and inconsistency assumptions, effect estimates and uncertainty. No studies evaluated preferences for different presentation options and only one study developed specific approaches for different NMA audiences. Knowledge of NMAs was variable among surveyed researchers; inadequate conduct and misinterpretations of published NMAs were sometimes reported. Common barriers to using NMAs included the large volume of data to synthesize, challenges interpreting and presenting data, evolving methodologies and a lack of expertise. Common facilitators included having the validity of NMA assumptions assessed; being able to take advantage of the methodological strengths of NMAs; and having consistent reporting standards. Knowledge, barriers and facilitators were reported from researchers' perspectives and no studies reported these outcomes among end-users such as clinicians and policymakers although many studies acknowledge these as key NMA audiences.

Comparison with Other Literature

Challenges interpreting NMAs have frequently been noted in the literature (Salanti et al., 2011, Cipriani et al., 2013, Mills et al., 2012). In addition, commonly identified limitations in conducting NMAs have also been noted (Song et al., 2009, Martin et al., 2013) including inadequate search methods, inappropriate pooling of heterogeneous data and suboptimal statistical methods. Some of the presentation approaches identified in this review may have been developed in response to some of these limitations. For example, the simplistic interpretations of treatment rankings that do not consider their associated uncertainty have often been noted (Salanti et al., 2011; Jansen et al., 2011). Numerous options for presenting treatment rankings and associated probabilities were identified in this review, perhaps, reflecting methodologists' attention to improving the presentation of ranking results in NMAs.

Knowledge of NMAs was found to be variable and inadequate conduct of NMAs was sometimes reported in the literature (Bafeta et al., 2013; Donegan et al., 2010; Nikolakopoulou et al., 2014; Song et al., 2009). In some cases, this was due to inadequate assessment of the required assumptions underlying NMAs. Clearly communicating NMA assumptions was also identified as a facilitator in this review. Therefore, approaches that are able to clearly communicate to end-users the underlying assumptions, their validity and the impact on results of these assumptions could

be of value. A recent review has documented the variety of methodological approaches for evaluating NMA assumptions and may be able to inform better presentation of this aspect of NMAs (Donegan et al., 2013). Other inadequate methodologies were that NMAs did not always incorporate standard systematic review or meta-analysis practices. This suggests that researchers may not always appreciate that a good quality NMA is founded on the same principles as traditional meta-analyses. Many of the principles underlying NMAs and meta-analyses, including assumptions and approaches related to combining heterogeneous studies, are similar (Fu et al., 2010). Therefore, making these linkages more explicit for researchers and end-users may strengthen conduct and understanding of NMAs.

The growing number of NMA publications (Bafeta et al., 2013), as well as facilitators identified from the literature in our systematic review, suggest that there is growing acceptance of NMAs, however, researchers' knowledge of NMAs is still variable. This may be due to the rapidly evolving methodologies in this field. Lack of expertise was frequently noted as a barrier and increased training in NMAs was recommended in the two studies that formally surveyed researchers (Abdelhamid et al., 2012; Coleman et al., 2012). Efforts could be considered to determine optimal approaches for educational initiatives for both researchers and end-users of NMAs. Educational initiatives targeted at researchers should ensure methods are adequately understood while end-users should understand how much confidence to place in NMA results and how to use them to inform decisions. Tutorials and tools have been developed, including critical appraisal checklists or questionnaires to enhance the credibility of NMAs for decision-makers (Ades et al., 2013; Jansen et al., 2014; Mills et al., 2012; Salanti et al., 2012), and may assist in building and expanding these initiatives.

The need for guidance on various NMA aspects, including conduct, reporting, presentation and critical appraisal was frequently raised as either a barrier or facilitator. Guidance on NMA reporting and critical appraisal standards is being developed (Chaimani & Salanti, 2013; GRADE Working Group 2013; Hutton et al., 2014). The implementation of these standards will facilitate the use of NMAs and increase their transparency. However, these approaches may be insufficient for non-technical end-users' understanding of NMAs and targeted knowledge translation approaches that identify optimal presentation formats will be needed. The large volume of data that is sometimes synthesized and the multitude of options for presenting these data in an NMA was often identified as a challenge. Therefore, considering approaches that focus end-users on the most relevant information while still providing a comprehensive overview of the NMA may be of value. Providing information in layers, i.e. high-level information first with more detailed information in subsequent layers, is one approach that has been used to increase accessibility of information (CHSRF 2010). For example, optimizing the use of supplementary

appendices, which is a common journal feature, may be one application of this principle for NMAs. One study developed presentation formats incorporating **graphical design features** that facilitate readability of figures and tables, including the flexibility to select what results to present (Tan et al., 2014, Bax et al., 2009). Other reviews have explored what numerical formats are most easily understood by lay audiences such as patients and the public and have noted that tables are widely accessible to different audiences, pictographs are better understood than bar charts and uncertainty is not widely understood (Hildon et al., 2012). Although no evidence on end-users' preferences in receiving NMAs was available in this review, policymakers' preferences for receiving evidence have been previously reported and include formats that are quickly scannable, emphasize bottom-line conclusions and clearly identify real-world implications (Murthy et al., 2012; Lavis et al., 2005; Dobbins et al., 2007). The standard application of these various established principles to new approaches and presentation formats developed by methodologists could enhance interpretation and use of NMAs by end-users.

Strengths and Limitations

Although some reviews, including some studies identified herein, have attempted to summarize different approaches to reporting or presenting NMAs, this is the first review to integrate these findings with aspects related to knowledge of NMAs as well as barriers and facilitators to using NMAs. Broad inclusion criteria were applied to this systematic review to explore these issues, however, some exclusion criteria such as employing a defined methodology were applied; therefore opinion papers or editorials discussing relevant issues related to the interpretation and presentation of NMAs would have been excluded. There has been considerable discussion about the challenges of conducting and interpreting NMAs (Cipriani et al., 2013; Mills et al., 2013; Salanti et al., 2012), however, no reviews, and few individual studies, have attempted to formally quantify knowledge of NMAs or barriers or facilitators to NMAs. The lack of original studies creates limitations in this dataset, including the heterogeneous indicators of NMA knowledge and the reliance on researchers' subjective perspectives for identifying barriers and facilitators. However, this review represents a first attempt to provide an overview and snapshot of common issues faced in the field of NMAs. Finally, although numerous approaches to communicating NMAs were identified in this review, no assessment of the methodological appropriateness of these approaches was conducted, therefore, it is uncertain if all presentation formats would be appropriate. In this rapidly evolving field, ensuring appropriate methodologies are clearly established before recommending presentation formats is important.

Policy Implications

Taking these factors into consideration, and the fact that, in the absence of head-to-head trials, NMAs may be used by clinicians to inform treatment decisions in practice and by policymakers to

inform funding and other policy decisions, the following implications for policy and development of NMA communication approaches should be considered:

1) Understanding End-Users' Perspectives: The NMA presentation options in this review were developed by researchers and, in most cases, evaluation or consideration of end-users' (e.g. policymakers, clinicians) preferences was limited. In this absence, the use of established principles and findings from other fields could be applied to enhance presentation of NMAs including research on graphical design standards, numerical formats best understood by lay audiences and policymakers' preferences for receiving evidence.

2) **Communicating Uncertainty**: Uncertainty can be a challenging concept to communicate but is essential for informing good decisions. Despite the methodological strengths of NMAs, a number of sources of uncertainty exist including rapidly evolving methodologies, whether underlying NMA assumptions are appropriately assessed and met and the robustness of treatment rankings and effect sizes. Standard approaches such as reporting confidence or credible intervals may be insufficient for exploring these kinds of uncertainties to lay audiences. Real-world decisions must still be made in the face of uncertainty; therefore, communication approaches that help end-users understand these uncertainties and place them in the appropriate context of policy and treatment decisions is important.

3) **Building on Similarities with Traditional Meta-Analyses**: Knowledge of NMAs is variable and, yet, results from NMAs inform similar policy and clinical practice decisions as traditional meta-analyses. Although some features and results produced by NMAs are unique, many of the methodological approaches (e.g. searches, heterogeneity) and some of the presentation formats (e.g. forest plots) build upon principles established in traditional meta-analyses. Considering how to link and incorporate meta-analytic features, with which audiences have more familiarity and comfort, while contrasting key differences between NMAs and meta-analyses (e.g. network diagrams, inconsistency, treatment rankings) and highlighting areas of greater complexity in NMAs, may increase NMA knowledge and enhance interpretation of NMAs.

Future Research and Conclusions

This review provides an overview of approaches for communicating NMAs and considers them in the context of current knowledge of NMAs and potential barriers and facilitators to using NMAs. However, evidence is available only from researchers', and not end-users', perspectives and there is little evaluation of the communication approaches. These should both be considered in future, as strategies to inform development of knowledge translation tools.

Many of the presentation formats identified in this review may better help researchers understand data patterns and the robustness of results. However, these approaches may be insufficient to allow non-technical audiences to understand what the results mean and how to use them. Parallels may been seen in the more established field of health economics, where standard presentation formats used by economists may not be easily understood by non-economists (Galani et al., 2008; Williams et al., 2008). Addressing some of these issues early in this growing field of NMAs may lead to better understanding and acceptance of NMAs in future.

Knowledge translation approaches have frequently been applied to enhancing understandability and accessibility of clinical trial evidence for policy makers, healthcare providers and patients (Akl et al., 2012; Chambers et al., 2012; Trevena et al., 2006). Expansion of these approaches to NMAs has not yet occurred, but could be applied to develop explanations of NMAs for policy makers or other non-technical audiences that are both technically accurate and easily understood. Development of presentation formats will allow end-users to better apply the results of NMAs and enhance the transparency and legitimacy of policy and treatment decisions informed by NMAs.

REFERENCES

Abdelhamid A, Loke Y, Parekh-Bhurke S, Chen Y, Sutton A, Eastwood A, Hollanda R, Song F. Use of indirect comparison methods in systematic reviews: a survey of Cochrane review authors. Research Synthesis Methods 2012, 3: 71-79.

Ades, A.E., Caldwell, D.M., Reken, S., Welton, N.J., Sutton, A.J., Dias, S. NICE DSU Technical Support Document 7: Evidence synthesis of treatment efficacy in decision making: a reviewer's checklist. 2012; available from http://www.nicedsu.org.uk

Akl EA, Guyatt GH, Irani J, Feldstein D, Wasi P, Shaw E, Shaneyfelt T, Levine M, Schünemann HJ. "Might" or "suggest"? No wording approach was clearly superior in conveying the strength of recommendation. J Clin Epidemiol 2012; 65: 268-275.

Bafeta A, Trinquart L, Seror R, Ravaud P. Analysis of the systematic reviews process in reports of network meta-analyses: methodological systematic review. BMJ. 2013 Jul 1;347:f3675.

Bafeta A, Trinquart L, Seror R, Ravaud P. Reporting of results from network meta-analyses: methodological systematic review. BMJ. 2014 Mar 11;348:g1741.

Bax L, Ikeda N, Fukui N, Yaju Y, Tsuruta H, Moons K. More Than Numbers: The Power of Graphs in Meta-Analysis. Am J Epidemiol 2009;169:249–255.

CADTH. Grey Matters: a practical search tool for evidence-based medicine [2nd edition]. Ottawa: Canadian Agency for Drugs and Technologies in Health; 2013.

Canadian Health Services Research Foundation (CHSRF). Reader Friendly Writing, 1:3:25. 2010. Available at: http://www.cfhi-fcass.ca/Migrated/PDF/CommunicationNotes/cn-1325_e.pdf [Last accessed: August 9, 2014].

Chaimani A, Higgins JPT, Mavridis D, Spyridonos P, Salanti G. Graphical Tools for Network Meta-Analysis in STATA. PLoS ONE, 2013, 8: e76654.

Chaimani A & Salanti G. Background Document. Statistical issues in NMA (Stream 2). CMIMG Meeting. Bristol: June 2013. Available at: http://cmimg.cochrane.org/sites/cmimg.cochrane.org/files/uploads/CMIMG%20Stream%202%20document%202013%2010%2001.pdf [Last accessed: August 9, 2014]

Chambers D, Wilson P. A framework for production of systematic review based briefings to support evidence-informed decision-making. Syst Rev 2012; 1: 32.

Cipriani A, Higgins JP, Geddes JR, Salanti G. Conceptual and technical challenges in network meta-analysis. Ann Intern Med 2013, 159 (2): 130-137.

Cochrane Collaboration: Cochrane Handbook for Systematic Reviews. Available from: www.cochrane.org [Last accessed May 19, 2014].

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.

Cooper NJ, Peters J, Lai MC, Juni P, Wandel S, Palmer S et al. How valuable are multiple treatment comparison methods in evidence-based health-care evaluation? Value Health 2011, 14(2): 371-80.

Cope S & Jansen J. Quantitative summaries of treatment effect estimates obtained with network meta-analysis of survival curves to inform decision-making. BMC Medical Research Methodology 2013, 13:147

Daniels N & Sabin J. Limits to health care: Fair procedures, democratic deliberations and the legitimacy problem for insurers. Philosophy & Public Affairs 1997; 303-350.
Dhalla I, Laupacis A. Moving from opacity to transparency in pharmaceutical policy. CMAJ 2008;178: 428-31.
Dobbins M, Jack S, Thomas H, Kothari A. Public Health Decision-Makers' Informational Needs and Preferences for Receiving Research Evidence. Worldviews on Evidence-Based Nursing 2007; 4(3), 156–163.

Donegan S, Williamson P, D'Alessandrob U, Smith C. Assessing key assumptions of network meta-analysis: a review of methods. Research Synthesis Methods 2013, 4: 291–323.

Donegan S, Williamson P, Gamble C, Tudor-Smith C. Indirect comparisons. A review of reporting and methodological quality. PLoS One 2010, 5(11):e11054.

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.

Galani C and Rutten FFH (2008). Self-reported health care decision makers' attitudes towards economic evaluations of medical technologies *Cur Med Res Op* **24:** 3049-3058.

GRADE Working Group. Selection of the best available evidence form an evidence network of treatment comparisons. An approach to rate the quality of direct, indirect and pooled evidence. Rome: 2013. Available at: www.gradepro.org/Rome2013/13aNMA.doc [Last accessed August 9. 2014]

Greenhalgh T, Robert G, Bate P, Kyriakidou O, Macfarlane F, Peacock R. How to spread good ideas A systematic review of the literature on diffusion, dissemination and sustainability of innovations in health service delivery and organisation. London: Report for the National Co-ordinating Centre for NHS Service Delivery and Organisation R&D (NCCSDO); 2004.

Hannes K. Chapter 4: Critical appraisal of qualitative research. In: Noyes J, Booth A, Hannes K, Harden A, Harris J, Lewin S, Lockwood C (editors), Supplementary Guidance for Inclusion of Qualitative Research in Cochrane Systematic Reviews of Interventions. Version 1 (updated August 2011). http://cqrmg cochrane org/supplemental-handbook-guidance 2011. Available from: URL: http://cqrmg.cochrane.org/supplemental-handbook-guidance [Last accessed August 9, 2014]

Hildon Z, Allwood D, Black N (2012). Impact of format and content of visual display of data on comprehension, choice and preference: a systematic review. Int J Qual Health Care, 24 (1): 55-64.

Hutton B, Salanti G, Chaimani A, Caldwell DM, Schmid C, Thorlund K et al. The quality of reporting methods and results in network meta-analyses. An overview of reviews and suggestions for improvement. PLoS One 2014, 9(3): e92508.

Jansen J, Fleurence R, Devine B, Itzler R, Barrett A, Hawkins N, Lee K, Boersma C, Annemans L, Cappelleri J (2011). 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 in Health 14: 417–428.

Jansen JP, Trikalinos T, Cappelleri JC, Daw J, Andes S, Eldessouki R et al. Indirect treatment comparison/network meta-analysis study questionnaire to assess relevance and credibility to inform health care decision making: an ISPOR-AMCP-NPC Good Practice Task Force report. Value in Health 2014, 17(2): 157-173.

Lavis, J., Davies, H., Oxman, A., Denis, J. L., Golden-Biddle, K., & Ferlie, E. (2005). Towards systematic reviews that inform health care management and policy-making. Journal of Health Services Research and Policy, 10 (Suppl 1), 35-48.

Martin A, Rizzo M, Iheanacho I. Faulty connections. Can criticisms of network meta-analyses in NICE submission be avoided? Value in Health Conference ISPOR 16th Annual European Congress Dublin Ireland, November 2013, 16(7): A608.

Mills E, Ionnidis J, Thorlund K, Schunemann H, Pulan M, Guyatt G. How to use an article reporting a multiple treatment comparison meta-analysis. JAMA, 2012, 308: 1246-53.

Mills E, Thorlund K, Ioannidis J. Demystifying trial networks and network meta-analysis. BMJ 2013, 346: f2914.
Murthy L, Shepperd S, Clarke MJ, Garner SE, Lavis JN, Perrier L, Roberts NW, Straus SE. Interventions to improve the use of systematic reviews in decision-making by health system managers, policy makers and clinicians. Cochrane Database of Systematic Reviews 2012, Issue 9. Art. No.: CD009401.

Nikolakopoulou A. Chaimani A. Veroniki AA. Vasiliadis HS. Schmid CH. Salanti G. Characteristics of networks of interventions: a description of a database of 186 published networks. PLoS ONE 2014, 9: e86754.

Ohlssen D, Price, K, Xia HA, Hong H, Kerman J, Fu H et al. Guidance on the implementation and reporting of a drug safety Bayesian network meta-analysis. Pharmaceut. Statist, 2014, 13: 55-70.

Salanti G. Indirect and mixed-treatment comparison, network, or multiple-treatments meta-analysis. Many names, many benefits, many concerns for the next generation evidence synthesis tool. Research Synthesis Methods 2012; 3:80-97.

Salanti G, Ades AE, Ioannidis JPA. Graphical methods and numerical summaries for presenting results from multiple-treatment meta-analysis: an overview and tutorial. J Clin Epidemiol 2011;64:163-71.

Salanti G, Del Giovane C, Chaimani A, Caldwell DM, Higgins JP. Evaluating the quality of evidence from a network meta-analysis. PLoS One 2014, 9(7): e99682.

Shea BJ, Hamel C, Wells GA, Bouter LM, Kristjansson E, Grimshaw J, Henry DA, Boers M. AMSTAR is a reliable and valid measurement tool to assess the methodological quality of systematic reviews. J Clin Epidemiol. 2009, 62:1013-20.

Sobieraj D, Cappelleri JC, Baker WL, Phung O, White CM, Coleman C. Methods used to conduct and report Bayesian mixed treatment comparisons published in the medical literature: a systematic review. BMJ Open, 2013, 3:e003111.

Song F, Loke YK, Walsh T, Glenny AM, Eastwood AJ, Altman DG. Methodological problems in the use of indirect comparisons for evaluating healthcare interventions: survey of published systematic reviews. BMJ 2009, 338: b1147.

Straus S, Tetroe J & Graham ID (Eds.), Knowledge translation in health care: Moving from evidence to practice, 2nd edition. Oxford: Blackwell Publishing, 2013.

Tan SH, Bujkiewicz S, Sutton A, Dequen P, Cooper N. Presentational approaches used in the UK for reporting evidence synthesis using indirect and mixed treatment comparisons. J Health Serv Res Policy 2013, 18(4): 224-232.

Tan SH, Cooper NJ, Bujkiewicz S, Welton NJ, Caldwell DM, Sutton AJ. Novel presentational approaches were developed for reporting network meta-analysis. J Clin Epidemol 2014, 67(6): 672-680.

Thomas J, Harden A. Methods for the thematic synthesis of qualitative research in systematic reviews. BMC Medical Research Methodology 2008;8: 45 doi:10.1186/1471-2288-8-45.

Trevena L, Davey H, Barratt A, Butow P, Caldwell P. A systematic review on communicating with patients about evidence. J Evalu Clin Practice 2006;12: 13-23.

Veroniki A, Vasiliadis H, Higgins JP, Salanti G. Evaluation of inconsistency in networks of interventions. Int J Epidemiol 2013, 42(1): 332-345.

Williams I, McIver S, Moore D, Bryan S. The use of economic evaluations in NHS decision-making: a review and empirical investigation. Health Technol Assess 2008; 12: 1-175.

TABLES AND FIGURES

Table 1: Characteristics of Included Studies.

STUDY CHARACTERISTICS			
Author and Year	**Study Design**	**Number of Participants**	**Location**
Ohlssen 2014	Case Study	1 NMA (31 studies of 7 therapies)	North America
Hutton 2014	Review	8 reviews of NMAs or IDCs	North America
Bafeta 2014	Review	121 NMAs	Europe
Nikolakopoulou 2014	Review	186 NMAs	Europe
Tan 2013	Review	19 NMAs	UK
Tan 2014	Case Study	1 NMA (95 studies of 8 drug classes)	UK
Chaimani 2013	Case Study	3 NMAs (47 studies of 14 therapies; 62 studies of 4 therapies; 27 studies of 6 therapies)	Europe
Cope 2013	Case Study	1 NMA (10 studies of 4 therapies)	North America
Bafeta 2013	Review	121 NMAs	Europe
Donegan 2010	Review	43 IDCs	UK
Song 2009	Review	88 IDCs	UK
Salanti 2011	Case Study	1 NMA (24 studies of 5 therapies)	Europe
Salanti 2014	Case Study	1 NMA (19 studies of 4 therapies)	Europe
GRADE Report 2013	Case Study	2 NMAs (35 trials of 5 therapies; 40 trials of 11 therapies)	International
Colemen 2012 & Sobieraj 2013	Review and Focus Group	43 NMAs (34 Bayesian, 9 frequentist) and 9 NMA Authors	North America
Abdelhamid 2012	Survey	48 Review Authors (57% response rate)	UK
Abbreviations: IDC=indirect comparison; NMA=network meta-analysis; UK=United Kingdom			

Table 2. Key Findings Related to Formatting Options for Communicating NMAs

Key Findings Related to High-Level Approaches to and Formats for Communicating NMAs

• Mix of multiple formats used to present NMAs, including tables, figures and text
• Some journal have limitations on number of figures or tables but supplementary materials option usually permits sufficient information be provided
• Different audiences may prefer different formats or information
• Graphics can support different aspects of NMAs including:
  ○ assessing assumptions
  ○ supporting quality assessments
  ○ reporting results
  ○ considering uncertainty
• Graphical design features to enhance readability of figures and interpretation of results include:
  ○ alternate shading of cells
  ○ enclosures (e.g. boxes) to emphasize groupings
  ○ placements (e.g. side-by-side) to facilitate comparisons
  ○ fitting information on one page or in one graph
  ○ flexible presentation options to reduce volume of comparisons or change order of outcomes
  ○ providing most important information first
  ○ providing most important numbers in darker font or creating greater contrast with shading options

Table 3. Key Findings on Approaches to Communicating NMA Data Sources

Key Findings on Approaches to Communicating NMA Data Sources Based on x Studies	
Evidence Networks	
• Network diagrams highlighting features by weighting or coloring nodes and edges or changing the style of edges, including: o sample sizes and number of patients per treatment arm o number of studies or patients per comparison, o presence of multi-arm studies o effect modifiers and/or key trial characteristics o risk of bias • Evidence network tables • Contribution plot identifying what direct and indirect evidence contributes to each NMA estimate and the overall network	
Trial Characteristics	
• Tabular format (characteristics not specified) • Network diagram incorporating key characteristics (not otherwise specified)	
Critical Appraisal	
RoB Assessments	• Tables reporting RoB assessment of individual trials along with other trial-level characteristics • Evidence network diagrams with coloured lines representing RoB assessments for individual trials • Pie chart showing RoB contributions to each network comparison • Bar graph showing RoB contributions to network estimates and direct estimates
GRADE-Related Assessments	• Table formats for presenting conclusions on GRADE quality assessment components o assessments presented separately for effect estimates and ranking estimates o assessments presented separately for direct, indirect and mixed estimates • Effect estimate assessments informed by contribution plots, RoB pie charts and/or RoB bar graphs • Ranking estimate assessments informed by rankograms and comparison of probabilities treatment rankings
Publication bias Assessments	• Comparison-adjusted funnel plot • Contour enhanced funnel plot
Abbreviations: GRADE=Grading of Recommendations Assessment, Development and Evaluation; NMA=network meta-analysis; RoB=Risk of Bias	

Table 4. Key Findings on Approaches to Communicating and Presenting NMA Assumptions

Key Findings on Approaches to Communicating and Presenting NMA Assumptions
Heterogeneity

• Forest plots including heterogeneity measures (e.g., I^2 and Tau^2)
• Tables including heterogeneity measures (e.g., I^2 and Tau^2)
• Network diagrams including heterogeneity measures (e.g., $I^{2)}$
• Heterogeneity assessments informed by:
  o Network diagrams incorporating potential effect modifiers
  o Predictive intervals included on forest plots

Inconsistency
• Inconsistency plot • Inconsistency assessments incorporated into GRADE-related conclusion tables

Table 5. Key Findings on Approaches to Communicating NMA Results

Key Findings on Approaches to Communicating NMA Results
Effect Estimates (Including Comparisons of Direct and Indirect/Mixed Estimates)

• Forest plots
• Caterpillar plots
• Matrix table with effect estimates for all treatment comparisons separated by the diagonal
  o Can include either or both mixed and direct estimates
• Tables of relative or absolute effects
• Summary forest plot matrix or a summary forest plot table
  o Both incorporate features from tables, forest plots and rankings
  o Uses graphical design features to facilitate comparisons
  o Flexibility in what effect estimates are presented
• Assessments of 'indirectness' incorporated into GRADE-related conclusion tables

Rankings
• Table providing list of rankings, probabilities each treatment is best, cumulative rankings and/or SUCRA %   o Uncertainty considered by comparing with magnitude of effect estimates • Rankograms • Bar charts with the probability each treatment has a specific rank • Cumulative probability ranking plots (can assess ranking uncertainty) • Plots of posterior distributions of rankings (can assess ranking uncertainty) • Plot of treatment ranking versus SUCRA % • Plot of treatment ranking versus MDS estimate • Threshold plot based on an acceptable tolerance threshold and where each treatment ranking falls in relationship to it • Clustered ranking plot • Summary forest plot matrix or the summary forest plot table incorporating rankings and associated probabilities • Median ranking chart (simplified rankings in large networks) • Survival rankings based on time-independent or time-varying measures (bar charts vs. line graphs) with rankings represented through rankograms, SUCRA% or probability (best)
Uncertainty
• Confidence intervals (frequentist methods) or credible intervals (Bayesian methods) • Predictive intervals and related forest plots to identify uncertainty • Uncertainty associated with rankings (see ranking section) • Graphs presenting uncertainty over time in survival rankings (see ranking section)

Table 6. Key Findings Related to Knowledge of NMAs

Key Findings Related to Knowledge of NMAs

• Increase over time in NMAs conducted
• Inadequate use of methods in some NMAs including:
  o Statistical approaches that don't preserve randomization
  o Statistical handling of 3-arm trials
  o Assessment of key NMA assumptions, e.g. heterogeneity, inconsistency
• Inadequate application of SR or MA principles underlying some NMAs
  o e.g. searches, data collection
• NMAs sometimes misinterpreted in literature
  o Better assessments of uncertainty may prevent misinterpretations
• Those conducting NMAs:
  o Very certain of its value/strength of approach
  o Considered NMA terminology not always used clearly
  o Approx. half thought reporting/conduct guidance on NMAs adequate
  o Most thought useful to include a network diagram
  o No clear trends identified regarding:
  ▪ if direct evidence takes precedence over indirect;
  ▪ if direct and indirect estimates both should always be considered and reported separately
  ▪ if comparisons not of direct relevance (e.g. placebo, older treatments) should be considered in an NMA.
• Most Cochrane reviewers:
  o Had some knowledge of NMAs but level of knowledge varied
  o Would consider indirect evidence, at least some, if not all of the time

Abbreviations: MA=meta-analysis; NMA=network meta-analysis; SR=systematic review

*Based on two qualitative studies questioning researchers about NMAs

** Based on indirect measures related to knowledge, barriers or facilitators identified in NMA case studies or review

Figure 1. PRISMA Flow Diagram for Results of Systematic Review Search.

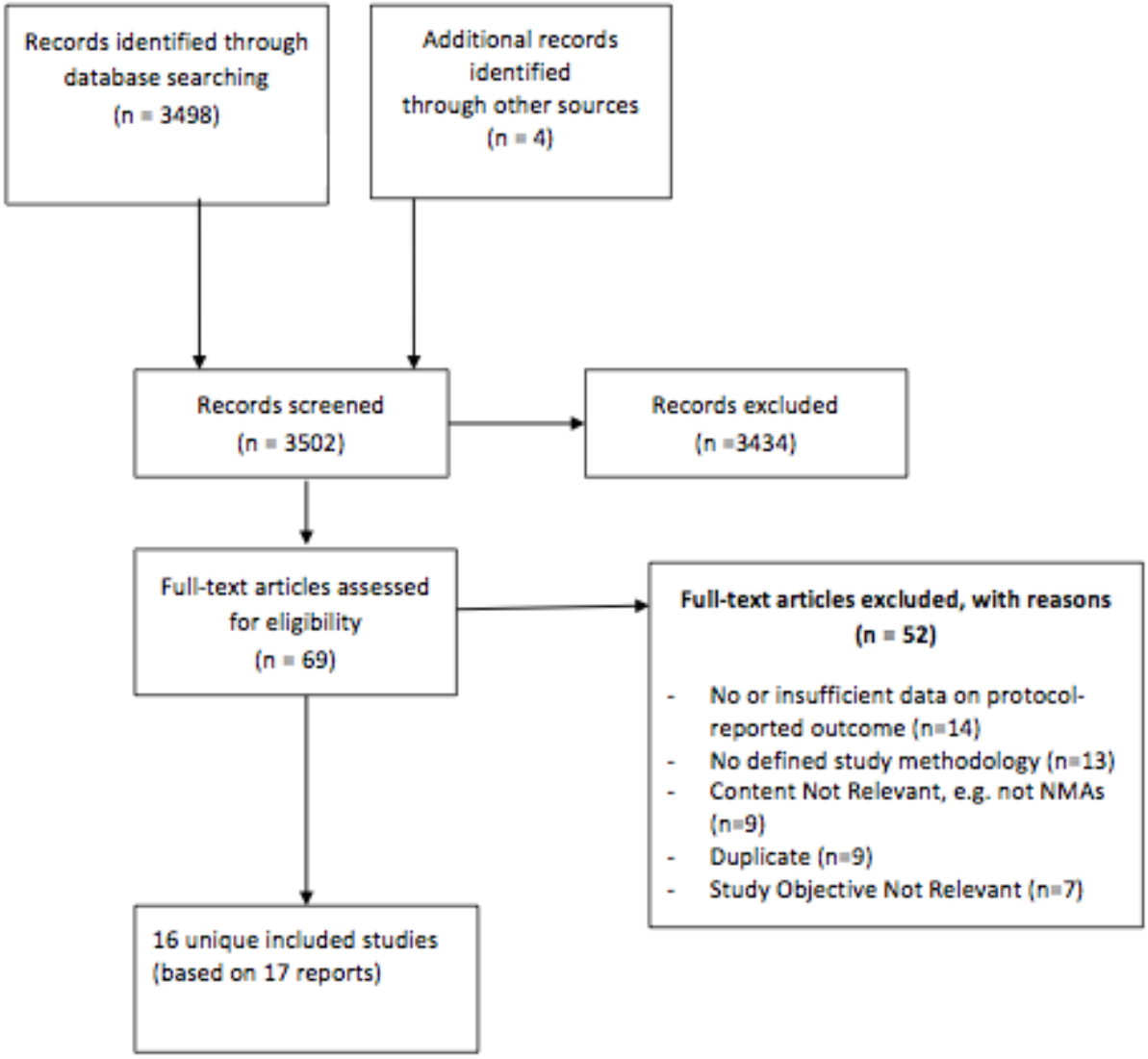


Figure 2A. Most Frequently Reported Types of Barriers to Using and Communicating NMAs.

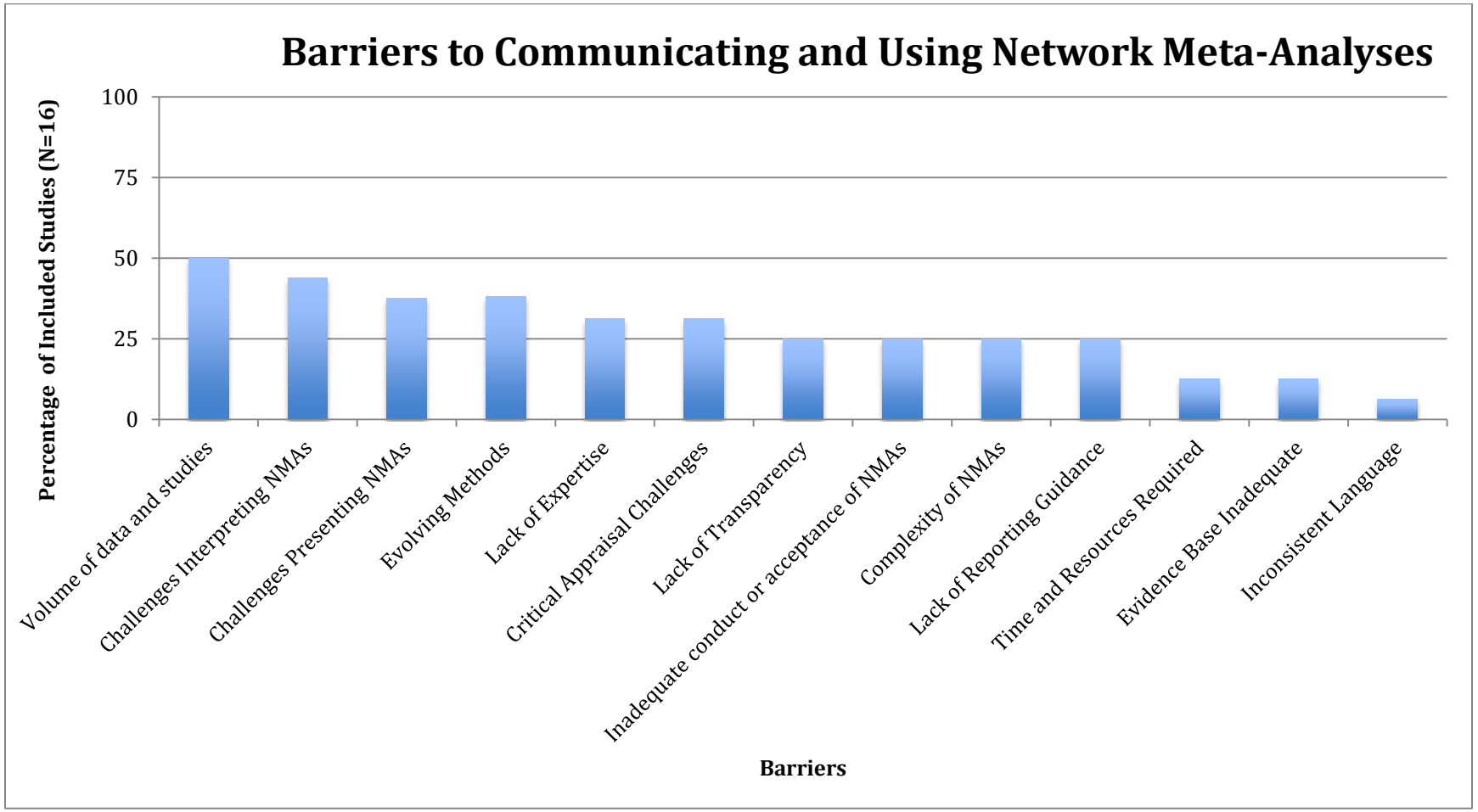
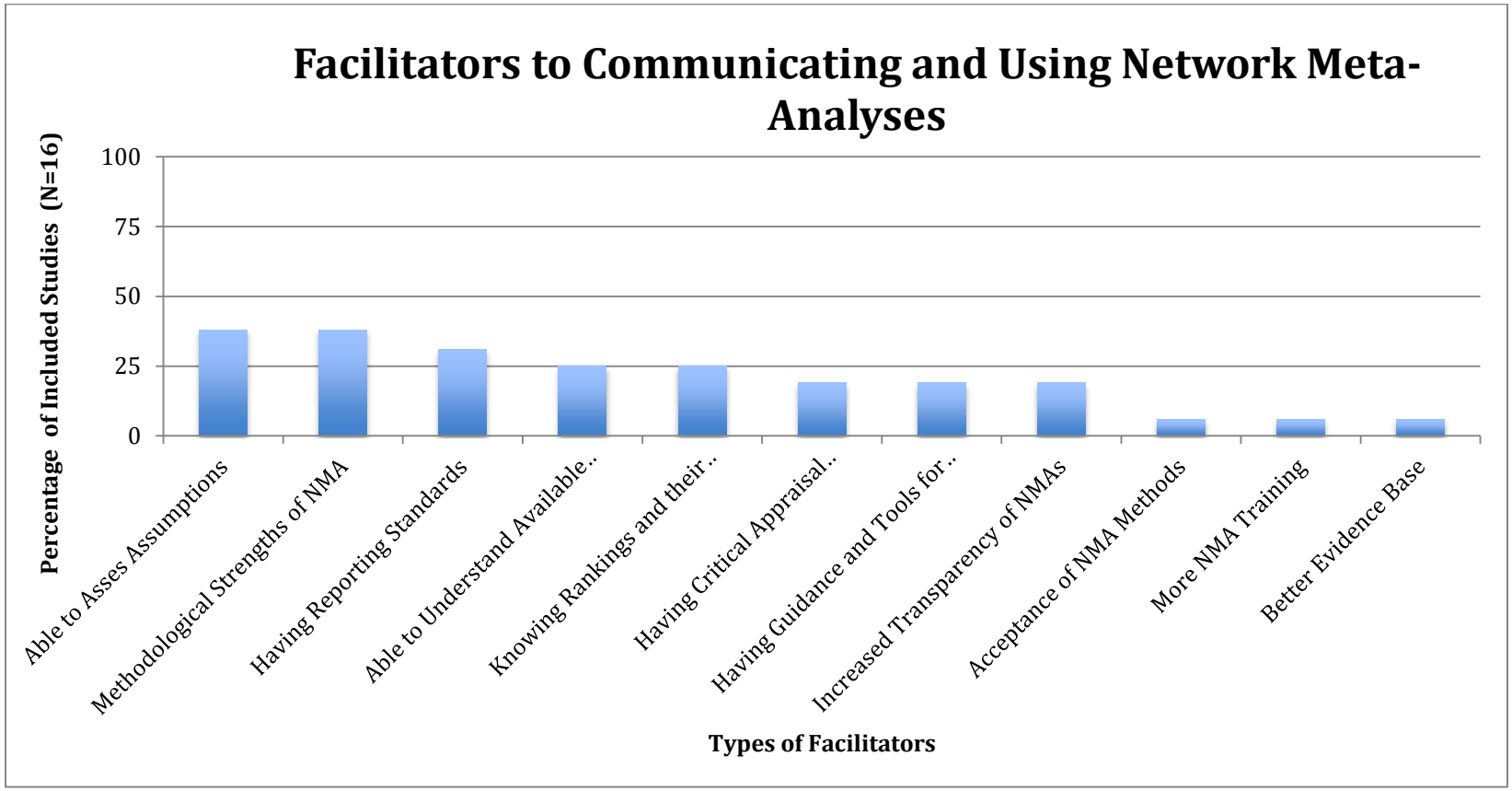


Figure 2B. Most Frequently Reported Types of Facilitators to Using and Communicating NMAs.



SUPPLEMENTARY MATERIAL

Supplemental Table 1A. Quality assessment of included reviews based on AMSTAR criteria.

REVIEWS	Research Protocol	Search and Inclusion	Data Extraction	Included Study Details	Analysis and Conclusions	Publication Bias	COI	Other Sources and Notes
Bafeta 2013		x				x		**Strengths**: pre-defined research question; multiple databases searched; independent data extraction; study characteristics reported; quality assessment key study objective; appropriate narrative synthesis; COI and funding reported. **Limitations**: no grey lit search; excluded studies not listed; publication bias unknown
Bafeta 2014		x			x	x		**Strengths**: pre-defined research question; multiple databases searched; double-data extraction; study characteristics reported; quality assessment aspects previously reported; appropriate narrative synthesis; COI and funding reported. **Limitations**: no grey literature search; excluded studies not listed; publication bias not considered; quality assessment not considered in conclusions.
Donegan 2010		x				x		**Strengths**: pre-defined research question; multiple databases searched; independent data extraction; study characteristics reported; quality assessment key study objective; appropriate narrative synthesis; COI and funding reported. **Limitations**: no grey literature

REVIEWS	Research Protocol	Search and Inclusion	Data Extraction	Included Study Details	Analysis and Conclusions	Publication Bias	COI	Other Sources and Notes
								search; excluded studies not listed; publication bias not considered;
Hutton 2014		x			x	x	x	**Strengths**: pre-defined research question; two databases searched; double-data extraction; study characteristics reported; appropriate narrative synthesis. **Limitations**: no grey literature search; excluded studies not listed; no quality assessment of studies; publication bias not assessed; conclusions don't consider study quality; industry-funding and potential COI
Nikolakopoulou 2014		x				x		**Strengths**: pre-defined research question; data independently extracted; individual study characteristics provided; quality assessment is key study objective; appropriate narrative synthesis; funding and COI reported **Limitations**: only one electronic database searched; excluded studies not listed; publication bias not assessed.
Song 2009		x			x	x		**Strengths**: pre-defined research question; data extraction independently verified; individual study characteristics provided; quality assessment is key study objective; appropriate narrative synthesis; COI reported **Limitations**: only one electronic database searched; excluded

REVIEWS	Research Protocol	Search and Inclusion	Data Extraction	Included Study Details	Analysis and Conclusions	Publication Bias	COI	Other Sources and Notes
								studies not listed or totals identified; publication bias not assessed.
Tan 2013		x	x		x	x	x	**Strengths**: pre-defined research question; individual study characteristics provided; appropriate narrative synthesis. **Limitations**: Search limited to unpublished studies based on scope of research question; excluded studies (N=187) not listed; duplicate data extraction not reported; no quality assessment of studies or consideration in conclusions; publication bias not assessed but included studies from grey literature; industry-funded and COI not reported
Abbreviations: COI=conflict of interest								

'x' – indicates an issue was identified during critical appraisal of the study

Supplemental Table 1B. Quality assessment of qualitative or mixed methods studies using checklists from Greenhalgh et al. 2004

QUALITATIVE or MIXED METHODS STUDIES	Research Question/ Design	Sampling/ Response	Data Collection/ Instrument	Analysis	Results/ Interpretation	Other Sources and Notes
Coleman 2012 * (Review and Focus Group)		x	x		x	**Strengths:** integration and triangulation of data form systematic reviews and focus groups; well-conducted systematic review; purposive interview sample; analysis methods well-described; many illustrative examples; government funded **Limitations:** focus group sample size insufficient for generalizations on NMA analysis methods; questionnaire validation not reported; biases of researchers on interpretations not considered; ethics approval not reported; COI not reported
Abdelhamid 2012 (Online Survey)					x	**Strengths:** purposive sample; sample size sufficient for conceptual generalizations; reasonable response rate (57%); survey development methods reported; analysis methods well-described and appropriate; illustrative examples provided; ethics approval reported; independent funding **Limitations:** biases of researchers on interpretations not considered; COI not reported

QUALITATIVE or MIXED METHODS STUDIES	Research Question/ Design	Sampling/ Response	Data Collection/ Instrument	Analysis	Results/ Interpretation	Other Sources and Notes
Abbreviations: COI=conflict of interest; NMA=network meta-analyses						

'x' – indicates an issue was identified during critical appraisal of the study

* Because Coleman 2012 provided more details on study methodology, the companion publication (Sobieraj 2013) was only assessed for supportive details as needed.

Supplemental Table 1C Quality assessment of case studies adapted from checklists in Greenhalgh et al. 2004

CASE STUDIES	**Research Question & Design**	**Case Study Selection**	**Case Study Characteristics**	**Analysis Framework**	**Results & Interpretation**	**Other Sources and Notes**
Chaimani 2013		x		x	x	**Strengths**: appropriate study design; applications to 3 cases increase generalizability; characteristics of case provided; interpretation of case study data provided; COI reported; independently funded. **Limitations**: development of case not grounded in theoretical framework; researchers' biases on case development and interpretation not considered;
Cope 2013		x		x	x	**Strengths**: appropriate study design; characteristics of case provided; interpretation of case study data provided; no funding source **Limitations**: one case insufficient for generalizations; development of case not grounded in theoretical framework; researchers' biases on case development and interpretation not considered; no COI reported but industry-related affiliations.
GRADE Working Group 2013					x	**Strengths**: appropriate study design; rationale for case selection provided; applications to 2 cases increase generalizability; development of case study grounded in theoretical framework; characteristics of cases provided; interpretation of case study data provided; subjectivity of researchers' biases on case development and interpretation considered. **Limitations**: COI and funding not reported.

CASE STUDIES	**Research Question & Design**	**Case Study Selection**	**Case Study Characteristics**	**Analysis Framework**	**Results & Interpretation**	**Other Sources and Notes**
Ohlssen 2014		x			x	**Strengths**: appropriate study design; development of case study grounded in Bayesian framework; characteristics of case provided; interpretation of case study data provided; **Limitations**: rationale for case selection not provided; one case insufficient for generalizations; researchers' biases on case development and interpretation not considered; COI and funding not reported; industry-affiliations.
Salanti 2011		x			x	**Strengths**: appropriate study design; rationale for case selection provided; grounded in Bayesian framework; characteristics of cases provided; interpretation of case study data provided. **Limitations**: one case insufficient for generalizations; researchers' biases on case development and interpretation not considered; COI and funding not reported.
Salanti 2014		x				**Strengths**: appropriate study design; development of case study grounded in theoretical framework; characteristics of cases provided; interpretation of case study data provided; subjectivity of researchers' biases on case development and interpretation considered; COI reported; independently funded. **Limitations**: rationale for case selection not provided; one case insufficient for generalizations.

CASE STUDIES	**Research Question & Design**	**Case Study Selection**	**Case Study Characteristics**	**Analysis Framework**	**Results & Interpretation**	**Other Sources and Notes**
Tan 2014		x			x	**Strengths**: appropriate study design; development of case study grounded in theoretical framework; characteristics of cases provided; interpretation of case study data provided; **Limitations**: rationale for case selection not provided; one case insufficient for generalizations; subjectivity of researchers' biases on case development and interpretation not considered; COI not reported; industry-funding.
Abbreviations: COI=conflict of interest						

'x' – indicates an issue was identified during critical appraisal of the study

Supplementary Table 2. Detailed Included Study Characteristics

Author and year	Study Objective	Type of NMA Information	How NMA Information Used
Abdelhamid 2012	To investigate the views and knowledge of Cochrane systematic review authors regarding the use of NMAs	• Methods Characteristics • Training Needs	To improve appropriate use of NMAs
Bafeta 2013	To examine whether NMAs follow key methodological recommendations for reporting and conduct of systematic reviews	• Reporting Characteristics	To inform NMA conduct and reporting guidance
Bafeta 2014	To examine how results of NMAs are reported	• Reporting Characteristics	To inform quality assessment or reporting standards
Chaimani 2013	To make NMA methodology accessible to non-statisticians by presenting and explaining a series of graphical tools via worked examples. To provide a set of STATA routines that can be easily applied to NMAs	• Presentation Formats	To enhance understanding of NMAs
Coleman 2012 & Sobieraj 2013	To summarize guidelines on conducting, reporting and interpreting NMAs. To summarize characteristics of published NMAs. To gain insight from NMA authors on conducting and reporting NMAs (barriers, strengths, training, preferences). Note: Sobieraj 2013 only reports on characteristics of Bayesian NMAs, not other components of Coleman 2012.	• Methods Characteristics • Reporting Characteristics • Training Needs	To inform guidance on NMA conduct, reporting, interpretation
Cope 2013	To present different alternatives for quantitative summaries of treatment effect estimates obtained with NMA of survival data to help inform decision-making	• Presentation Formats	To enhance understanding of NMAs
Donegan 2010	To systematically review the quality of published indirect comparisons to add to existing empirical data suggesting that improvements can be made when reporting and applying indirect comparisons	• Reporting Characteristics	To inform NMA quality criteria
GRADE Working Group 2013	To present an approach of the GRADE Working Group to identifying the best available evidence from NMAs	• Presentation Formats • Quality Assessment	To enhance understanding of NMAs
Hutton 2014	To examine findings of peer reviewed literature on the key limitations in the reporting of NMAs as part of the development of a statement extending PRISMA to cover NMAs. To help assess the need for reporting guidance and focus the PRISMA guideline on those features of most importance to clear reporting of NMAs.	• Reporting Characteristics	To inform development of reporting guidelines

Author and year	Study Objective	Type of NMA Information	How NMA Information Used
Nikolakopoulou 2014	To provide an overview of the characteristics of published NMAs. To direct NMA methods developments applicable to the majority of NMAs and not just special cases. To explore uptake of new NMA methods. To investigate whether the choice of NMA method is associated with a network's structural characteristics	• Methods Characteristics • Network Characteristics	To inform NMA methods development
Ohlssen 2014	To produce a checklist for the conduct and reporting of Bayesian NMAs	• Methods Characteristics • Reporting Characteristics	To inform NMA conduct and reporting guidance
Salanti 2011	To present simple graphical and quantitative ways to assist interpretation and improve presentation of results from NMAs	• Presentation Formats	To improve presentation of NMA results
Salanti 2014	To propose an approach to evaluating and determining the quality of and confidence in the output of an NMA based on the GRADE framework	• Presentation Formats • Quality Assessment	To enhance understanding of NMAs
Song 2009	To investigate basic assumptions and other methodological problems in the application of IDCs in systematic reviews of competing healthcare interventions	• Methods Characteristics	To improve NMA methods
Tan 2013	To establish guidance and current practice on presentation of NMAs in the UK. To provide recommendations to improve NMA reporting. To identify research priorities for improved NMA presentation	• Presentation Formats	To inform standards for presentation of NMAs
Tan 2014	To present graphical tools for reporting NMA results aiming to increase the accessibility, transparency, interpretability, and acceptability of NMAs	• Presentation Formats	To enhance understanding of NMAs

Abbreviations: GRADE= Grading of Recommendations Assessment, Development and Evaluation; IDC=indirect comparison; NMA=network meta-analysis; PRISMA=Preferred Reporting Items for Systematic Reviews and Meta-Analysis; UK=United Kingdom

Supplementary Table 3. Study-Level Details on Formatting Options for Presenting NMAs

Study	Details on High-Level Approaches and Formats for Communicating NMAs
Bafeta 2013	• 89% results were presented in tables and only 21% were presented in graphs.
Bafeta 2014	• Format of results provided for 54% NMAs: 41% tables, 6% forest plots, 5% matrices, 3% figures. • Number of outcomes assessed (median=2, range=1 to 3). • 58% (n=70/121) had a supplement or appendix. • Number of printed pages (median=11, range=9 to 15). • Number of tables/network (median=5, range=3 to 6). • Number of figures/network (median=3, range=2 to 5).

Chaimani 2013	Pros and cons of 9 graphical options discussed and presented:
	• Network plot
	• Contribution plot
	• Inconsistency plot
	• Comparison-adjusted funnel plot
	• Predictive intervals plot
	• Ranking plots for a single outcome using probabilities
	  o Cumulative rankograms
	  o Ranking plot using SUCRA values
	• Ranking plot for a single outcome using multidimensional scaling
	• Clustered ranking plot for two outcomes

Coleman 2012 and Sobieraj 2013	Length and use of tables and figures in published NMAs and potential impact of journal limitations reported:
	• **Journal Characteristics:**
	  o 66% word count limit, 44% table and figure limit, 78% appendix allowed
	• **Bayesian NMAs**:
	  o 71% tables, 94% text, 62% figures used to report results.
	  o Mean number printed pages 16.6 +/- 36.3 and 58.8% had a supplement or appendix.
	• **Frequentist NMAs**:
	  o 33% tables, 89% text, 67% figures used to report results.
	Mean number printed pages 16.1 +/- 16.0 and 67% had a supplement or appendix

Cope 2013	Six figures explored to most accurately and easily interpret NMA survival results Rankings of survival outcomes using time-independent or time-varying measures (bar charts vs. line graphs) with rankings represented through rankograms, SUCRA% or probability of being best
GRADE Working Group 2013	Summary Table reporting GRADE quality assessment
Hutton 2014	Optimal use of tables and figures to convey results to readers recommended

Study	Details on High-Level Approaches and Formats for Communicating NMAs
Salanti 2011	Pros and cons of 11 different graphical options discussed and presented: • Forest plot with mean effect and 95% credible intervals • Correlation graphs between effect sizes • Predictive intervals • Table of treatment rankings, probabilities of best, cumulative probabilities and SUCRA% ○ e.g. Ranking table with SUCRA%, median rank (95% CrI) and probability no worse than a threshold value • Rankograms with rank probabilities versus possible ranks • Bar chart for ranking probabilities • Cumulative ranking probability plots • SUCRA numerical summaries • Plots of posterior distribution for rank of each treatment with median, 95% credible intervals • Threshold plot with probability for each treatment to be within a range from the best treatment • Summary table of treatment effects with credible and predictive intervals
Salanti 2014	Ten figures and tables recommended to assist in GRADE-related quality assessment: • Evidence network diagram • Table summary of direct results with heterogeneity estimates • Contribution plot • Table of rankings and probabilities of each treatment achieving each possible rank • Risk of bias bar chart for network and direct estimates • Risk of bias pie chart for each comparison • Forest plot with effect estimates, confidence or credible intervals and predictive intervals • Summary table of GRADE-related quality assessment for each comparison • Rankograms • Comparison-adjusted funnel plot

Tan 2013	• 84% NMAs presented effect estimates in tables and 25% also presented in either forest plots or caterpillar plots
	• Various presentation formats recommended for different audiences:
	  ○ Network tables for academics and decision-makers
	  ○ Analysis codes and model data for academics and statistical analysts
	  ○ Summary forest plots and/or tables of relative effects for selected comparisons for academics
	  ○ Matrix tables of all comparison and summary forest plot or caterpillar plot for decision makers
	  ○ Tables or rankograms with probability best statistics and treatment rankings for

Study	Details on High-Level Approaches and Formats for Communicating NMAs
Tan 2014	Three figures created with formatting options that support readability and flexibility: • Summary forest plot matrix • Summary forest plot table • Median ranking chart
Abbreviations: GRADE= Grading of Recommendations Assessment, Development and Evaluation; NMA=network meta-analysis; NR=not reported; SUCRA= Surface under the cumulative ranking curve	

Note: Abdelhamid 2012, Donegan 2010, Nikolakopoulou 2014, Ohlssen 2014, Song 2009 did not contain data on NMA formatting and presentation options

Supplementary Table 4. Study-Level Details on Approaches to Communicating NMA Data Sources, Assumptions and Results.

Category	Suggestions for Presentation	Frequency of Reporting in Published NMAs
NMA Data Sources		
Evidence Networks	• Network diagrams including features such as sample sizes, the number of studies per comparison and the presence of multi-arm studies • Network diagram that incorporates features such as sample size, effect modifiers, risk of bias by weighting or coloring nodes and edges • Contribution plot that identifies how much direct and indirect evidence contributes to each NMA estimate as well as the overall network	• 74% NMAs provided a network diagram (based on 1 review) • 21% NMAs provided a network diagram, 58% provided evidence network tables and 11% provided both (based on 1 review)
Trial Characteristics	• Provide in tabular format • Incorporate key characteristics into network diagram	• 21% to 88% NMAs reported direct trial characteristics (based on 3 reviews)
Critical Appraisal	• RoB presented in tables with other trial-level characteristics • RoB incorporated into evidence network diagrams • Table formats for presenting GRADE-related quality assessments • Critical appraisal of effect estimates and ranking estimates separately • Contribution plots, RoB pie charts and/or RoB bar graphs to inform GRADE quality assessment of NMA treatment effects • Rankograms and tables of associated metrics to inform GRADE quality assessment of NMA rankings • Comparison-adjusted funnel plot to assess potential publication bias and inform GRADE quality assessments	• 58% NMAs did not report RoB assessments for individual studies (based on 1 review) • 84% did not report methods for assessing publication bias (based on 1 review)
NMA Methods		
Heterogeneity	• Presenting mean summary effects with predictive intervals to facilitate interpretation of impact of potential effect modifiers and heterogeneity • Inclusion of Tau^2 on forest plot of effect estimates • Between-study variance, Tau^2, frequently suggested measure to report NMA heterogeneity but different approaches to presenting this information not identified	• 32% NMAs evaluated heterogeneity in the NMA (based on 1 review) • 61% evaluated heterogeneity in accompanying traditional pairwise MA (based on 1 review) • 32% of Bayesian NMAs reported an evaluation of heterogeneity (based on 1 review)

Category	Suggestions for Presentation	Frequency of Reporting in Published NMAs
Inconsistency	• Inconsistency plot • Incorporation into GRADE-related quality assessment table	• 11% to 71% NMAs reported and evaluated inconsistency (based on 7 reviews)
NMA Results		
Effect Estimates	• Forest plots or tabular formats (e.g. relative and absolute effect tables, matrix tables) frequently used for presenting various effect estimates • Recommend direct and indirect evidence be presented separately • Direct and indirect evidence should be compared as part of GRADE quality assessment of 'indirectness' • Recommend considering which comparisons, effect estimates should be reported • Use of forest or caterpillar plots to present effect estimates • Tabular formats used more frequently than graphical formats • Generally, three table formats used: matrix of effects for all treatment comparisons separated by the diagonal, tables of relative effects, tables of absolute effects • Matrix tables can be used to present both NMA and traditional pairwise effect estimates • Summary forest plot matrix or a summary forest plot table, which incorporate features from both tables and forest plots and apply graphical design features such as shading and juxtapositioning of information to allow for easy comparisons (can also incorporate rankings in these figures) • Overall variability in what effect estimates are presented (e.g. based on direct, indirect or mixed evidence, which treatment comparisons, absolute or relative measures, how many outcomes)	• 56% of NMAs reported direct results from individual trials (based on 1 review) • 52% to 89% of NMAs reported direct results from traditional pairwise meta-analyses (based on 3 reviews)
Uncertainty	• Confidence intervals (frequentist methods) or credible intervals (Bayesian methods) most frequently presented • Predictive intervals and related plots to identify uncertainty • Uncertainty associated with rankings should also be considered (see ranking section) • Uncertainty associated with survival estimates over time presented through various graphical options	• All NMAs presented either confidence or credible intervals (based on 3 reviews)

Category	Suggestions for Presentation	Frequency of Reporting in Published NMAs
Rankings	• Table providing list of probabilities each treatment is best • Rankograms • Bar charts with the probability each treatment has a specific rank • Plot based on SUCRA % (surface under the ranking line) • Plot based on an acceptable tolerance threshold and where each treatment ranking falls in relationship to it • Treatment ranking versus multidimensional space (MDS) estimate • Summary forest plot matrix or the summary forest plot table incorporating rankings and associated probabilities • Median ranking chart to simplify rankings in large networks • Rankings of survival outcomes using time-independent or time-varying measures (bar charts vs. line graphs) with rankings represented through rankograms, SUCRA% or probability (best) • Uncertainty in rankings presented through cumulative probability plots, plots of posterior distributions of rankings, side-by-side consideration of magnitude of treatment effect estimates	• 26% to 62% NMAs presented treatment rankings or associated probabilities treatment is best (based on 4 reviews)

Abbreviations: GRADE= Grading of Recommendations Assessment, Development and Evaluation; IDC=indirect comparison; MA=traditional pairwise meta-analysis; MDS=multi-dimensional space; NMA=network meta-analysis; RoB=Risk of Bias; SUCRA= Surface under the cumulative ranking curve

* Could include information on either suggested ways of presenting NMAs or information on how reported in completed NMAs

Supplementary Table 5. Study-Level Details on NMA Knowledge Outcomes

Study*	Outcomes
Qualitative Survey Responses Indicating NMA Knowledge	
Abdelhamid 2012	• 25% knew about IDC and had used it, 35% knew about IDC but had not used it, 27% heard about IDC but did not know much about it, 13% didn't know about IDC. • When considering if they would use indirect evidence when making a decision, 40% would consider evidence from IDC when making a decision, 36% would sometimes, 9% would not and 15% were not sure. • 81% of researchers, 58% of clinicians and 88% of individuals who were clinician-researchers said they "would" or "would sometimes" consider indirect evidence, • 88% individuals who had stated they knew about IDC, 61% who had heard about IDC but didn't know much, 33% who didn't know about IDC said they "would" or "would sometimes" consider indirect evidence
Coleman 2012	• 100% agree combining indirect and direct evidence provides valuable information not available from head-to-head trials • 67% agree combining indirect and direct evidence gives a more precise estimate. • 55% thought term 'NMA' used ambiguously • 44% considered 'MTC" to be more clear. • 56% thought current guidance on NMA conduct and reporting is sufficient • 78% thought NMAs should provide graphical depiction of evidence network. • No clear trends were identified re: if direct evidence takes precedence over indirect; if direct and indirect estimates both should always be considered and reported separately and if comparisons not of direct relevance (e.g. placebo, older treatments) should be considered in an NMA.
Inadequate Conduct or Interpretation of NMAs reported in Published Studies	
Bafeta 2013	• inadequate conduct of standard meta-analysis components, e.g. searches and data collection

Salanti 2011	• Misinterpretation of uncertainty in effect estimates
	  • could be prevented by use of 95% PrIs to identify uncertainty associated with effect estimates
	• Misinterpretation of uncertainty associated with treatment ranking claims
	  ○ could be prevented by use of ranking graphics to identify uncertainties

Song 2009	• inadequate statistical approaches used that don't preserve randomization • inadequate methods used to assess key NMA assumptions, e.g. heterogeneity, inconsistency • inadequate statistical handling of 3-arm trials
Abbreviations: IDC=indirect comparison; MTC=multiple treatment comparison; NMA=network meta-analysis; PrI=Predictive interval	

* None of the other 11 studies provided indicators of individuals' knowledge of NMAs

Supplementary Table 6. Study-Level Details of Barriers and Facilitators

Study	Barriers	Facilitators
Based on Survey and Focus Group Responses in Qualitative Studies		
Abdelhamid 2012	Lack of awareness of NMA methodology, no expertise to conduct an NMA, perception that indirect evidence is not valid or available evidence does not meet NMA assumptions, time, resources, incomparable trials and low quality of trials.	Training on NMA methods, interpretation of results or clarification of NMA assumptions.
Coleman 2012	Poor data quality, heterogeneity of data, volume of data, inability to form a network, lack of methodological knowledge, issues associated with ease of using software, availability of researchers with expertise, readers' ability to interpret NMAs, volume of data to digest, adequate reporting in manuscripts, acceptance of methods and knowing how to present results.	Rankings, ability to use both direct and indirect evidence, to analyze comparisons not conducted directly, high quality, practical factors, we do 'informal' MTC everyday, model estimation, multi-arm trial adjustment, handling uncertainty, impact, intuitive interpretation, growing in acceptance, comprehensive picture of the evidence, no adjustment for zero cells, thorough check of available evidence, novel, answers questions that are otherwise unanswerable, presentation of results with Bayesian inference.
Extracted from Included Study Publications and Grouped by Theme		
Based on All 16 Included Study Publications	**Complexity of NMAs** ○ basic assumptions more complicated than those underlying traditional meta-analysis ○ more complicated and larger (number of studies and therapies) statistical analyses ○ too complex ○ complexity of rating indirect estimates (i.e. more than first-order loops) increases workload but unlikely to change conclusions **Inconsistent Language** ○ terminology and frameworks for assumptions may vary **Evolving Methods** ○ Growing number of publications and new NMA methods ○ Variations in statistical methods approaches ○ inappropriate or no methods to evaluate inconsistency form unreliable basis for choosing best available therapies ○ infancy of some methods so hard to make recommendations ○ current ranking methods assume the treatment effects are constant over time ○ assumptions not clearly distinguished leading to methodological and practical problems in use and	**Able to Asses Assumptions** • Awareness of assumptions crucial to improve methodological reporting • able to assess if similar but not identical interventions grouped (e.g. different doses in one node) • graphical tools that help identify inconsistencies or inform judgments on plausibility of consistency • assessment of assumptions vital • clinical knowledge needed to truly assess similarity assumption - potential patient characteristics that could influence IDC results may only be known to those in clinical area • more discussion of underlying assumptions • develop methods for investigating heterogeneity • reasons for discrepancies between direct/indirect investigated • explicitly assess consistency assumption • 77% clarification on basic assumptions needed. • Involvement of both clinical/methodological experts for assessment of similarity/consistency assumptions

	interpretation of IDCs or MTCs. ○ is a work in progress ○ methodological uncertainties in NMAs remain **Lack of Reporting Guidance** ○ need for reporting guidance ○ no guidelines re: reporting/assessment of IDCs ○ variable reporting styles/lack of standard presentation formats creates interpretation challenges **Lack of Expertise** ○ need for more training ○ only for researchers with advanced computational and statistical knowledge ○ NMA methods still in realm of expert statisticians ○ not aware of NMA methods ○ didn't have expertise to conduct an NMA, ○ thought IDC evidence not valid ○ availability of researchers with expertise ○ lack of methodological knowledge, ○ issues associated with ease of using software **Critical Appraisal Challenges** ○ no guidelines re: reporting/assessment of IDCs ○ need for framework for peer-review of NMAs ○ knowing probability each treatment is best may not convey entire picture; appropriately interpreting/criticizing NMAs in peer-review literature ○ how to apply assessment to GRADE summary of findings table but optimum format still unclear ○ grading the evidence assumes the analysis was adequately conducted - adequacy of methods isn't assessed here ○ RoB assessments don't apply equally to the network and to rankings ○ ·need for more transparency and completeness of reporting ○ inadequate reporting may lower confidence in NMA results and conclusions ○ inadequately conducted or reported IDCs **Lack of Transparency** ○ lack of transparency and subjectivity on assessment of trial similarity, i.e. relies on expert opinion ○ adequate reporting in manuscripts	**Methodological Strengths of NMA** • transparent framework vs. back of envelope IDCs, can reduce uncertainty b/c inclusion of more data, preserves within trial randomization • ability to use both direct and indirect evidence • to analyze comparisons not conducted directly • we do 'informal' MTC everyday • handling uncertainty • high quality • answers questions that are otherwise unanswerable • appropriately analyze multi-arm trials • multi-arm trial adjustment • understand choices b/w frequentist and Bayesian methods. • no adjustment for zero cells • can be more reliable than direct evidence in some cases • strengths of Bayesian methods • methodological advantages • the adjusted indirect comparison can take advantage of the strength of RCTs in making unbiased comparisons **Having Reporting Standards** • consistent approach allows better understanding of evidence synthesis • Guidelines on reporting and conduct needed for NMAs • guidance for reporting NMA results • more clear guidance on conducting and reporting may lead to more consistent approaches • a guidance document for reporting may also be beneficial **Able to Understand Available Evidence** • Able to assess amount of evidence in network, relative effect sizes and uncertainty • knowing precise shape of network • knowing number of trials and patients b/c affects reliability of estimates • justifications for why trials excluded (e.g. placebo-

	Challenges Interpreting NMAs ○ choosing among direct, indirect, mixed evidence especially when discrepancies b/w direct and indirect estimates ○ variable reporting styles/lack of standard presentation formats creates interpretation challenges ○ Results not easy to interpret ○ assumptions too difficult to evaluate ○ provides outputs not easily understood and interpreted by clinicians ○ lack of tools to evaluate assumptions and present results ○ assumptions not clearly distinguished leading to methodological and practical problems in use and interpretation of IDCs or MTCs. ○ subjectivity in assessments but if apply quantitative approach can be misinterpreted, over interpreted ○ readers' ability to interpret NMAs ○ little attention to when to use indirect evidence, how to judge confidence in results **Challenges Presenting NMAs** ○ How to present and interpret results ○ knowing probability each treatment is best may not convey entire picture; appropriately interpreting/criticizing NMAs in peer-review literature ○ larger number of outcomes so more difficult graphical representation and interpretation of results ○ graphs can be over-interpreted or interpreted in isolation ○ current guidance documents provide few details and recommendations on presentation formats ○ variety of audiences so need different presentations ○ can't make prescriptive recs on presentation formats b/c depends on question of interest ○ no graphical presentation standards available for NMA ○ provides outputs not easily understood and interpreted by clinicians, ○ lack of user friendly implementation framework ○ lack of tools to evaluate assumptions and present results ○ variable reporting styles/lack of standard presentation formats creates interpretation challenges ○ figure and table formats challenging if large network and number of treatments	controlled, head-to-head) • systematic literature searches used • comprehensive picture of the evidence • thorough check of available evidence • intuitive interpretation • able to compare effect estimates from direct, indirect and mixed evidence • selects 'best available' evidence to focus on, i.e. direct, indirect, mixed • presenting direct and indirect evidence separately **Knowing Rankings and their Uncertainty** • should be able to examine uncertainty in rankings of interventions b/c differences may be small and not clinically relevant • providing comprehensive summaries of probabilities/rankings to avoid over-interpretation of probabilities associated with being the 'best' • ranking should be done using probabilistic methods to take into account uncertainty • rankings • presentation of results with Bayesian inference, **Having Critical Appraisal Standards** ○ need to consider confidence that can be placed in results ○ Able to assess amount of evidence in network, relative effect sizes and uncertainty ○ need to convey quality clearly so audience can make informed judgment on how to use results ○ selects 'best available' evidence to focus on, i.e. direct, indirect, mixed ○ builds on established GRADE framework **Having Guidance and Tools for Presenting NMAs** ○ guidance for reporting NMA results ○ presentational tools for reporting NMA results ○ allow built-in alternative display options for different audiences ○ more concise and informative presentation of results **Increased Transparency of NMAs** ○ using checklist will improve transparency

	○ knowing how to present results. **Volume of data and studies** ○ large number of studies and treatments ○ large volume of results in NMAs challenging to present ○ larger number of outcomes so more difficult graphical representation and interpretation of results ○ efficiently summarizing results from multiple pairwise comparison NMAs ○ simple methods are easy to use and transparent but inconvenient when lots of therapies ○ complexity of large number of competing treatments and interpretation of multiple effect sizes ○ figure and table formats challenging if large network and number of therapies ○ volume of data ○ journal limitations on figures/tables space **Inadequate conduct or acceptance of NMAs** ○ Inadequate SR methods can decrease confidence in NMA results ○ inadequate reporting may lower confidence in NMA results and conclusions ○ inappropriate statistical methods for IDCs ○ acceptance of methods AHRQ ○ inadequately conducted or reported IDCs **Evidence Base Inadequate** ○ trials included did not meet assumptions required for IDC ○ incomparable trials ○ low quality of trials ○ potential for effect modification ○ poor data quality ○ heterogeneity of data ○ inability to form a network **Time and Resources Required** ○ time and resources	○ transparency to allow readers to make their own judgment according to explicitly presented discrepancies ○ transparent framework vs. back of envelope IDCs, can reduce uncertainty b/c inclusion of more data, preserves within trial randomization **Acceptance of NMA Methods** ○ growing in acceptance, **More NMA Training** ○ methodological training **Better Evidence Base** ○ improve consistency and comparability among primary trials when designing RCTs (e.g. standardized outcome measures). **Standard Terminology** ○ promoting standardized terminology to improve understanding
Abbreviations: AHRQ=Agency for Healthcare Research and Quality; GRADE=Grading of Recommendations Assessment, Development and Evaluation; IDC=indirect comparisons; MTC=mixed treatment comparisons; NMA=network meta-analysis; RCT=randomized controlled trials; RoB=Risk of Bias; SR=systematic reviews;		

CHAPTER 6.

GUIDELINE STANDARDS FOR PRESENTING NETWORK META-ANALYSES: A SYSTEMATIC REVIEW

CHAPTER OVERVIEW

Purpose and Relationship to Other Manuscripts:

One of the goals of this thesis is to understand researchers' approaches to communicating and presenting network meta-analyses (NMAs). Therefore, the objective of this fourth manuscript is to systematically review guidelines for conducting and/or reporting NMAs for information on how to present NMAs. This work will be used to identify what guidance researchers, i.e. knowledge producers, are given on presentation formats for communicating NMAs. Consideration will also be give to how these findings differ from how researchers are currently presenting NMAs and from presentational formats that researchers have independently developed for presenting NMAs (as identified in Chapter 5, a systematic review of literature on knowledge, barriers, facilitators and approaches to communicating NMAs). This chapter is similar to Chapter 4, which provided an overview of guidelines for conducting economic evaluations.

Supporting Material:

None

Manuscript Status:

Submitted for review and awaiting a decision. All information in this chapter, including the supplementary materials, has been submitted to the journal.

Authors: Shannon M Sullivan (SMS), Doug Coyle (DC), George Wells (GW)

Author Roles (SMS, DC, GW) and Contributions:

SMS, DC, GW were responsible for the study conception.
SMS, GW were responsible for the search for trials.
SMS, GW were responsible for the selection of trials.
SMS, GW were responsible for the data extraction.
SMS, GW were responsible for the data analysis.
SMS, DC, GW were responsible for in the interpretation of results.
SMS, GW were responsible for drafting the manuscript.

All authors (SMS, DC, GW) provided comments on drafts and revisions to the manuscript, agree with the manuscript's results and conclusions, and read and approved the final manuscript.

Title: What Guidance are Researchers Given on How to Present Network Meta-Analyses? A Systematic Review.

Authors: Shannon M. Sullivan MSc [1] Doug Coyle MA, MSc, PhD [2] George Wells PhD [1,2]

Affiliations:

1 University of Ottawa Heart Institute, Ottawa, Ontario

2 University of Ottawa, Department of Epidemiology and Community Medicine, Ottawa, Ontario

ABSTRACT

Introduction: Standards for reporting and critically appraising network meta-analyses (NMAs) are important. However, consideration should also be given to identifying optimal presentation formats for NMA, i.e. determining not just 'what' to report but 'how' best to report. While good reporting practices are important to follow and can contribute to clarity, they may be insufficient for good communication to audiences unfamiliar with NMAs including other researchers or policymakers. This paper systematically reviews NMA guidelines for information on how to present NMAs.

Methods: Electronic databases and supplementary sources were searched for NMA guidelines. Presentation format details were extracted and frequencies tabulated.

Results: Seven guidelines were included. Current guidelines focus on how to conduct NMAs but provide limited guidance to researchers on how to best present analyses. None of the guidelines provided reporting templates. Few provided advice on tailoring presentations to different end-users, such as policymakers, despite many guidelines being developed in HTA organizations. Available guidance on presentation formats focused on evidence networks, characteristics of individual trials, comparisons between direct and indirect estimates and assumptions of heterogeneity and/or inconsistency. Some guidelines also provided examples of figures and tables that could be used to present information.

Conclusions: Limited guidance exists for researchers on how best to present NMAs in an accessible format for policymakers. This requires further thought, and integration with policymakers' needs, when NMAs are used to support health policy decision-making. Developing presentation formats that enhance understanding and accessibility of NMAs could also enhance the transparency and legitimacy of policy decisions informed by NMAs.

Introduction:

Transparency is a key principle underlying fair and legitimate health technology assessment (HTA) processes and related policy decisions (1). Transparency requires not only providing sufficient information, but also providing information in an accessible and understandable format for end-users. This is especially relevant when complex methods such as network meta-analyses (NMA) form the basis of a HTA.

In the absence of head-to-head trials of relevant comparators, NMAs frequently inform cost-effectiveness evaluations and therapeutic or drug class reviews (2). One frequently cited concern with NMAs is that the complex statistical methods do not permit the end user to understand how the results were obtained or if they are valid (3,4). Creating simple but accurate explanations of NMAs for policymakers, and those impacted by policy decisions such as clinicians and patients can be challenging. This is further complicated by the variable expertise among researchers in conducting and interpreting NMAs (3,5) and the rapidly evolving developments in NMA methods. (6-9) Although NMAs build on many of the same concepts as traditional meta-analyses, these similarities are often not recognized and the same critiques that meta-analyses once faced (e.g. heterogeneity and combining studies) are frequently applied to NMAs. (10)

Although work is ongoing to develop standards for reporting NMAs and for critically appraising NMAs,(8,11-12) consideration should also be given to identifying optimal presentation formats, i.e. determining not just 'what' to report but 'how' best to report it, and how to tailor information to different audiences who may be unfamiliar with NMAs. While good reporting practices are important to follow and contribute to clarity and transparency, they may be insufficient for good communication to non-technical audiences. Some studies have focused on helping end-users interpret NMAs, (13-14) but it is unclear what guidance researchers have on how to optimally present NMAs to these end-users.

A number of studies have surveyed current practices for presenting NMAs or explored different options for presenting NMAs (7,15-17). However, determining what guidance researchers are provided on how to present a NMA and determining how it aligns with end-users needs is also an important piece of this puzzle. This can contribute to developing optimal presentation formats and knowledge translation approaches applicable to NMAs. (18) A systematic approach to knowledge translation could improve the accessibility of HTAs that are based on NMAs, thereby enhancing the legitimacy of and confidence in health policy decision-making processes. (1) Therefore, this study systematically reviewed current guidelines for conducting or reporting NMAs to determine what guidance researchers are provided on how to present NMAs.

Methods:

Systematic Review: A systematic review was conducted following Cochrane methodology (19). Guidelines for conducting or reporting on network meta-analyses were included that targeted policy makers, statisticians, researchers, healthcare professionals or other end-users of these analyses. Because of the nascent field of NMA, articles that were not specifically developed as formal guidelines, but which provided direction on conducting or reporting network meta-analyses were also included (e.g. working group documents or interim guidance). Studies were excluded if they were: not guidance on conducting or reporting NMAs (e.g. guidance on how to interpret or critically appraise NMAs, editorials or opinion papers; original methodological articles on conducting NMAs), overviews or reviews of existing NMAs; guidelines not primarily focused on NMAs; or not the most recent or comprehensive versions of the guidelines. Outcomes of interest were the type and frequency of information on presentation formats.

Search Strategy: Databases searched included Medline (1996 to June 2014), EMBASE (1980 to June 2014) and the Cochrane Database of Systematic Reviews, using the earlier date limit of 2000 but no language restriction. The search concepts were ‘guidelines’, ‘indirect comparisons’,

'network meta-analyses' and 'multiple or mixed treatment comparisons'. Grey literature was searched for unpublished reports using the CADTH Grey Matters checklist as a guide, in addition to other relevant resources (20). Studies were also obtained through hand searching of selected journals and select authors, reviewing reference lists of potentially relevant studies and suggestions from experts in NMAs.

Study Selection: Citations were screened for relevance by one review author based on the title and abstract of identified articles. Two review authors independently reviewed the full text of potentially relevant guidelines to assess exclusion or inclusion.

Critical Appraisal: Guideline quality was assessed using the AGREE II instrument (21). AGREE II was designed, in part, to assess the quality of clinical practice guidelines. No instruments currently exist to assess methodological guidelines or guidelines for NMAs. Therefore, minor modifications to the AGREE II instrument were made for this study (see Supplementary Materials Table 1 for details).

Data Extraction: One author extracted data, which was verified by a second author and disagreements were resolved by a third reviewer. The following data were extracted from guidelines (see Supplementary Materials Table 2 for detailed variable definitions). Items related to presentation format were selected based on identification of key principles related to traditional meta-analyses and network meta-analyses (10) To be extracted, information was required on the format for presenting the information, not just that the information be provided or reported (i.e., focusing on 'how' to report not just 'what' to report)

(1) *Guideline characteristics*: guideline purpose, guideline scope, general or disease-specific, target audience, year, geographic region, author affiliations, has a reporting template, sample tables included, sample figures included, has a glossary.
(2) *Presentation formats*: details on how to present data sources (evidence networks, individual trial characteristics, critical appraisal of individual trials), analysis methods

(assumptions, heterogeneity and inconsistency, methodological concerns) and results (comparison with direct estimates, uncertainty, rankings, implications of findings).

Detailed guidance on how to conduct a NMA was not extracted (e.g. analytic approaches) and is not the focus of this systematic review.

Data Analysis: The frequency of recommendations for presenting NMAs was tabulated and common trends assessed.

Results:

Of the 1251 citations identified, 14 reports, representing 7 guidelines, were included (see Figure 1).(2, 22-34) Thirty reports were excluded, including a background document from the Cochrane Collaboration on the history of discussions within Cochrane on comparing multiple interventions in Cochrane Reviews, which will be used to inform future guidance for authors. (35)

INSERT FIGURE 1. PRISMA Flow Diagram for Systematic Review of NMA Guidelines

Key characteristics of the guidelines are outlined in Table 1 with the purpose of each guideline described in Supplementary Table 3.

INSERT TABLE 1. Characteristics of Network Meta-Analysis Guidelines

All guidelines provided guidance on how to conduct NMA and three (43%) also provided guidance on reporting NMA. (22-31). None of the guidelines were directed to any specific disease area or intervention and were generally applicable across all health technologies. Most guidelines (n=5, 71%) were developed by HTA organizations around the world with the others

developed by collaboratives with an interest in HTA, including the International Society For Pharmacoeconomics and Outcomes Research (ISPOR) and the European network for Health Technology Assessment (EUnetHTA). The target audience for all guidelines was researchers and decision-makers; two guidelines (29%) also identified health care professionals as part of the target audience. (29-31) Many guidelines (n=5, 71%) specifically acknowledged how decision-makers or other non-technical audiences use NMAs. However, very few provided specific guidance regarding how to present NMAs to non-technical end-users. In one of the few examples of this, one guideline suggested converting outcomes to measures decision-makers might prefer such as RR, ARR, NNT. (29-30) None of the guidelines provided a full glossary of technical terminology used in NMAs, however three of the guidelines (43%) provided some definitions. (29-30,32,34)

The earliest guideline was published in 2008. (34) Most guidelines recognized the rapidly evolving field of NMA and noted that updates would be required and guidelines would be monitored for these changes. Guideline quality was generally low when critically appraised. Common limitations were related to narrow stakeholder involvement, lack of systematic development and limited details related to implementation and applicability of guidelines. Although recommendations were clearly identified in most guidelines, differences in terminology when comparing across guidelines may create confusion and lead to lack of clarity. More details on the critical appraisal are provided in Supplementary Table 4.

Presentation Formats Identified in Recommendations, Sample Figures and Sample Tables

Of the seven guidelines that met the inclusion criteria, only four (57%) provided recommendations or guidance on presentation formats, as described in more detail in sections below. None provided an example template for how to report NMA. Although actual recommendations on how to present NMAs were not provided in all guidelines, some example figures (n=5, 71%) and

tables (n=4, 57%) were provided when illustrating how to conduct NMAs. Sample figures and tables were generally related to presenting data sources (e.g. evidence network diagrams, trial characteristics tables) or results (e.g. forest plots, tables comparing direct and indirect estimates). These presentation formats are described in more detail below as they relate to each section. The different areas of NMAs for which presentation formats were identified or recommended in guidelines and are outlined in Table 2.

INSERT TABLE 2. NMA Areas with Frequency of Presentation Formats Identified in Guidelines

Presenting Data Sources/Included Data

Three guidelines (43%) made recommendations on presenting evidence networks. (22-30,34) All three recommended a graphical schematic of the evidence structure, which are often referred to as 'network diagrams'. These diagrams outline relationships between the included studies and where direct and indirect evidence exists between therapies in the network. Diagram characteristics in the sample figures (n=3, 43%) generally included features such as the use of solid lines for direct evidence and dashed lines for indirect evidence relationships; providing on the connecting line between two therapies the name, number of trials or direct results contributing to a comparison; arrowheads indicating which therapy is favoured in the comparison; and labeling or identifying the network geometry, e.g. star shapes, closed loops. However, specific recommendations on which features to include in the network diagram were not provided in any of the guidelines.

Two of the guidelines recommended flow diagrams outlining included and excluded studies and/or tables or lists identifying included and excluded studies. (22-30) When presenting lists or tables of included and excluded studies, clarifying which studies were identified in the systematic review versus which studies had sufficient information to be included in the network meta-analysis was requested. (22-30)

Four guidelines (57%) made recommendations on how to present details of individual trials. Two guidelines recommended presenting information in a table format (22-30) while the other two indicated it could be discussed in the text. (31, 34) Individual trial details included what treatments were compared, trial level data used in the analysis, trial level covariate values, or if individual participant data are available. (22-28) Other factors that could be effect modifiers were recommended for inclusion in tables such patient age, length of time with a disease, history of treatment or geographic region (ISPOR). Two guidelines only made broad recommendations that sufficient study-level characteristics be provided and discussed. (31, 34) One guideline also specifically recommended presenting an assessment of individual trial quality as discussion in the text. (22-28)

When considering sample formats that were sometimes provided in guidelines (n=4, 57%), additional formats and characteristics were noted. For example, trial sponsor, population, therapies and doses, trial duration, size of treatment group, primary and secondary endpoints, blinding and study conclusions were included in one sample table (33) Three guidelines also recommended a table where each row is a study and each column is a treatment, with cells populated with absolute frequencies from the trials, which can clarify where direct evidence exists in the network while also demonstrating the raw data included in the analysis. (2,22-30) In one guideline, additional columns reporting relevant trial characteristics were also added to the table. (22-30)

Presenting Analysis Methods

Two guidelines (29%) recommended presenting information on general assumptions and on general methodological concerns as a description in the text. (22-30) When looking at more specific assumptions, four guidelines (57%) made recommendations on how to present issues related to heterogeneity and inconsistency. Multiple formats for presenting this information were identified including text descriptions, graphics and numerical estimates. For example, NICE

guidance suggested graphics generated by 'node splitting', numerical estimates of the degree of heterogeneity, discussion of extent and sources of heterogeneity and plotting posterior mean deviances to identify inconsistencies. (22-28) Another guideline suggested presenting sensitivity analyses of including/excluding trials and providing a discussion of potential sources of heterogeneity. (34) Two other guidelines focused on presenting descriptions that included explicit statements and step-by-step description of analyses (29-30) or distinguishing between clinical, methodological and statistical heterogeneity. (31)

When considering sample formats presented in guidelines (n=1, 14%), various plots were included that would allow one to explore methodological assumptions when conducting an NMA such as leverage plots, density plots, interaction plots and residual deviance plots. (22-28) Sample table formats were also identified that included measures of model fit and heterogeneity, summary results for consistency and inconsistency models and that outlined results from different sensitivity analyses with and without covariate interactions.

Presenting Results

Four guidelines (57%) recommended how to present and compare direct estimates with indirect or mixed estimates. Both tables formats and forest plots were recommended in two guidelines but there was no preference noted for one format over the other. (22-30) Generally, these formats were recommended so as to allow easy visual comparison of different results, e.g. outcomes by treatment arm, absolute and relative effect measures from trials, pairwise meta-analysis results and/or pooled NMA results. (22-30) Presenting results using a common reference standard was noted in two guidelines (22-30) and one guideline also recommended that all relevant pairwise comparisons should also be provided. (29-30) Two guidelines also recommended the use of multiple estimators to present results (e.g., RR, OR, RD, NNT). (29-30, 34) Tables providing relative treatment effects alongside estimates of heterogeneity were recommended in one

guideline. (22-28) One guideline recommended a discussion of why the estimates differed but did not specify more details. (31)

When considering sample formats that were sometimes provided in guidelines, additional formats were noted. Four guidelines (57%) provided sample tables that outlined direct, indirect and/or mixed estimates in table columns while each row represented a comparison of interest. (2, 22-30,33) Other types of results were also compared in this format (e.g. analytic approaches such as Bayesian vs. frequentist or different sensitivity analyses). Two of these guidelines also provided samples of forest plots that included direct, indirect and mixed estimates and allowed for easy visual comparisons. (29-30, 33)

Three guidelines (43%) recommended presenting the uncertainty of estimates as credible intervals or confidence intervals around the point estimate, depending on the statistical analysis approach. (22-30, 34) In guidelines that provided sample figures or tables (n=4, 57%), these intervals were usually provided in table columns or as a component of forest plots, along with the point estimate.

Two guidelines (29%) recommended presenting treatment rankings as either graphics or tables. (22-30) There was also considerable discussion in the guidelines on the challenges of presenting information on rankings so that it is not misinterpreted. For example, ensuring the information on the spread of rankings is provided (29-30) and that the probabilities of being best, second best, etc. are calculated was recommended in both guidelines. (22-30) One guideline specifically suggested presenting this information in the format of a rank-o-gram, which can incorporate multiple outcomes in one display. (22-28) Other possible formats published in peer-review literature were also recognized to be useful at times. (22-28)

Only one guideline provided a sample format of how to present rankings. (29-30) The guideline provided a table with each row being a different treatment versus the reference comparators and columns of results (ORs, % CrI) and columns of rankings and different probabilities of being best.

Two guidelines (29%) made recommendations on presenting the implications of findings as descriptions in the text. One guideline specifically referred to the implication of assumptions on results (22-28) The other identified implications with respect to validity of results; expectations compared with existing evidence, clinical rationale or biological rationale; relevance to real-world clinical and policy decisions; and the extent of possible bias and if it could lead to a different conclusion than a non-biased analysis. (29-30)

Discussion:

Key Findings

Current NMA guidelines focus on how to conduct analyses but provide limited guidance to researchers on how to best present analyses. None of the guidelines provided reporting templates. Few provided specific advice on tailoring presentations to different end-users or extensive glossaries for non-technical readers, despite many guidelines being developed in HTA organizations for use, in part, by policymakers. Only four of the seven included guidelines provided advice on presentation formats or provided sample figures and tables to guide presentation of the NMA. This guidance focused primarily on presenting evidence networks, characteristics of individual trials, comparisons between direct and indirect estimates and assumptions of heterogeneity and/or inconsistency.

Comparison with Other Literature

Although reporting guidelines exist in other fields (e.g., economic evaluations, systematic reviews), due to the relatively innovative nature of NMAs, guidelines for the reporting and

conducting of NMAs are less well-developed. Currently, NMA guidelines focus on defining standard methodological approaches and increasing the consistency and transparency of how NMAs are conducted. Corroborating this, two other reviews of NMA guidelines were identified but they compared methodological recommendations across guidelines and did not explore best practices of how to present NMAs. (36-37) With the growing need for more expertise and understanding of NMAs determining how best to present NMAs has taken on greater prominence. (3) The limited advice provided in the existing NMA guidelines suggests a lack of consensus on how researchers present NMAs or tailor NMAs to non-technical end-users such as policymakers. However, the areas that were frequently considered in existing guidelines provide a starting point for focusing this guidance:

- **Inclusion of trial details and network diagrams**. This will also align with policymakers' need to understand the applicability of the analysis to their specific context. Providing graphical schematics (i.e., evidence network diagrams) may also help audiences better understand the relationships among the included studies. Looking at the shape of the network itself (e.g. star diagram, number of closed loops) can provide information on characteristics of the included evidence. (37,38) In addition, there are many other styles of network diagrams that have been reported in literature and other trial characteristics that can be included (e.g. sample size, risk of bias). (11) Even though many guidelines provided examples of network diagrams, few guidelines made specific recommendations on their design and more guidance should be considered on the optimal presentation of evidence network diagrams.
- **Comparison between direct and indirect results.** Questions often arise on why direct and indirect estimates may differ from each other and explanations could assist end-users. Although, NMA results may be complex, building on familiar approaches used to present traditional meta-analyses (e.g. forest plots) may provide end-users and researchers with greater comfort with and understanding of how to interpret NMAs. (10)
- **Heterogeneity and inconsistency assumptions.** Identifying presentation formats that provide audiences with clarity on the assumptions made in the analysis and help end-users

understand their validity would be of value. Critical appraisal checklists have been developed and could also assist end-users understand an NMAs robustness. (8) Although other assumptions are important in conducting NMAs (e.g. goodness of fit, random effects), different methods of presenting the validity of these assumptions were not often noted.

Further developing guidelines for researchers on how to best present these aspects of NMAs to end-users of NMAs such as policymakers or others who want to enhance their understanding of NMAs should be considered in light of: (a) policymakers preferences when presenting evidence syntheses; (b) common principles related to the presentation of traditional meta-analyses; (c) how researchers currently present NMAs; and (d) common limitations when conducting NMAs.

a) Although policymakers may have limited experience with NMAs, they have often expressed views on how to present systematic reviews and traditional meta-analyses. Studies have reported that policymakers prefer presentation formats that can be quickly scanned, emphasize the bottom-line conclusions and clearly identify real-world implications (39-41). Policymakers have also expressed a desire for understanding if the review is robust enough from a methodological perspective, to support decision-making. (12) Incorporating policymakers' preferences at this stage in the development of NMA guidelines would narrow the gap between policymakers and researchers when using and providing NMAs. This would also allow for the creation of NMAs that adequately balance both rigour and accessibility for those who must apply their results and understand their implications in the 'real-world'. A recent ISPOR task force partially addresses this gap by developing a questionnaire for decision-makers to assess the relevance and credibility of NMAs. (12)

b) NMAs are based on many of the same principles as traditional meta-analyses. (10,42). Leveraging some of these same principles may be one approach that could help end-users better understand and appreciate NMAs. For example, conducting systematic searches for evidence, verifying homogeneity of populations when combining individual study results and

presenting results with appropriate measures of precision or uncertainty are principles applicable to both traditional meta-analyses and NMAs.

c) Current presentation of NMAs by researchers is extremely variable. For example, Tan et al (2013) reviewed 19 indirect treatment comparisons (IDCs) from NICE health technology assessments and found that researchers most frequently presented evidence network diagrams, model descriptions and tables and forest plots of results. (15) Based on this work approaches that may be most useful for non-technical audiences were identified and standardized graphical tools were developed that compare direct with indirect results and that summarize rankings. (16) An analysis from Chiamani et al (2013) has also proposed new graphical options for presenting NMA methods and results. (7) For example, network diagrams that incorporate key trial characteristics and risk of bias assessments and contribution plots that emphasize how much direct versus indirect evidence is available in the analysis. Evaluating these different tools and formats to determine if they influence use or understanding of NMAs would be beneficial.

d) Common limitations in NMAs may need to be presented very clearly for non-technical audiences. In a survey of 88 indirect comparisons, limitations included: an unclear understanding of underlying assumptions; inappropriate search for or selection of included trials; inappropriate or flawed methods; trial similarity not being objectively assessed; and an inappropriate combination of direct and indirect evidence. (43) A report on indirect comparisons submitted to NICE reported similar results, finding that common flaws were related to systematic review search methods, inappropriate pooling of heterogeneous data and suboptimal statistical methods. (44) More consistent reporting standards and optimized presentation formats in these areas may result in more clarity around expectations when conducting NMAs and ultimately enhance their quality.

Strengths and Limitations

This is the first work to systematically review the guidance that researchers are provided on how to present NMAs and consider this in the context of end-users of NMAs, including policymakers and other non-technical audiences who are impacted by policy decisions informed by NMAs. Identification of variables to extract was grounded in theoretical concepts of how NMAs relate to traditional meta-analyses. (10) Using this approach and building on these concepts could provide one avenue to facilitate understanding of NMAs among audiences familiar with traditional meta-analyses. Methodological guidelines do not exist to critically appraise the quality of NMA guidelines and a tool for critically appraising clinical practice guidelines was adapted. Therefore, concepts unique to methodological guidelines may have been missed, however, the appraisal allowed identification of some key issues in the guidelines' development. Extracting qualitative data from various guideline documents requires judgment. The heterogeneity of guideline formats and objectives created challenges in data synthesis, however, common factors such as their development by HTA organizations and consideration of both researchers and policymakers suggests they have a similar overall intent. Also, the limited information available on presentation formats and broad data definitions applied in this review suggests that misclassification of data was unlikely to occur, e.g. it was unlikely that a figure presenting data sources would be missed or misclassified as a figure presenting results. A comprehensive search of the grey literature, electronic databases and other sources allowed for identification of a broad set of possible NMA guidelines. Although partial guidelines were excluded from this review, the novel but evolving nature of NMA guidance suggests they were not likely to yield additional data.

Policy Implications/Interpretation

Some policy implications should be considered in light of this NMA guideline review and are summarized below:

(1) As methodological standards become more clear for conducting NMAs, guidelines should begin to consider how best to present NMAs. Researchers themselves have identified

areas of confusion related to NMAs and non-technical audiences are likely in need of greater support in understanding NMAs.

(2) The focus on developing presentation formats should be in areas that were most frequently raised in current NMA guidelines and are important to researchers, including: trial details and evidence network diagrams; comparisons between direct and indirect results; and heterogeneity and inconsistency assumptions.

(3) It is important that presentation formats align with policymakers needs and are tailored to a non-technical audience, given that policymakers are important end-users of NMAs and the considerable challenges users and researchers have had in appropriately conducting and interpreting NMAs.

Future Research and Conclusions

Knowledge translation approaches have frequently been applied to enhancing understandability and accessibility of clinical trial evidence for policy makers, healthcare providers and patients (45-47). Tailored knowledge translation approaches have presented evidence syntheses in formats such as 1000 faces in decision aids and key messages in briefing notes tailored to policy makers (45-46). Expansion of these approaches to NMAs has not yet occurred, but could be applied to develop technically accurate but simplified explanations of NMAs for policy makers or other non-technical audiences. For example, in addition to providing technical reports and scientific publications of NMAs, decision support tools could be developed. Tools could also be used to inform other educational initiatives or supplementary resources for policy makers and other general audiences to understand NMAs. Development of presentation formats will allow policy makers will be able to better apply the results of NMAs and enhance the transparency and legitimacy of HTA-informed decision-making processes.

Researchers' current focus is on how to present details of included trials and the evidence network; comparisons between direct and indirect estimates; and assumptions of heterogeneity and inconsistency. However, limited guidance exists for researchers on how best to present

NMAs in an accessible format for decision-makers and requires further thought, and integration with decision-makers needs, if NMAs are going to be used to strengthen and support policy decision-making.

ACKNOWLEDGEMENTS:

None

REFERENCES

1. Daniels N & Sabin J. Limits to health care: Fair procedures, democratic deliberations and the legitimacy problem for insurers. Philosophy & Public Affairs 1997; 303-350.

2. CADTH Health Technology Assessment Report. Indirect Evidence. Indirect treatment comparisons in meta-analysis. Ottawa: Canadian Agency for Drugs and Technologies in Health; 2009.

3. Abdelhamid A, Loke Y, Parekh-Bhurke S, Chen Y, Sutton A, Eastwood A, Hollanda R, Song F. Use of indirect comparison methods in systematic reviews: a survey of Cochrane review authors. Research Synthesis Methods 2012, 3: 71-79.

4. Pocock SJ. Safety of drug-eluting stents. Demystifying network meta-analysis. Lancet 2007; 370:2099-100.

5. 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.

6. Donegan S, Williamson P, D'Alessandrob U, Smith C. Assessing key assumptions of network meta-analysis: a review of methods. Research Synthesis Methods 2013, 4: 291–323.

7. Chaimani A, Higgins JP, Mavridis D, Spyroidonos P, Salanti G. Graphical tools for network meta-analysis in STATA. PLoS One 2013, 8 (10): e76654.

8. Salanti G, Del Giovane C, Chaimani A, Caldwell DM, Higgins JP. Evaluating the quality of evidence from a network meta-analysis. PLoS One 2014, 9(7): e99682.

9. Veroniki A, Vasiliadis H, Higgins JP, Salanti G. Evaluation of inconsistency in networks of interventions. Int J Epidemiol 2013, 42(1): 332-345.

10. Salanti G. Indirect and mixed-treatment comparison, network, or multiple-treatments meta-analysis. Many names, many benefits, many concerns for the next generation evidence synthesis tool. Research Synthesis Methods 2012; 3:80-97.

11. Hutton B, Salanti G, Chaimani A, Caldwell DM, Schmid C, Thorlund K et al. The quality of reporting methods and results in network meta-analyses. An overview of reviews and suggestions for improvement. PLoS One 2014, 9(3): e92508.

12. Jansen JP, Trikalinos T, Cappelleri JC, Daw J, Andes S, Eldessouki R et al. Indirect treatment comparison/network meta-analysis study questionnaire to assess relevance and credibility to inform health care decision making: an ISPOR-AMCP-NPC Good Practice Task Force report. Value in Health 2014, 17(2): 157-173.

13. Mills E, Ionnidis J, Thorlund K, Schunemann H, Pulan M, Guyatt G. How to use an article reporting a multiple treatment comparison meta-analysis. JAMA 2012; 308: 1246-1253.

14. Cipriani A, Higgins JP, Geddes JR, Salanti G. Conceptual and technical challenges in network meta-analysis. Ann Intern Med 2013, 159 (2): 130-137.

15. Tan SH, Bujkiewicz S, Sutton A, Dequen P, Cooper N. Presentational approaches used in the UK for reporting evidence synthesis using indirect and mixed treatment comparisons. J Health Serv Res Policy 2013, 18(4): 224-232.

16. Tan SH, Cooper NJ, Bujkiewicz S, Welton NJ, Caldwell DM, Sutton AJ. Novel presentataional approaches were developed for reporting network meta-analysis. J Clin Epidemol 2014, 67(6): 672-680.

17. Donegan S, Williamson P, Gamble C, Tudor-Smith C. Indirect comparisons. A review of reporting and methodological quality. PLoS One 2010, 5(11):e11054.

18. Straus S, Tetroe J & Graham ID (Eds.), Knowledge translation in health care: Moving from evidence to practice, 2nd edition. Oxford: Blackwell Publishing, 2013.

19. Cochrane Collaboration: Cochrane Handbook for Systematic Reviews. Available from: www.cochrane.org [Last accessed May 19, 2014].

20. CADTH. Grey Matters: a practical search tool for evidence-based medicine [2nd edition]. Ottawa: Canadian Agency for Drugs and Technologies in Health; 2013.

21. Brouwers MC, Kho ME, Browman GP, Burgers JS, Cluzeau F, Feder G, Fervers B, Graham ID, Grimshaw J, Hanna SE, Littlejohns P, Makarski J, Zitzelsberger L; AGREE Next Steps Consortium. AGREE II: advancing guideline development, reporting and evaluation in health care. J Clin Epidemiol. 2010 Dec;63(12):1308-11. doi: 10.1016/j.jclinepi.2010.07.001.

22. Dias, S., Welton, N.J., Sutton, A.J., Ades, A.E. NICE DSU Technical Support Document 1: Introduction to evidence synthesis for decision making. 2011; last updated April 2012; available from http://www.nicedsu.org.uk

23. Dias, S., Welton, N.J., Sutton, A.J., Ades, A.E. NICE DSU Technical Support Document 2: A generalised linear modelling framework for pair-wise and network meta-analysis of randomised controlled trials. 2011; last updated April 2012; available from http://www.nicedsu.org.uk

24. Dias, S., Sutton, A.J., Welton, N.J., Ades, A.E. NICE DSU Technical Support Document 3: Heterogeneity: subgroups, meta-regression, bias and bias-adjustment. 2011; last updated April 2012; available from http://www.nicedsu.org.uk vaila

25. Dias, S., Welton, N.J., Sutton, A.J., Caldwell, D.M., Lu, G. & Ades, A.E. NICE DSU Technical Support Document 4: Inconsistency in Networks of Evidence Based on Randomised Controlled Trials. 2011; last updated April 2012; available from http://www.nicedsu.org.uk vaila

26. Dias, S., Welton, N.J., Sutton, A.J., Ades, A.E. NICE DSU Technical Support Document 5: Evidence synthesis in the baseline natural history model. 2011; last updated April 2012; available from http://www.nicedsu.org.uk

27. Dias, S., Sutton, A.J., Welton, N.J. & Ades, A.E. NICE DSU Technical Support Document 6: Embedding evidence synthesis in probabilistic cost-effectiveness analysis: software choices. 2011; last updated April 2012; available from http://www.nicedsu.org.uk

28. Ades, A.E., Caldwell, D.M., Reken, S., Welton, N.J., Sutton, A.J., Dias, S. NICE DSU Technical Support Document 7: Evidence synthesis of treatment efficacy in decision making: a reviewer's checklist. 2012; available from http://www.nicedsu.org.uk

29. Jansen J, Fleurence R, Devine B, Itzler R, Barrett A, Hawkins N, Lee K, Boersma C, Annemans L, Cappelleri J (2011). 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 in Health 14: 417–428.

30. Hoaglin D, Hawkins N, Jansen J, Scott D, Itzler R, Cappelleri J, Boersma C, Thompson D, Larholt K, Diaz M, Barrett A (2011). 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 in Health 14:429-43.

31. 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/.

32. European Network for Health Technology Assessment (EUnetHTA). Guideline. Comparators & Comparisons: Direct and indirect comparisons. Final version, February 2013. Available:http://www.eunethta.eu/sites/5026.fedimbo.belgium.be/files/Direct%20and%20indirect%20comparisons.pdf

33. Haute Autorite de Santé (HAS). Summary Report. Indirect comparisons, Methods and validity. Final version, July 2009. Available: http://www.has-sante.fr/portail/upload/docs/application/pdf/2011-02/summary_report__indirect_comparisons_methods_and_validity_january_2011_2.pdf

34. Report of the Indirect Comparisons Working Group (ICWG) to the Pharmaceutical Benefits Advisory Committee: Assessing Indirect Comparisons. Australian Government, Department of health and Ageing; Dec 2008 [Last accessed July 7, 2014]. Available at:. http://www.pbs.gov.au/industry/useful-resources/PBAC_feedback_files/ICWG%20Report%20FINAL2.pdf

35. Becker L, Caldwell D, Higgins J, Li T, Salanti G, Schmid C. Comparing Multiple Interventions in Cochrane Reviews. 2013. Available:http://cmimg.cochrane.org/sites/cmimg.cochrane.org/files/uploads/Comparing%20Multiple%20Interventions%20in%20Cochrane%20Reviews%20-%202003%2003%2023.pdf

36. Hawkins N, Laws AJ, Kendall RM. Review and comparison of different national guidelines on the implementation of network meta-analysis. Value in Health. Conference ISPOR 17th Annual International Meeting Washington, DC United States 2012, 15(4): A33.

37. Salanti G, Kavvoura FK, Ioannidis JP. Exploring the geometry of treatment networks. Ann Intern Med. 2008 Apr 1;148(7):544-53.

38. Nikolakopoulou A. Chaimani A. Veroniki AA. Vasiliadis HS. Schmid CH. Salanti G. Characteristics of networks of interventions: a description of a database of 186 published networks. PLoS ONE 2014, 9(1):e86754.

39. Murthy L, Shepperd S, Clarke MJ, Garner SE, Lavis JN, Perrier L, Roberts NW, Straus SE. Interventions to improve the use of systematic reviews in decision-making by health system

managers, policy makers and clinicians. Cochrane Database of Systematic Reviews 2012, Issue 9. Art. No.: CD009401. DOI: 10.1002/14651858.CD009401.pub2

40. Lavis, J., Davies, H., Oxman, A., Denis, J. L., Golden-Biddle, K., & Ferlie, E. (2005). Towards systematic reviews that inform health care management and policy-making. Journal of Health Services Research and Policy, 10 (Suppl 1), 35-48.

41. Innvaer, S., Vist, G., Trommald, M., & Oxman, A. (2002). Health policy-makers' perceptions of their use of evidence: A systematic review. Journal of Health Services Research and Policy, 7, 239-244.

42. Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. Ann Intern Med, 151(4).

43. Song F, Loke YK, Walsh T, Glenny AM, Eastwood AJ, Altman DG. Methodological problems in the use of indirect comparisons for evaluating healthcare interventions: survey of published systematic reviews. BMJ 2009, 338: b1147.

44. Martin A, Rizzo M, Iheanacho I. Faulty connections. Can criticisms of network meta-analyses in NICE submission be avoided? Value in Health Conference ISPOR 16th Annual European Congress Dublin Ireland, November 2013, 16(7): A608.

45. Chambers et al. (2011). Maximizing the impact of systematic reviews in health care decision making. A systematic scoping review of knowledge translation resources. Milbank Q, 89: 131-156.

46. Trevena L, Davey H, Barratt A, Butow P, Caldwell P. A systematic review on communicating with patients about evidence. J Evalu Clin Practice 2006;12: 13-23.

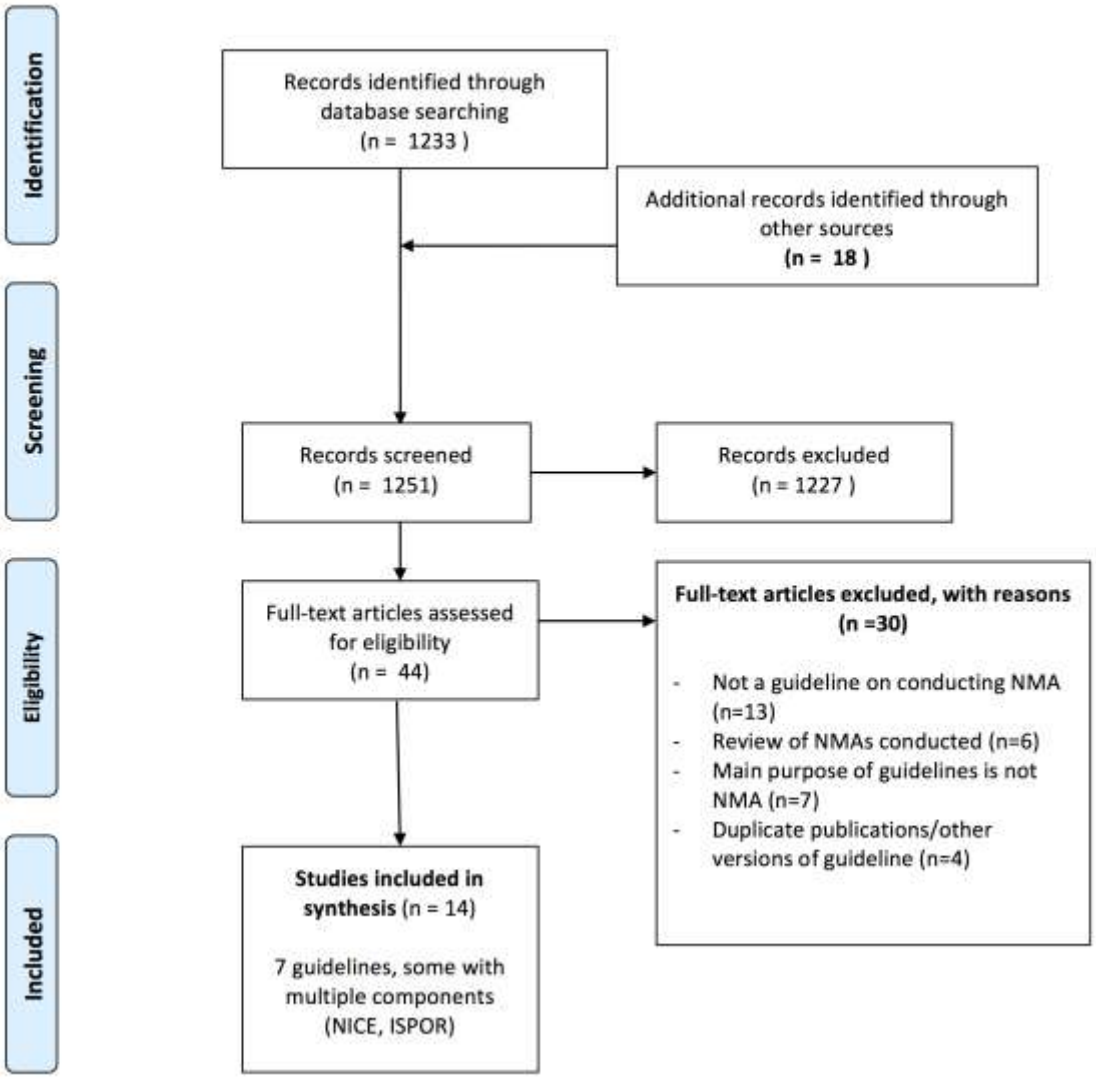
47. Akl EA, Guyatt GH, Irani J, Feldstein D, Wasi P, Shaw E, Shaneyfelt T, Levine M, Schünemann HJ. "Might" or "suggest"? No wording approach was clearly superior in conveying the strength of recommendation. J Clin Epidemiol 2012; 65: 268-275.

TABLES AND FIGURES

FIGURE LEGENDS

Figure 1. PRISMA Flow Diagram for Systematic Review of NMA Guidelines.

Figure 1. PRISMA Flow Diagram for Systematic Review of NMA Guidelines



TABLES

Table 1. Characteristics of Network Meta-Analysis Guidelines

Guideline	Geographic Region	Scope	Presentation Formats Recommended*	Sample Figures*	Sample Tables*	Reporting Template	Glossary or Definitions Provided	Acknowledging Non-Technical End-Users**
ISPOR 2011	International Collaboration	Reporting and Conducting	Yes	Yes	Yes	No	Yes	Yes
CADTH 2009	Canada	Conducting	No	Yes	No	No	No	Yes
NICE DSU Series 2011	UK	Reporting and Conducting	Yes	Yes	Yes	No	No	Yes
PBAC 2008	Australia	Conducting	Yes	No	No	No	Yes	Yes
HAS 2009	France	Conducting	No	Yes	Yes	No	No	No
AHRQ 2010	USA	Reporting and Conducting	Yes	No	No	No	No	Yes
EUnetHTA 2013	European Collaboration	Conducting	No	Yes	No	No	Yes	No

Abbreviations: AHRQ=Agency for Healthcare Research and Quality; CADTH=Canadian Agency for Drugs and Technologies in Health; EUnetHTA=European network for Health Technology Assessment; HAS=Haute Autorite de Santé; ISPOR=International Society for Pharmacoeconomics and Outcomes Research (ISPOR); NICE=National Institute for Health and Clinical Excellence; PBAC=Pharmaceutical Benefits Advisory Committee

*Although actual recommendations on how to present NMAs were not provided in all guidelines, some example figures and tables were provided when illustrating how to conduct NMA, which could inform how to present NMAs

** Although most guidelines acknowledged there were non-technical end-users of NMAs, only one (ISPOR) provided specific guidance on how to present information to them.

Table 2. NMA Areas with Frequency of Presentation Formats Identified in Guidelines

Areas of NMA	**Recommendations on Presentation Formats n (%), N=7**	**Formats Presented in Sample Figures or Tables n (%), N=7**
Included Data		
Trial Network	3 (43)	3 (43)
Individual Trial Characteristics	4 (57)	4 (57)
Critical Appraisal	1 (14)	0
Methods		
Assumptions	2 (29)	1 (14)
Heterogeneity and/or Inconsistency	4 (57)	1 (14)
Methodological Concerns	2 (29)	0
Results		
Comparison of Direct and Indirect Effects	4 (57)	4 (57)
Uncertainty	3 (43)	4 (57)
Rankings	2 (29)	1 (14)
Implications of Findings	2 (29)	0

SUPPLEMENTARY MATERIAL

SUPPORTING INFORMATION

Supplementary Table 1. AGREE II (Unmodified)* for Critical Appraisal of Clinical Practice Guidelines

* To apply AGREE II to the critical appraisal of methodological guidelines for network meta-analyses (rather than to clinical practice guidelines), the following minor modifications were made:

- For item 2 rather than 'health' question, the 'methodological question(s) covered by the guideline is (are) specifically described.
- For item 11 rather than 'health benefits, side effects and risks/, the 'advantages and disadvantages have been considered in formulating the recommendations'
- Item 16 was removed: 'the different options for management of the condition or health issue are clearly presented'

Domain	Item
Scope and purpose	1. The overall objective(s) of the guideline is (are) specifically described.
	2. The health question(s) covered by the guideline is (are) specifically described.*
	3. The population (patients, public, etc.) to whom the guideline is meant to apply is specifically described.
Stakeholder involvement	4. The guideline development group includes individuals from all the relevant professional groups.
	5. The views and preferences of the target population (patients, public, etc.) have been sought.
	6. The target users of the guideline are clearly defined.
Rigor of development	7. Systematic methods were used to search for evidence.
	8. The criteria for selecting the evidence are clearly described.
	9. The strengths and limitations of the body of evidence are clearly described.
	10. The methods for formulating the recommendations are clearly described.
	11. The health benefits, side effects and risks have been considered in formulating the recommendations.*
	12. There is an explicit link between the recommendations and the supporting evidence.
	13. The guideline has been externally reviewed by experts prior to its publication.
	14. A procedure for updating the guideline is provided.

Clarity of presentation	15. The recommendations are specific and unambiguous.
	16. The different options for management of the condition or health issue are clearly presented.*
	17. Key recommendations are easily identifiable.
Applicability	18. The guideline describes facilitators and barriers to its application.
	19. The guideline provides advice and/or tools on how the recommendations can be put into practice.
	20. The potential resource implications of applying the recommendations have been considered.
	21. The guideline presents monitoring and/ or auditing criteria.
Editorial independence	22. The views of the funding body have not influenced the content of the guideline.
	23. Competing interests of guideline development group members have been recorded and addressed.

Supplementary Table 2. Definitions of Variables and Data Extracted in Systematic Review

Variable	Definition
Guideline Characteristics	
Audience	Who is the intended audience for the guideline
Scope	Was the scope of the guideline related to conducting NMAs, reporting NMAs or both
Author Affiliations	Affiliations of guideline developers and authors including professional or academic collaborations, health technology assessment organizations, academic institutions, industry.
Presentation Formats	
Data Sources/Included Data Formats	
Trial Network	How to present the relationships between data points and structure of the evidence included in the NMA
Individual Trial Characteristics	How to present characteristics of the direct evidence the network is built upon
Risk of Bias	How to present methodological characteristics of the direct evidence that the network is built upon
Analysis Methods Formats	
Assumptions	How to present assumptions that are made and/or required for a NMA
Heterogeneity and Inconsistency	How to present potential sources of heterogeneity and/or inconsistency or any results evaluating these assumptions
Methodological Concerns	How to present any specific methodological issues that arise in an NMA
Results Formats	
Comparison of Direct and Indirect Effects	How to present comparisons of results that are obtained from direct evidence or from indirect and/or mixed evidence
Uncertainty	How to present issues around the uncertainty (or certainty) of NMA results
Rankings	How to present rankings in an understandable way to users, such that they are not misinterpreted
Implications of Findings	How to present the context surrounding NMA results, i.e. their impact and the implications of the finding

Supplementary Table 3. Purpose of NMA Guideline Documents

Guideline	Purpose of Guidance
ISPOR (Pt1 and2)	1) Provides guidance on the interpretation of indirect treatment comparisons and NMA to assist policymakers and health-care professionals in using its findings for decision-making (2) Provides guidance on technical aspects of conducting network meta-analyses
CADTH 2009	To identify and review the different methods available for making indirect treatment comparisons; select methodological objectives using the Bucher indirect treatment comparison approach; to illustrate the application of various methods for indirect treatment comparisons
NICE DSU Series	To support those presenting submissions to National Institute for Health and Clinical Excellence (NICE) by addressing the issues of the relative and absolute efficacy of interventions as it relates to evidence synthesis methods in decision-making
Australia PBAC	To provide a framework to help assess each indirect comparison to help judge the degree of confidence that can be given to relying on this assumption
France HAS	To provide a literature review to (1) present the background to the development of indirect comparison methods; (2) identify situations in which indirect comparisons can make a real contribution (3) present the various reliable indirect comparison methods (4) describe the main advantages and limitations of each methods (5) examine the validity of indirect comparisons (6) propose an evaluation and critical review method for indirect comparison studies
USA -AHRQ	To summarize publicly available guidance for, and current use of, meta-analytic methods for multiple treatment comparison 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
EUnetHTA	To describe the main methods of direct, indirect and mixed treatment comparison available; to explain the main strengths and weaknesses of the methodologies; to discuss common issues that must be considered when interpreting results; to provide recommendations on the use of direct and indirect comparisons in a relative effectiveness assessment

Supplementary Table 4. **Quality of Network Meta-Analysis Guidelines as Assessed by modified AGREEII Instrument***

Guideline	Scope	Stakeholder Involvement	Rigour	Clarity	Applicability	Editorial Independence	Comments
ISPOR (Pt1 and2)			x		x	x	**Strengths:** clear objective and audience; reviewed by experts before publishing; key stakeholders (researchers, industry, decision-makers) invited to join task force and could provide comments on drafts; reference list and citations throughout text; strengths and limits of evidence noted; includes accompanying tools; recommendations specific and clear; author affiliations reported **Limits:** no systematic search for evidence; process for formulating recommendations unclear; pros and cons of recommendations not identified; no process for updating but areas for future research identified; implementation issues not discussed; no monitoring criteria; COI not reported; editorial independence from ISPOR unclear
CADTH 2009	x	x	x	x	x	x	**Strengths:** clear objectives; comprehensive literature search and selection of commonly reported methodologies; reference list and citations throughout text; strengths and limits of evidence noted; includes accompanying tools; COI reported; funded by decision-makers **Limits:** audience not identified; reviewed by external academics/researchers before publishing but no other stakeholders involved (e.g. industry, decision-makers); recommendations not clearly identifiable and unambiguous (but providing recommendations was not a key objective); pros and cons of recommendations not discussed; process for formulating recommendations unclear; no process for updating but areas for future research identified; implementation issues not discussed; no monitoring criteria
NICE DSU-TSD Series		x	x		x	x	**Strengths:** clear objective and audience; is regularly updated; reviewed by experts before publishing; reference list and citations throughout text; strengths and limits of evidence noted; recommendations specific and clear; includes accompanying tools; provides details on how to implement methods; author affiliations reported; editorial independence from funder (i.e. NICE) **Limits:** no systematic search for evidence reported; involvement of broad stakeholder community unclear (e.g. decision-makers, industry); process for formulating recommendations unclear; pros and cons of recommendations not identified; no monitoring criteria; COI not reported

Australia PBAC		x	x		x	x	**Strengths:** clear objectives and audience; systematic search for evidence; reference list and citations throughout text; strengths and limits of evidence noted; recommendations specific and clear; developed and funded by decision-makers **Limits:** stakeholder involvement or review of drafts outside of working group not reported (e.g. industry); criteria for selecting evidence unclear; process for formulating recommendations unclear; pros and cons of recommendations not identified; no process for updating or monitoring; no accompanying tools; implementation issues not discussed; COI not reported
France HAS	x	x	x		x	x	**Strengths:** clear objectives; systematic search for evidence conducted and selection criteria outlined; reference list and citations throughout text; strengths and limits of evidence noted; recommendations specific and clear; includes accompanying tools; developed and funded by decision-makers **Limits:** audience unclear; stakeholder involvement or review of drafts not reported (e.g. industry); process for formulating recommendations unclear; pros and cons of recommendations not identified; no process for updating or monitoring; implementation issues not discussed; COI not reported
USA -AHRQ		x	x		x		**Strengths:** clear objectives and audience; reference list and citations throughout text; strengths and limits of evidence noted; recommendations specific and clear; to be updated as evidence evolves; COI reported; editorial independence from funder (i.e. AHRQ) **Limits:** stakeholder involvement unclear; no systematic search for evidence reported; selection of evidence unclear; process for formulating recommendations unclear; pros and cons of recommendations not identified; no process for monitoring; no accompanying tools; implementation issues not discussed
EUnetHTA		x	x		x	x	**Strengths:** objective and audience clearly defined; reviewed and validated by experts in HTA agencies who are target audience but other stakeholders not engaged; systematic search for evidence conducted and selection criteria outlined; recommendations are clear; reference list provided and cited within text; funded by decision-makers **Limits:** process for formulating recommendations unclear; no process for updating or implementing recommendations; no accompanying tools; authorship and COI not reported

'x' – indicates an issue was identified during critical appraisal of the guideline

* Minor modifications were made to the AGREEII instrument to assess methodological guidelines rather than clinical practice guidelines, as described in the Methods section of the Supplementary Materials.

Section/topic	#	Checklist item	Reported on page #
TITLE			
Title	1	Identify the report as a systematic review, meta-analysis, or both.	1
ABSTRACT			
Structured summary	2	Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number.	2
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of what is already known.	3
Objectives	4	Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).	4
METHODS			
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number.	NA
Eligibility criteria	6	Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale.	5
Information sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.	5
Search	8	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.	5
Study selection	9	State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).	5

Data collection process	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.	6
Data items	11	List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made.	6
Risk of bias in individual studies	12	Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis.	6
Summary measures	13	State the principal summary measures (e.g., risk ratio, difference in means).	NA
Synthesis of results	14	Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g., I^2) for each meta-analysis.	NA

Section/topic	#	Checklist item	Reported on page #
Risk of bias across studies	15	Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies).	6
Additional analyses	16	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified.	NA
RESULTS			
Study selection	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram.	7
Study characteristics	18	For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations.	7-8
Risk of bias within studies	19	Present data on risk of bias of each study and, if available, any outcome level assessment (see item 12).	8
Results of individual studies	20	For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group (b) effect estimates and confidence intervals, ideally with a forest plot.	31-32

Synthesis of results	21	Present results of each meta-analysis done, including confidence intervals and measures of consistency.	NA
Risk of bias across studies	22	Present results of any assessment of risk of bias across studies (see Item 15).	7-8
Additional analysis	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see Item 16]).	NA
DISCUSSION			
Summary of evidence	24	Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., healthcare providers, users, and policy makers).	13-16
Limitations	25	Discuss limitations at study and outcome level (e.g., risk of bias), and at review-level (e.g., incomplete retrieval of identified research, reporting bias).	16-17
Conclusions	26	Provide a general interpretation of the results in the context of other evidence, and implications for future research.	17-18
FUNDING			
Funding	27	Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review.	19

From: Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. PLoS Med 6(6): e1000097. doi:10.1371/journal.pmed1000097

For more information, visit: **www.prisma-statement.org**.

CHAPTER 7

CANADIAN AND INTERNATIONAL PERSPECTIVES ON KNOWLEDGE TRANSLATION OF ECONOMIC EVALUATIONS AND NETWORK META-ANALYSES: KEY INFORMANT INTERVIEWS

CHAPTER OVERVIEW

Purpose and Relationship to Other Manuscripts:

One of the goals of this thesis is to understand barriers and facilitators to communicating economic evaluations and network meta-analyses (NMAs) from the perspective of end-users and to identify international best practices and tools for communicating these analyses that could be adapted to the Canadian context. Therefore, the primary objective of this fifth manuscript is to conduct key informant interviews with Canadian policymakers, i.e. knowledge users, to identify their informational needs related to understanding and using economic evaluations and NMAs to support policy decisions. A secondary objective is to conduct key informant interviews with HTA organizations around the world, i.e. knowledge producers, to identify a sample of knowledge tools related to economic evaluations and NMAs.

This work will be used to inform the informational needs and optimal presentation of economic evaluations and NMAs for policymakers. It will be able to confirm if findings in the literature related to knowledge, barriers, facilitators and approaches to communicating economic evaluations (Chapter 3) and NMAs (Chapter 5) are applicable to the Canadian context. It will also provide critical information on policymakers' needs with respect to NMAs as no evidence evaluating policymakers' preferences was identified in the systematic review on approaches to communicating NMAs (Chapter 5).

Supporting Material:

Appendix B. Ethics Approval

Appendix C. Interview Guides for Key Informant Interviews

Manuscript Status:

Submitted for review and awaiting a decision. All information in this chapter, including the supplementary materials, has been submitted to the journal.

Authors:

Shannon M Sullivan (SMS), George Wells (GW), Doug Coyle (DC), Elaine MacPhail (EM), Barbara Davies (BD)

Author Roles (SMS, GW, DC, EM, BD) and Contributions:

SMS, GW, DC were responsible for the study conception.

SMS, GW, DC, EM, BD were responsible for the design of the data collection tools

SMS, EM and DC were responsible for identifying and/or recruiting target populations

SMS was responsible for conducting interviews

SMS, BD were responsible for the qualitative data analysis

SMS, GW, DC, EM, BD were responsible for the interpretation of results.

SMS, BD were responsible for drafting the manuscript.

All authors (SMS, GW, DC, EM, BD) provided comments on drafts and revisions to the manuscript, agree with the manuscript's results and conclusions, and read and approved the final manuscript.

TITLE:
Canadian and International Perspectives on the Knowledge Translation of Economic Evaluations and Network Meta-Analyses for Policymakers. A Qualitative Analysis

SHORT TITLE:
Economic Evaluations and NMAs for Policymakers

ABSTRACT:

Objective: Health policy and drug funding decisions may be informed by economic evaluations and network meta-analyses (NMAs), which can be complex analyses. Policymakers are important end-users of economic evaluations and NMAs. Therefore it is important to understand their informational needs when communicating them. Enhancing the relevance, transparency and accessibility of economic evaluations and NMAs may increase confidence in the legitimacy of health policy decisions informed by these analyses. Applications of knowledge translation (KT) principles may assist.

Methods: Key informant interviews were conducted with (1) Canadian policymakers, i.e. knowledge users, to identify their informational needs related to understanding and using economic evaluations and NMAs to support policy decisions and (2) senior individuals in health technology assessment (HTA) organizations around the world, i.e. knowledge producers, to identify a sample of existing tools that could inform best practices and existing challenges in this area. Data were synthesized qualitatively

Results: For both economic evaluations and NMAs, Canadian policymakers (N=22) value information that includes (1) content explaining the evidence base, limitations in the analysis and uncertainty of results; (2) contextual information on local and real-world policy applications of the evidence and (3) formatting approaches to enhance clarity and key messages. Internationally (N=17 key informants), few KT approaches specific to economic evaluations and NMAs exist, despite an acknowledged need and interest in better approaches.

Conclusions: Approaches to enhance the presentation of economic evaluations and NMAs for policymakers were identified. These could inform development of tools tailored to policymakers that could enhance transparency and accessibility of economic evaluations and NMAs.

KEY WORDS:
economic evaluation; network meta-analysis; decision-making; knowledge translation; qualitative methods

ACKNOWLEDGEMENTS:
None

SOURCES OF FUNDING:
Canadian Institutes of Health Research

INTRODUCTION

Transparency is a key component of credible and legitimate decision-making processes (1). Full transparency includes, not only providing information, but also presenting information in an accessible and understandable format. Health policy and drug funding decisions may be informed, in part, by assessments of health technologies that include economic evaluations and, sometimes, network meta-analyses (NMAs), also known as indirect comparisons. Economic evaluations and NMAs can be complex analytical techniques and various audiences have expressed challenges interpreting or applying these analyses to actual decisions (2-5). Policymakers, i.e., knowledge users, are important audiences for economic evaluations and NMAs. Therefore it is important to understand their informational needs and how researchers, i.e. knowledge producers, can best communicate these analyses to them. Clearer communication could enhance the relevance, transparency and accessibility of economic evaluations and NMAs, and subsequently, increase confidence in the legitimacy of health policy decisions informed by these analyses.

Understanding how to communicate information to different audiences is a cornerstone of knowledge translation (KT). KT can be defined as "a dynamic and iterative process that includes synthesis, dissemination, exchange and ethically-sound application of knowledge to improve health, provide more effective health services and products and strengthen the health care system." (6). The Knowledge-to-Action (KTA) framework (see Supplementary Figure 1) is one framework that identifies key concepts and steps in knowledge creation (7-8). In the KTA framework, after a problem is identified, barriers and facilitators to knowledge use are assessed. This assessment provides an evidence-base to develop tools that are tailored to target audiences (e.g. policymakers) and that can be adapted to the local setting (e.g. different provinces).

In Canada, each province and territory is responsible for healthcare and health policy-related decisions. As it relates to public drug funding, Canadian policymakers are supported by a common national review process, which provides recommendations to Canadian provinces (excluding Quebec), territories and federal drug programmes (9). Although economic evaluations and, more recently, NMAs, are frequently used to inform Canadian health policy decisions, the informational needs of Canadian policymakers in this area have never been formally surveyed and published. Many other countries also have health technology assessment (HTA) processes that use economic evaluations, and sometimes NMAs (10-11). Although the national context of how HTAs are conducted and used varies across countries, HTA organizations worldwide can benefit from each others' collective knowledge and experiences as some issues related to using evidence to inform health policies are universal (12). Therefore, approaches or solutions that have been applied to communicating and sharing economic evaluations and NMAs in the

international context could inform best practices for supporting Canadian policymakers in these fields.

Therefore, the primary objective of this study is to conduct key informant interviews with Canadian policymakers, i.e. knowledge users, to identify their informational needs with respect to understanding and using economic evaluations and NMAs to support policy decisions. A secondary objective is to conduct key informant interviews with senior individuals in HTA organizations around the world, i.e. knowledge producers, to identify a sample of existing knowledge tools that could inform Canadian approaches to enhancing the transparency and accessibility of economic evaluations and NMAs for policymakers.

METHODS

(See Supplementary Materials, Box 1 for additional methodological detail)

Study Design

Semi-structured telephone interviews were conducted with purposive samples of (1) Canadian policymakers from provincial, territorial and federal publicly-funded drug programmes and (2) senior individuals in HTA organizations around the world. Semi-structured telephone interviews with key informants were conducted in English by one researcher between March-April 2014 (Canadian participants) and May-July 2014 (international participants). Interviews were approximately 30-45 minutes. Interviews were audio-recorded with the permission of participants, and transcribed by a professional transcriptionist. Ethical approval was obtained from the Ottawa Health Science Network Research Ethics Board.

Recruitment

Canadian participants were recruited from public drug programmes participating in national HTA review processes and who are involved in making drug-funding decisions, i.e. the Common Drug Review participants. Participants included policymakers, internal staff, external consultants or expert advisory committee members and represented diverse knowledge users. International participants were recruited from senior individuals in HTA organizations with membership in the International Network of Agencies for Health Technology Assessment (INAHTA) or Health Technology Assessment international (HTAi) and were intended to represent perspectives of knowledge producers. In both samples, key informants were individuals with knowledge of how economic evaluations and NMAs are used in their programme to support health policy decision-making.

Canadian contacts were obtained through membership on a national advisory committee whose members represent all of the Canadian public drug funding programmes. International contacts were obtained through HTAi and INAHTA contact lists and/or informal networks. In both samples, the organizational contacts were responsible for identifying key informants.

Data Collection

The KTA framework was prospectively applied to develop two unique interview guides, which included questions related to information about both economic evaluations and NMAs.

- For Canadian participants, questions focused on the informational needs of policymakers, how information is used, perceived knowledge, and perceived barriers and facilitators to use and understanding of this information.
- For international participants, questions focused on KT tools and approaches applied to this information

Interview guides were validated and piloted prior to widespread use.

Supporting characteristics of Canadian samples were obtained from the Canadian Institute of Health Information (CIHI) and Statistics Canada (13-14) and of international samples were obtained from INAHTA and HTAi websites or organizational websites (15-16).

Data Analysis:

Thematic content analysis was conducted on the data (17-19). Transcripts were coded using a descriptive content analysis approach and themes identified through an iterative process. Themes were verified by a second researcher experienced in qualitative research and any disagreements were resolved through further discussion and input from a third researcher.

RESULTS

KNOWLEDGE USERS: CANADIAN POLICYMAKER INTERVIEWS

Canadian Key Informants from Provincial, Territorial and Federal Drug Programmes

Key informants were identified by an advisory committee representing 18 publicly funded provincial, territorial and federal drug programmes in Canada. Some key informants represented several programmes and/or were a shared resource between smaller drug programmes. Telephone interviews were conducted with 22 of 24 key informants identified by the advisory

committee (response rate, 92%). See Supplemental Figure 1 for recruitment of Canadian policymakers.

Participants included senior management or directors (n=9/22, 41%), expert committee members who make recommendations to the drug programmes (n=6/22, 27%) and staff who support policymakers (n=7/22, 32%). Two of the expert committee members were on a national committee, and not on a specific provincial committee. Most participants had more than 10 years (n=14/22, 64%) or between 5-10 years experience (n=7/22, 32%) related to drug funding and reimbursement work. The majority of participants were health care professionals (n=14/22, 64% pharmacists; n=5/22, 23% physicians). Two participants (9%) had graduate-level training in health economics.

In 2011, the population across Canadian provinces and territories ranged from 34,200 people in Nunavut to 13,263,500 in Ontario and the percent of prescription drug expenditures that were publicly funded ranged from 29.7% in New Brunswick to 64% in the Northwest Territories (see Table 1) (16-17). When considering a combination of both of these metrics the largest provinces (i.e. large population and high expenditures) were Ontario, Alberta and British Columbia. Similar metrics were unavailable for the Canadian federal drug programmes. All programmes had access to national resources through the Canadian Agency for Drugs and Technologies in Health (CADTH) and participated in drug pricing negotiations to some extent, either nationally through the Pan-Canadian Pricing Alliance (PCPA) or locally (20-21). In addition to relying on recommendations from the national Canadian Drug Expert Committee (CDEC), most drug programmes also had a local or regional expert committee to assist in making drug-funding decisions. Some of these drug programmes also had in-house economic expertise, although the majority of the work they conducted was related to budget impact assessments, rather than cost-effectiveness analyses. Only one drug plan reported having access to in-house NMA expertise.

INSERT TABLE 1. Characteristics of Canadian Provinces, Territories and Public Drug Programmes That are Represented in Key Informant Interviews

General types of economic information used by policymakers included national funding recommendations, reports that are written based on a review of pharmaceutical manufacturer's submitted economic models and reports, cost tables comparing drug prices of relevant therapies, budget impact assessments and published, scientific articles on cost-effectiveness. NMA information that was relied upon included therapeutic reviews and recommendations provided by CADTH; NMAs provided by the pharmaceutical industry and critiqued by CADTH; and published

peer-reviewed literature and/or reports from other organizations. Similar types of information were used by international HTA organizations.

A number of overarching themes frequently emerged related to the use of economic evaluations and NMAs across the different questions that were asked, as listed in Table 2. A detailed description of results is provided below for the primary study question, "What are policymakers' informational needs" and in Supplementary Material Box 2 for the remaining interview questions explored with Canadian policymakers (i.e. how policymakers use information; their knowledge of the information; and barriers and facilitators).

INSERT TABLE 2. Overarching Themes from Canadian Key Informant Interviews, Organized by Question

What are Policymakers' Informational Needs?

Three themes contributed to economic evaluations or NMAs that meet policymakers informational needs, relating to (A) the quality of the content, (B) local and real-world contextual information and (C) the clarity of the how information is presented. These themes are described below, with specific suggestions on content elements that are valued by policymakers noted in Supplemental Table 1.

(A) CONTENT: Understanding potential limitations, the evidence-base, and uncertainty associated with results

Economic Evaluations: Policymakers indicated that understanding the content and quality of the economic information was important. Being aware of potential limitations in the economic model and the accuracy of the resulting incremental cost-effectiveness ratios was frequently raised. As one participant suggested sometimes "we don't know, was that a good number or not"" (I14). Participants often raised the uncertainty associated with economic estimates as an issue. It was noted that being presented with a variety of related estimates e.g. best and worst-case scenarios, a range of cost-effectiveness estimates, or various sensitivity analyses under different assumptions, was useful in understanding uncertainty. It was noted that policy makers may not have "a very good framework for how to approach uncertainty" (I19) but that this would be beneficial to develop. Finally, participants often raised the issue that cost-effectiveness estimates sometimes differed when independent researchers critically appraised pharmaceutical industry analyses. Being able to clearly understand the reasons for the differences between industry estimates and the independent estimates was important to policymakers, particularly when price

negotiations were required in advance of a drug funding decision: "we need to understand what the disagreements are just to help us…navigate through the negotiations" (I11).

NMAs: Similar to economic analyses, policymakers indicated that understanding the content and quality of NMAs was important when NMAs are used in a policy context. Policymakers noted that knowing what studies are included in the analysis, their relevance to real-world scenarios and any limitations of the analysis was very important. Policymakers frequently raised the issue that "the quality of the clinical data that's even going into the analysis to start with" (I16) is limited. Policymakers often had questions around the studies used in the NMA: "What were the studies that they used? What did they exclude? Why did they exclude them? (I20). It was frequently noted that authors "have to be clear of what the limitations were" (I2) and many suggested that completed quality checklists would be a valuable component of NMAs. As one participant asked, "What makes a good indirect comparison versus the one that's not so good?...what do we accept? What do we not accept?" (I13). Similar to economic analyses, policymakers raised the issue of understanding the uncertainty associated with NMA results, i.e. "where we struggle is really how certain should we be? How much weight should we put on this indirect comparison in our decision" (I11). As one policymaker, expressed, results alone aren't always enough when making a decision as they "often want numbers and what the numbers mean" (I2).

(B) CONTEXT: Identifying relevant local and real-world factors when presenting information

Economic Evaluations: In addition to considering different end-users when presenting economic information, participants noted that if the information is "tailored to each situation" (I2) it would be most useful. Identifying contextual factors related to the specific drug or technology in question and the subsequent local environment upon implementation of a policy decision was important to participants. Participants also noted the economic information is often interpreted in the context of social or ethical factors and having this context identified in the provided information is useful. For example, knowing if "there are other concerns that aren't captured in the economic analysis for which maybe we'd pay a higher cost per QALY in some circumstances than others" can help them understand how the economic results could apply to a policy decision (I19).

NMAs: Understanding the context in which an NMA is conducted is also important to policymakers. They noted that understanding "how it relates to the real world, either in terms of dosing or populations - that's important" (I16). Similarly, being clear on the included therapies in an NMA is important because "if you don't have the right drugs in there to start with, it's still

limited" (I9). From a policymaker's perspective it is important to "try to make it [relate] back to the impact it would have on our population. So we always look to see what is currently available and what does this mean to change it" (I2).

(C) CLARITY: Tailored presentation formats providing clarity on conclusions and rationale with appropriate layering of information

Economic Evaluations: Three preferences emerged related to the format of how both economic information and NMAs are provided to policymakers. Policymakers value formats that layer information, are tailored to their audience and that provide clarity. It was noted that because policymakers must be able to explain decisions to others that are based upon economic evaluations, the clarity of the information provided is essential. Clarity on the rationale for any conclusions or recommendations was important, as was being able to compare rationales across previous decisions to ensure consistent and legitimate decision-making. As one participant discussed "it has to be clear what the justification for that is because you ultimately could get questions, well why did you do this, when last time you did [that]" (I5). Key informants also noted that different types of policymakers have differing levels of familiarity with or expertise in economics and want varying amounts of detail from an economic analysis. Respondents discussed ways in which this could be achieved, including layering information with the most important details provided upfront, noting that "providing the overall report but also synthesizing it down to some key points and some key bottom-line recommendations is helpful for us" (I17).

NMAs: Policymakers expressed similar ideas around the format and presentation of NMAs with elements related to layering, tailoring and clarity being important. Different participants preferred varying levels of NMA detail. For example, one respondent noted "often I don't need to know all this...but some want to go more in-depth" (I2). Clarity of the analysis was also emphasized. As one respondent suggested, providing "just transparency, throughout the review, so you know exactly what you're looking at" would be beneficial. Another noted that "being more clear on the rationale and the justification for certain decisions that are made during the review" (I14) would be helpful. Finally, having information targeted to policymakers was raised by one participant who indicated "you just want to hone in on what's relevant" (I15).

How Do Policymakers Use Economic Evaluations and NMAs?

Three themes emerged related to how policymakers use economic evaluations: using them as part of an evidence-based decision process; adapting them to the local or real-world context of

policy decisions and using them to inform drug price negotiations. Two themes emerged related to how policymakers use NMAs: adapting them to the real-world and local context of policy decisions and using them when it is the best evidence available to inform a decision. See Supplemental Box 2 for more details.

What are Policymakers' Perceptions of Knowledge About Economic Evaluations and NMAs?

Three unique themes emerged related to policymakers' knowledge of economic evaluations and NMAs. In both fields, there was variable expertise but interest in training initiatives. For economic evaluations, there was a broad understanding of complex issues and for NMAs there was acknowledgement that they are complex methods but there was an evolving comfort with these methods. See Supplemental Box 2 for more details on themes.

What are Policymakers' Barriers and Facilitators to Using and Understanding Economic Evaluations and NMAs?

Six key barriers to using economic evaluations and NMAs emerged. In both types of analyses, the limited quality of analyses, the limited relevance to the real-world or local context and the limited availability of expertise were important barriers. For economic evaluations, drug pricing challenges, such as the existence of confidential price negotiations, and budget constraints were also important barriers. For NMAs, the lack of clarity in analyses was a key barrier.

Three key facilitators to using economic evaluations and NMAs emerged. In both types of analyses, having good quality studies and methodological approaches was a facilitator, as was having access to experts or training. For NMAs, having analyses that were applicable to the local or real-world context was also a facilitator. See Supplemental Box 2 for more details on themes. Specific suggestions of how these barriers and facilitators could be addressed in the information that is provided to policymakers are noted in Supplemental Table 2.

Overall pattern analyses by provincial size and policymaker role are noted in Supplemental Box 2.

KNOWLEDGE PRODUCERS: INTERVIEWS WITH INTERNATIONAL HTA ORGANIZATIONS

International Key Informants from HTA Organizations

Seventeen key informants from 14 HTA organizations were identified; telephone interviews (n=13) or written responses (n=3) were provided by 16 of them (key informant response rate, 94%). Six HTA organizations did not respond to the call for key informants (organizational

response rate, 70%, n=14/20), however, regional targets that were specified *a priori* for the environmental scan sample (i.e. two organizations from each geographic region) were closely achieved (90%, n=9/10). The exception was Asia, where only one of two eligible organizations participated. See Supplemental Figure 2 for recruitment of HTA organizations.

The 20 eligible HTA organizations represented five geographic regions. Participation was achieved from 75% (n=3/4) American, 80% (n=4/5) Northern Europe, 57% (n=4/7) Southern Europe, 50%(n=1/2) Asian and 100% Oceanic (n=2/2) organizations. Most organizations (93%, n=13/14) conducted HTAs to inform drug-funding policies, with some organizations also conducting HTAs of other non-drug technologies. Mandates varied across organizations, however, most were intended to support policies and policymakers in maximizing health and optimizing resource allocation by conducting evidence syntheses (see Table 3).

INSERT TABLE 3. Characteristics and Knowledge Translation Approaches of HTA Organizations Around the World (i.e. Knowledge Producers) That Are Represented in Key Informant Interviews

What are International Approaches to the Knowledge Translation of Economic Evaluations and NMAs?

Three overarching themes emerged related to the communication and KT of economic evaluations and NMAs in HTA organizations around the world: general KT approaches existed for HTAs; KT approaches specific to economic evaluations or NMAs were limited; interest in enhancing communication of economic evaluations and NMAs was consistently reported. These themes are described below and listed in Table 4. A more detailed list of KT approaches is provided in Supplemental Table 3. Fewer suggestions were obtained on NMAs due to limited use or familiarity with the methods.

INSERT TABLE 4. Approaches to the Knowledge Translation of Economic Evaluations and NMAs Identified by International Key Informants

General Knowledge Translation Approaches Exist for HTAs

The ability of an HTA organization to extensively support KT approaches (e.g. tailoring to different audiences, adapting evidence to different local settings, dissemination and stakeholder engagement activities) generally depended on the availability of resources to the organization. While all recognized the importance of such activities, it was not always a realistic goal given available resources or timelines. For example, one respondent indicated that “we do [the work] in a very short time” (I2), which doesn’t permit such activities. Many organizations employed

general communications staff and dissemination techniques or tools that could be applied to general evidence syntheses as well as economic evaluations or NMAs. These included publishing reports on the organization's website and providing press releases. However, most organizations did not devote extensive resources to developing project-specific tools or products. Those organizations that did invest in KT strategies noted approaches such as engaging with stakeholders, key points targeting different stakeholder audiences, policy briefs, patient leaflets, executive summaries as well as "newer tools like infographics and some things that are specific to social media" (I1)

Few Knowledge Translation Approaches Specific to Economic Evaluations or NMAs

Even among organizations that invested in KT approaches, specific strategies related to economic evaluations and NMAs, and not other types of evidence syntheses, were not usually developed. A number of organizations noted the use of guidelines, critical appraisal checklists, report templates and training workshops. One organization described a graphical tool demonstrating how the results of an economic evaluation could be applied to different local settings. One organization indicated that researchers were provided specific language and template wording to try to explain economic concepts to lay users. Another organization described lay reviewers, including a professional journalist and non-research staff who indicate when "this expression is incomprehensible to me." (I9). One organization indicated that future plans may include developing "more lay friendly guides that will explain a little bit more about the economics and trying to take it down away from such technical language." (I8) No organization identified specific approaches for enhancing the communication of NMAs and, as one respondent noted "it's not your normal ABC's" (I9). Many organizations are focusing first on ensuring their researchers understand how to conduct and critique NMAs before exploring how to better communicate these techniques to lay audiences.

Interest in Enhancing Communication of Economic Evaluations and NMAs

Despite the lack of specific tools or approaches for communicating economic evaluations and NMAs, HTA organizations consistently indicated an interest in and a need for better approaches. As one respondent noted, "we are now communicating to the patient why the product was rejected when reasonably physicians would want to use it…I'm also now explaining to the government why I would expect them to sink another billion dollars into the pharmaceutics area and they are asking me so what are we going to get? And I can't give them an incremental cost effectiveness ratio as an answer" (I6). Although NMA techniques are more novel and there is less experience with them, respondents echoed one individual's comment that "there is a growing need" in this area (I4)

Overall pattern analyses by geographic region and years of experience are noted in Supplemental Box 2.

DISCUSSION

Summary of Findings

Canadian drug policymakers indicated they value economic evaluations and NMAs that (1) have the evidence-base, the limitations and the uncertainty clearly explained (2) identify and relate the analyses to real-world and local contextual factors and (3) are presented with adequate clarity and emphasize key messages. Perceptions were that policymakers had variable expertise in both economic evaluations and NMAs but that there was interest in further training. With respect to economic evaluations, it was noted that there was generally a broad understanding of complex economic modeling issues and this was sufficient for policymaking purposes when accompanied by access to expert resources. For NMAs, a relatively novel analytic technique, policymakers acknowledged its complexity but noted that there was evolving comprehension and comfort with NMAs and an openness to accept NMAs as the best available evidence in appropriate circumstances. Common barriers and facilitators to understanding and using economic evaluations and NMAs included similar themes such as available expertise, quality of evidence, relevance and clarity of information. Additional themes unique to economic evaluations were related to the existence of confidential drug pricing negotiations and budget constraints. Interviews with HTA organizations world-wide found that few organizations have developed KT approaches specific to economic evaluations or NMAs but that there is an acknowledged need and strong interest in enhancing approaches to communicating this information to non-technical audiences.

Comparison with Literature

Preferences for how health economic information is communicated, as reported in literature, include disaggregating results, presenting key principles and results; identifying practical applications of results, identifying implications for real-world decision-making and tailoring presentation formats to different end-users (2-4). The current study conducted in the Canadian context confirmed many of these findings. Furthermore, the literature identified that common barriers to understanding economic evaluations include a lack of clarity; not communicating principles, and unclear model assumptions (2-4), which were also identified in the current study. Studies surveying policymakers' knowledge of economics were also consistent with the current study in that they found that economic training is variable but there is interest to learn more. NMA literature indicates people are still challenged in understanding NMAs (5, 22). There is likely

more clarity around economic evaluations than NMA because it is a more established analysis approach.

When comparing the findings of this study with general approaches to communicating evidence to policymakers, consistency in themes is observed. Systematic reviews (23-24) of factors enhancing policymakers' uptake of evidence have identified a single clear message (25), a summary with clear recommendations (26), details on applicability of research to the local context and providing information that can be quickly scanned (27). These were also supported by a recent study on policymakers pharmaceutical information needs (28). The current study confirmed these findings, however, practical consideration of how these principles apply to economic evaluations and NMAs compared with traditional evidence policymakers receive such as systematic reviews and meta-analyses, is required. There is evolving research on presenting and reporting NMAs (29-30) as well as research on how to assist policymakers in using and interpreting NMAs and economic evaluations (31-32). Connecting this research with the real-world needs of policymakers identified in the current study will help guide the optimal presentation of NMAs for policymakers and other non-technical audiences.

This current study was prospectively structured around the KTA framework (7-8). Results of this study confirmed that the transparency and accessibility of economic evaluations and NMAs is valued by policymakers. Barriers and facilitators were identified that can inform development of tools and approaches to tailoring economic evaluations and NMAs to policymakers and adapting the analyses to local contextual factors. The key themes policymakers raised also align closely with another framework that guides the implementation of evidence in practice, Promoting Action on Research Implementation in Health Services (PARIHS) (33). In the PARIHS framework, successful implementation of an intervention such as a knowledge tool is a function of the nature of the evidence, the quality of the context of implementation and appropriate approaches to facilitation. In the current study, the "nature of the evidence" could be considered the ability to provide robust information on economic evaluations and NMAs that identifies the evidence-base, potential limitations and uncertainty in results. The "context of the implementation" is being able to provide information on economic evaluations and NMAs in a way that they are adaptable to local or real-world contextual factors that are important to policymakers. Consistent presentation formats may enhance the clarity of economic evaluations and NMAs, and thus, could be one "approach to facilitation". Therefore, creating effective knowledge tools for enhancing the transparency and accessibility of economic evaluations and NMAs can be considered a function of three interacting factors:

Effective Tools = Explanations of Potential Limitations and Uncertainty Associated with Results (EVIDENCE/CONTENT) + Tailoring to Local and Real World Factors (CONTEXT) + Presentation Formats that Enhance Clarity (FACILITATION/CLARITY)

Strengths and Limitations

This is the first work to systematically survey and publish findings related to Canadian drug policy-makers informational needs on economic evaluations and NMAs. While a number of international studies have been conducted in the area of economic evaluations, the nascent field of NMAs has not been studied in the same and this is highly original research. The current work is further strengthened by the attempt to identify international approaches to KT in the area of economic evaluations and NMAs that one could build upon when developing tools. Finally, recruitment and perspectives from HTA organizations around the world were not as comprehensive or as extensive as the Canadian component of this study, however, they provide a window on current KT practices, as was the original intent of this study element. This also provided an opportunity to confirm findings across different settings and suggest some practices that may be transferable across some world-wide HTA organizations.

Policy Implications

Policymakers use economic evaluations and NMA to inform drug funding and resource allocation decisions. Therefore, to ensure legitimacy of these decisions, analyses should be clearly and transparently communicated. The lack of knowledge tools and approaches specific to economic evaluations and NMAs, despite keen interest in the area, suggests a universal need for this kind of tool. Although researchers often focus on results and technical methodological approaches (34-35), policymakers have emphasized understanding the thought process behind the analysis, wanting recommendations and conclusions from individuals who are experts in the field, needing clarity on the rationale for different decisions made in the analysis and understanding how much weight to place on limitations and being able to understand how to adapt results to a real-world decision and local contextual factors. Therefore, the current study offers insights on three areas of focus when developing tools for policymakers related to economic evaluations and NMAs.

- **Content**: Clearly understanding the evidence-base and limitations of economic evaluations and NMAs assists policymakers in understanding uncertainty associated with the results and how much weight to place upon them.
- **Context**: The Canadian context for drug funding policy is complex with the local context varying in each province (21, 36). Therefore, clearly identifying real-world considerations or where there may be local variability that could change the cost-effectiveness of therapies

assists policymakers and could inform local priorities, price negotiations and implementation of funding decisions.

- **Clarity**: Highlighting issues important to policymakers (e.g. rationale for analysis decisions, contextual factors, etc.) and providing information in easily scannable documents that layers information is beneficial, e.g. providing key messages upfront and detailed information in subsequent layers (37).

Furthermore, incorporating these concepts into educational initiatives would contribute to building capacity in economic evaluations and NMAs.

Conclusions

Canadian policymakers value information on economic evaluations and NMAs in which its quality is clearly understood, is presented in a format providing clarity on rationales and conclusions, and which is relevant and adaptable to the local context for implementation of a policy decision. Although similar needs exist internationally for enhanced presentation of economic evaluations and NMAs, little work has been done on developing specific KT approaches or tools. Suggestions provided for enhancing the presentation of economic evaluations and NMAs to policymakers can inform future development of tools to support policymakers. These tools will contribute to enhancing transparency and increasing confidence in the legitimacy of health policy decisions that are informed by these types of analyses.

REFERENCES

1. Daniels N & Sabin J. Limits to health care: Fair procedures, democratic deliberations and the legitimacy problem for insurers. Philosophy & Public Affairs 1997; 303-350.

2. Galani C and Rutten FFH (2008). Self-reported health care decision makers' attitudes towards economic evaluations of medical technologies Cur Med Res Op 24: 3049-3058

3. Thurston SJ, Craig D, Wilson P and Drummond MF (2008). Increasing decision makers' access to economic evaluations: alternative methods of communicating the information. Intl J Technol Assess Health Care 24: 151-157.

4. Williams I, McIver S, Moore D, Bryan S. The use of economic evaluations in NHS decision- making: a review and empirical investigation. Health Technol Assess 2008; 12: 1-175.

5. Abdelhamid A, Loke Y, Parekh-Bhurke S, Chen Y, Sutton A, Eastwood A, Hollanda R, Song F. Use of indirect comparison methods in systematic reviews: a survey of Cochrane review authors. Research Synthesis Methods 2012, 3: 71-79.

6. Canadian Institutes of Health Research (2012). Guide to Knowledge Translation Planning at CIHR: Integrated and End-of-Grant Approaches. Ottawa: CIHR.

7. Straus S, Tetroe J & Graham ID (Eds.), Knowledge translation in health care: Moving from evidence to practice, 2nd edition. Oxford: Blackwell Publishing, 2013.

8. Graham ID, Logan J, Harrison MB, Straus SE, Tetroe J, Casell W et al. Lost in knowledge translation: time for a map: J Contin Ed Health Prof 2006; 26(1):13-24.

9. Tierney M, Manns B & Members of the Canadian Expert Drug Advisory Committee. Optimizing the use of prescription drugs in Canada through the Common Drug Review. CMAJ 2008, 178: 432-5.

10. Mathes T, Jacobs E, Morfeld JC, Pieper D. Methods of international health technology assessment agencies for economic evaluations--a comparative analysis. BMC Health Serv Res. 2013 Sep 30;13:371.

11. Oortwijn W1, Broos P, Vondeling H, Banta D, Todorova L Mapping of health technology assessment in selected countries. Int J Technol Assess Health Care. 2013 Oct;29(4):424-34.

12. Lavis JN, Wilson MG, Grimshaw JM, Haynes RB, Ouimet M, Raina P et al. Supporting the use of health technology assessments in policy making about health systems. International Journal of Technology Assessment in Health Care, 26:4 (2010), 405–414

13. Canadian Institute for Health Information (CIHI). Drug Expenditure in Canada, 1985 to 2011. Available at: https://secure.cihi.ca/free_products/DEIC_1985_2011_EN.pdf [Last accessed: August 9, 2014].

14. Statistics Canada,. Population by year, by province and territory (Number), 2009 to 2013. Available at: http://www.statcan.gc.ca/tables-tableaux/sum-som/l01/cst01/demo02a-eng.htm [Last accessed: August 5, 2014].

15. International Network of Agencies for Health Technology Assessment (INAHTA). Available at: http://www.inahta.org [Last accessed: August 7, 2014]

16. Health Technology Assessment international (HTAi). Available at: http://www.htai.org [Last accessed: August 7, 2014]

17. Miles M & Huberman A. *Qualitative data analysis. An expanded sourcebook*, 3rd edition. Thousand Oaks, CA. Sage Publications Inc., 2013

18. Hsieh H & Shannon S. Three approaches to qualitative content analysis. Qualitative Health Research, 2005, 15: 1277-1288.

19. Sandelowski M (2000). Focus on research methods. Whatever happened to qualitative description? *Research in Nursing & Health*, 23: 334-340.

20. Pan-Canadian Pricing Alliance (PCPA). Available at: http://www.conseildelafederation.ca/en/initiatives/358-pan-canadian-pricing-alliance [Last accessed: August 9, 2014].

21. Morgan S, Thomson P, Daw J, Friesen M. Inter-jurisdictional cooperation on pharmaceutical product listing agreements: views from Canadian provinces. BMC Health Services Research 2013; 13:34

22. Salanti G. Indirect and mixed-treatment comparison, network, or multiple-treatments meta-analysis. Many names, many benefits, many concerns for the next generation evidence synthesis tool. Research Synthesis Methods 2012; 3:80-97.

23. Chambers et al. Maximizing the impact of systematic reviews in health care decision making. A systematic scoping review of knowledge translation resources. Milbank Q, 2011, 89: 131-156.

24. Perrier L, Mrklas K, Lavis J, Straus S. Interventions encouraging the use of systematic reviews by health policymakers and managers: A systematic review. Implementation Science 2011; 6:43.

25. Murthy L, Shepperd S, Clarke MJ, Garner SE, Lavis JN, Perrier L, Roberts NW, Straus SE. Interventions to improve the use of systematic reviews in decision-making by health system managers, policy makers and clinicians. Cochrane Database of Systematic Reviews 2012, Issue 9. Art. No.: CD009401. DOI: 10.1002/14651858.CD009401.pub2

26. Innvaer, S., Vist, G., Trommald, M., & Oxman, A. (2002). Health policy-makers' perceptions of their use of evidence: A systematic review. Journal of Health Services Research and Policy, 7, 239-244.

27. Lavis, J., Davies, H., Oxman, A., Denis, J. L., Golden-Biddle, K., & Ferlie, E. Towards systematic reviews that inform health care management and policy-making. Journal of Health Services Research and Policy, 2005, 10 (Suppl 1), 35-48.

28. Greyson DL, Cunningham C & Morgan S. Information behaviour of Canadian pharmaceutical policy makers. Health Information and Libraries Journal, 2011, 29: 16–27

29. Hutton B, Salanti G, Chaimani A, Caldwell DM, Schmid C, Thorlund K et al. The quality of reporting methods and results in network meta-analyses. An overview of reviews and suggestions for improvement. PLoS One 2014, 9(3): e92508.

30. Tan SH, Cooper NJ, Bujkiewicz S, Welton NJ, Caldwell DM, Sutton AJ. Novel presentational approaches were developed for reporting network meta-analysis. J Clin Epidemol 2014, 67(6): 672-680.

31. Jansen JP, Trikalinos T, Cappelleri JC, Daw J, Andes S, Eldessouki R et al. Indirect treatment comparison/network meta-analysis study questionnaire to assess relevance and credibility to inform health care decision making: an ISPOR-AMCP-NPC Good Practice Task Force report. Value in Health 2014, 17(2): 157-173.

32. Caro J, Eddy DM, Kan H et al. Questionnaire to assess relevance and credibility of modeling studies for informing health care decision making: an ISPOR-AMCP-NCP Good Practice Task Force report. Value in Health 2014, 17: 174-182.

33. Rycroft-Malone J &Bucknall T. Models and Frameworks for Implementing Evidence-Based Practice: Linking Evidence to Action. Wiley-Blackwell: West Sussex, UK. 2010.

34. CADTH. Guidelines for the economic evaluation of health technologies: Canada [3rd Edition]. Ottawa: Canadian Agency for Drugs and Technologies in Health; 2006.

35. CADTH Health Technology Assessment Report. Indirect Evidence. Indirect treatment comparisons in meta-analysis. Ottawa: Canadian Agency for Drugs and Technologies in Health; 2009.

36. Daw JR & Morgan S. Stitching the gaps in the Canadian public drug coverage patchwork? A review of provincial pharmacare policy changes from 2000 to 2010. Health Policy 2012, 104: 19–26.

37. Canadian Health Services Research Foundation (CHSRF). Reader Friendly Writing, 1:3:25. Available at: http://www.cfhi-fcass.ca/SearchResultsNews/10-06-01/d497a465-5398-4ec8-addf-d7cbf86b1e43.aspx. [Last accessed: August 9, 2014].

38. Hutton B, Salanti G, Chaimani A, Caldwell DM, Schmid C, Thorlund K et al. The quality of reporting methods and results in network meta-analyses. An overview of reviews and suggestions for improvement. PLoS One 2014, 9(3): e92508.

39. Moher D, Weeks L, Ocampo M, Seely D, Sampson M, Altman DG et al. Describing reporting guidelines for health research: a systematic review. J Clin Epidemiol 2011, 64: 718-742.

TABLES AND FIGURES

Table 1. Characteristics of Canadian Provinces, Territories and Public Drug Programmes That are Represented in Key Informant Interviews

Province and/or Drug Plan*		Characteristics						
		No. Key informants (N=22)**	Local Expert Committee	In-House Economic Expertise	Access to National Resources	Sometimes Participate in Price Negotiations	Population, 2011 (in thousands)	% Prescription Drug Expenditure that is Publicly-Funded 2011
BC		3	Yes	Partial	Yes	Yes	4,499.1	41.2
AB		3	Yes	Partial	Yes	Yes	3,790.2	48.3
SK		2	Yes	Partial	Yes	Yes	1,066.3	49.0
MB		2	Yes	Partial	Yes	Yes	1,233.7	44.9
ON		2	Yes	Partial	Yes	Yes	13,263.5	45.0
Atlantic Provinces	NB	4	Yes (Atlantic)	Partial	Yes	Yes	755.5	29.7
	NS						944.5	37.9
	PEI						144.0	34.2
	NL						525.0	34.7
Territories	YK	1	No	No	Yes	Yes	35.4	61.7
	NWT						43.5	64.2
	NU						34.2	58.6
Federal Drug Plans		3	Yes	Partial	Yes	Yes***	N/A****	N/A****

Note: QC was not included in the original sampling frame as it is not a full participant in the pan-Canadian processes such as the Common Drug Review

Note: economic expertise was considered partial if it constituted expertise in budget impact assessments and not cost-effectiveness analyses

* For smaller drug plans (e.g. Atlantic, Territories, Federal), key informants were either not available for all plans or were shared across plans

** An additional 2 key informants were members of a national expert committee and not affiliated with any specific drug plan or province

*** Price negotiations occur only at local level or pan-federally, not through the pan-Canadian Pricing Alliance (PCPA)

**** Provincial populations and drug expenditures are reported based on the geographic region and not the size of the population covered by public drug funding plans, therefore, there is no corollary population size or drug expenditure data available from CIHI or Statistics Canada for federal drug plans. However, the total drug expenditure per capita was $664.6 million in 2011, representing 2.5% of total expenditure on prescribed drugs. The federal government provides drug benefits to Aboriginal Canadians, veterans, members of the military and the RCMP, prisoners in federal correctional facilities and refugees.

Table 2. Overarching Themes from Canadian Key Informant Interviews, Organized by Question

Economic Themes	n (%) of policy makers mentioning theme (N=22)	NMA Themes	n (%) of policy makers mentioning theme (N=22)
What are Policymakers Informational Needs?			
Applicable to Local or Real-World Context	21 (95)	Knowing Potential Content Limitations	18 (82)
Knowing Potential Content Limitations	19 (86)	Applicable to Local or Real-World Context	15 (68)
Clearly Formatted Information	19 (86)	Clearly Formatted Information	13 (59)
How Do Policymakers Use Economic Evaluations and NMAs?			
Adapted to Local Context	21 (95)	Adapted to Real-World and Local Contexts	15 (68)
Informs Price Negotiations	18 (82)	Used When is the Best Available Evidence	15 (68)
Part of Evidence-Based Decision Process	8 (36)		
What are Policymakers' Perceptions of Knowledge About Economic Evaluations and NMAs?			
Variable Expertise and Interest in Training	20 (91)	Variable Expertise and Interest in Training	16 (73)
Broad Understanding of Complex Issues	19 (86)	Complex Methods but Evolving Comfort	8 (36)
What are Facilitators to Policymakers' Understanding and Using Economic Evaluations and NMAs?			
Having Economic Expertise or Training	20 (91)	Good Quality Studies and Methods	18 (82)
Good Quality Evidence and Models	19 (82)	NMA Expertise or Training Available	16 (73)
		Applicable to Local or Real-World Context	15 (68)
What are Barriers to Policymakers' Understanding and Using Economic Evaluations and NMAs?			
No or Unclear Relevance to Local Context	21 (95)	Limitations in Quality of NMA	18 (82)
Limited Economic Expertise	20 (91)	Limited Expertise in NMA	16 (73)
Limitations in Quality of Evidence and Models	19 (86)	Limited Real-World Relevance	15 (68)
Budget Constraints and Challenges	19 (86)	Lack of Clarity	13 (59)
Pricing Challenges	17 (77)		

Table 3. Characteristics and Knowledge Translation Approaches of HTA Organizations Around the World (i.e. Knowledge Producers) That Are Represented in Key Informant Interviews

Country or Region	Organization	Purpose *	Knowledge Translation Approaches		
			General	Economics	NMA
Australia	PBAC**	Recommends new medicines for listing on the Pharmaceutical Benefits Scheme	Yes	No	No
Belgium	KCE	Advises policymakers about possibilities to obtain an efficient allocation of scarce health care resources that optimizes the quality and accessibility of health care.	Yes	Some	No
Canada	CADTH	Supports informed decisions about the uptake and use of drugs and other health technologies in Canada.	Yes	Some	No
Finland	FinOHTA	Promotes the effectiveness of Finnish health care, concentrating on technologies that deal with major health problems or are of economic importance for the country	Yes	Some	No
France	HAS	Improves the quality of care delivered to patients and assesses the expected and actual clinical benefit of drugs, medical devices, and diagnostic and therapeutic procedures and advises the authorities on their reimbursement.	Yes	Some	Some
New Zealand	PHARMAC**	Ensures citizens get the best possible health outcomes from money the government spends on medicines.	Yes	Some	No
Poland	AHTAPol	Informs Minister of Health on public funding decisions on public funding of health technologies, especially controversial, expensive procedures and provides guidance to policymakers and hospitals on offering the highest quality of evidence-based health care.	No	No	No
Quebec	INESSS	Promotes clinical excellence and the efficient use of resources in the field of health and social services	Yes	No	No

Country or Region	Organization	Purpose *	Knowledge Translation Approaches		
			General	Economics	NMA
Scotland	HIS-SMC	Supports healthcare providers in Scotland to deliver high quality, evidence-based, safe, effective and person-centered care and scrutinizes those services to provide public assurance about the quality and safety of that care	Yes	Some	Some
Spain	AETS	Assesses different health technologies as a basis for formulating policies on technology selection and implementation in the National Health System and promotes the appropriate use of existing technologies	No	No	No
Sweden	SBU	Identifies effective and ineffective health practices in different patient groups and thereby improves health services.	Yes	Some	No
Taiwan	CDE	Enhances the efficiency and quality of drug evaluation, thus promoting public health and welfare including timely access to innovative medicines.	No	No	No
UK	NICE**	To improve outcomes for people using the National Health Services and other public health and social care services.	Yes	No	Some
USA	AHRQ	Conducts and supports studies of the outcomes and effectiveness of diagnostic, therapeutic and preventive health care services and procedures	Yes	No	No

Abbreviations: AETS=Agencia de Evaluación de Tecnologias Sanitarias; AHRQ=Agency for Healthcare Research and Quality; AHTAPol=Agency for Health Technology Assessment in Poland; CADTH=Canadian Agency for Drugs and Technologies in Health; CDE=Center for Drug Evaluation; FinOHTA=Finnish Office for Health Technology Assessment; HAS=Haute Autorité de Santé; HIS-SMC=Healthcare Improvement Scotland – Scottish Medicines Consortium; INESSS=Institut national d'excellence en santé et en services; KCE=Belgian Healthcare Knowledge Centre; NICE=National Institute for Healthcare and Excellence; PBAC=Pharmaceutical Benefits Advisory Committee; PHARMAC=Pharmaceutical Management Agency of New Zealand; SBU=Swedish Council on Technology Assessment in Health Care

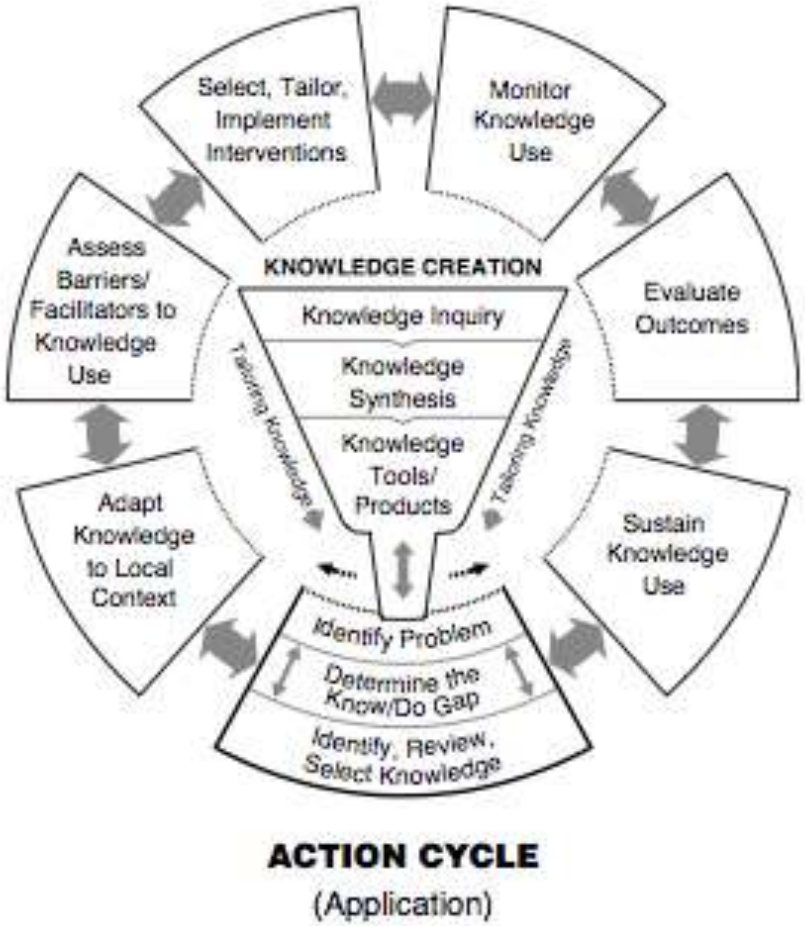
* Information was obtained from the INAHTA website descriptions of organizations as of Aug 2014.

** For organizations that are not INAHTA members, descriptions were obtained from the HTAi website and/or the HTA organization's website

Table 4. Approaches to the Knowledge Translation of Economic Evaluations and NMAs Identified by International Key Informants

General KT Approaches That Could be Applied to Economic Evaluations and NMAs
• Developing different types of tools including decision aids, practice guidelines and patient leaflets • Targeted approaches for policymakers and clinicians including key messages and policy briefs • Summary documents of reports and deliberations written by non-researchers • Approaches to writing reports and type of information to include • Use of editorial and review processes • Stakeholder interactions to enhance uptake and understanding • Applying integrated KT strategies during HTA conduct and implementation
Economic-Specific KT Approaches*
• Brief summaries of economic evaluation reports • Approaches to writing economic evaluations including templates, language and points to emphasize • Tools that assist with planning and implementation of policy decisions • Stakeholder interactions to enhance understanding • Providing resources including guidelines, training workshops and reports explaining key economic issues.
NMA-Specific KT Approaches*
• Brief summaries of NMAs • Approaches to writing NMAs including what points to explain • Stakeholder interactions to enhance NMA uptake • Providing resources including guidelines, training and NMA literature

SUPPLEMENTARY MATERIALS.



Supplemental Figure 1. Knowledge to Action (KTA) Framework.
From Straus et al., 2013 (used with permission from John Wiley and Sons)

Supplemental Box 1. Additional Methodological Details and Considerations.

ADDITIONAL METHODOLOGICAL DETAILS

Selection and Recruitment Details
Among Canadian participants, a target of 2-3 individuals per programme was established a priori, with the expectation that due to the small size and organizational structures of some drug programmes that represent smaller populations, fewer key informants may be identified or may be shared amongst programmes.

Selection criteria applied to international HTA organizations and considerations around sample size are outlined here. INAHTA members are non-profit organizations that are at least 50% funded from public sources; that assess technology in health care; and that relate to a regional or national government. HTAi members are individuals or organizations that come together to create an international network of experts and leaders in HTA. Only organizations from countries in the Organisation for Economic Co-operation and Development (OECD) and with a publicly funded healthcare system were included to ensure comparability with the Canadian context. Finally, only organizations that conduct or critically appraise HTAs on drugs and provide advice on or make decisions related to drug funding policies were included. Expert opinion was used to refine the list and include or exclude additional organizations. Key informants were senior individuals with knowledge of how economic evaluations and NMAs are shared and communicated with knowledge users. A target of at least two organizations in each of five geographic regions (Americas, Northern Europe, Western Europe, Asia, Oceania) was established a priori, for a target sample size of ten. This was considered adequate for providing a reasonable sample of current approaches and tools as the intent of this component of the study was to conduct an environmental scan and not to provide a comprehensive review of all existing KT tools and approaches.

For each contact or key informant identified, a maximum of two emails were sent to obtain an initial response. On the second attempt of contacting each key informant, an option was also provided to submit written responses to the interview questions, rather than scheduling a teleconference. If no response was received, attempts were made to identify an alternate organizational contact or key informant. If no response was received from the alternates, no further attempts were made to contact individuals.

Interview Guide Development
When developing interview guides, content was validated by an advisory committee with expertise in economics, statistics, health policy and KT. Interview questions were then piloted with a small number of policymakers who were not part of the recruited population. Comments were received on the clarity, completeness, readability and ease of use of the interview guide and feedback was incorporated.

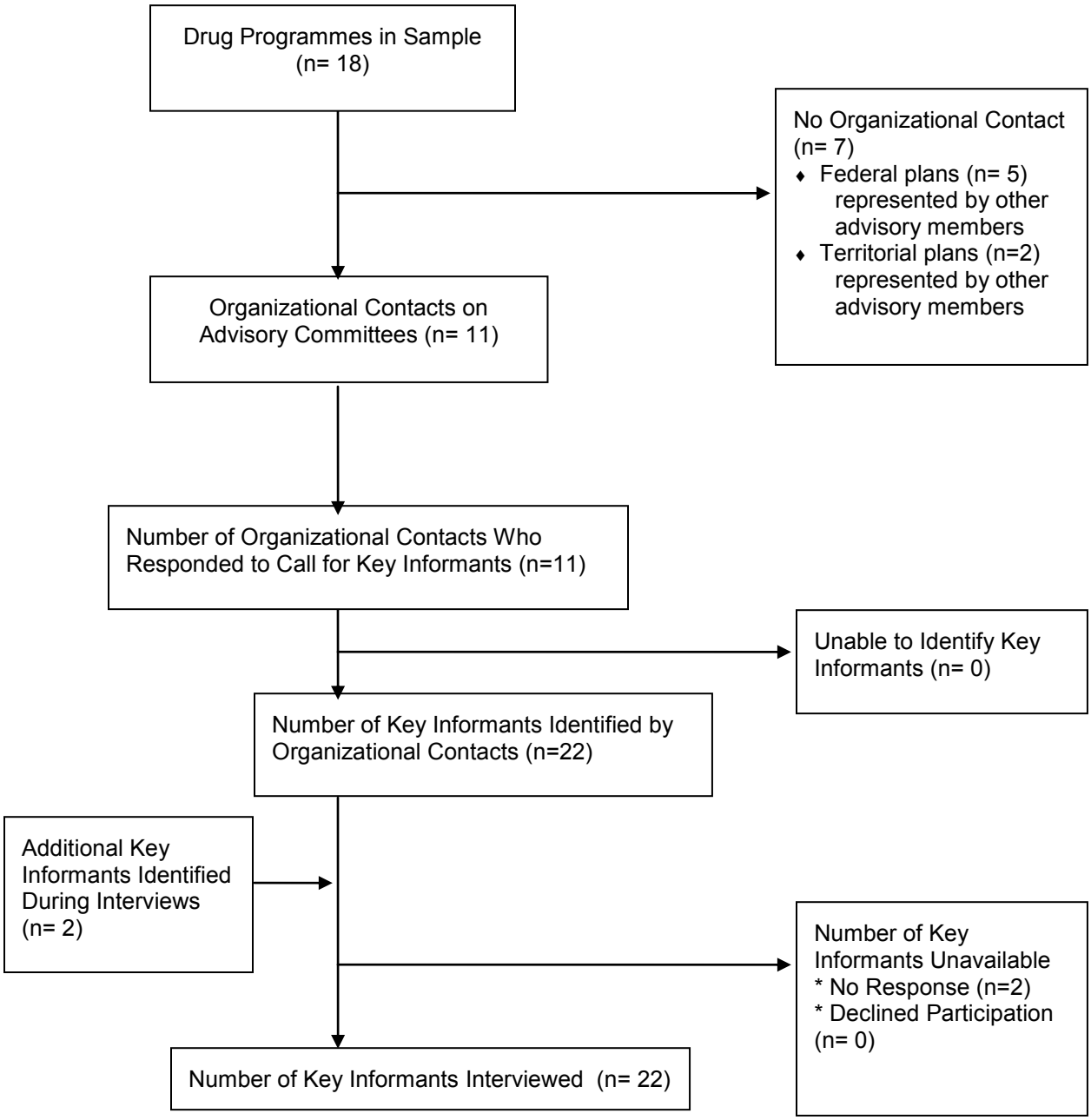
Qualitative Data Analysis:
When coding data, responses to each question were grouped based on questions in the interview guide. Additional codes were generated based on a general overview of the transcripts and themes from published literature. Codes were then were applied and iteratively revised. Iterations included inductive coding using the words of participants, organizing recurring responses into themes, identifying common ideas and responses across questions, displaying themes in coding matrices to identify patterns based on the characteristic of the respondent (e.g. type of policymaker) or the organization (e.g. size of the drug programme, geographic region of HTA organization) and re-coding and condensing themes.

CONSIDERATION OF SOME METHODOLOGICAL STRENGTHS AND LIMITATIONS

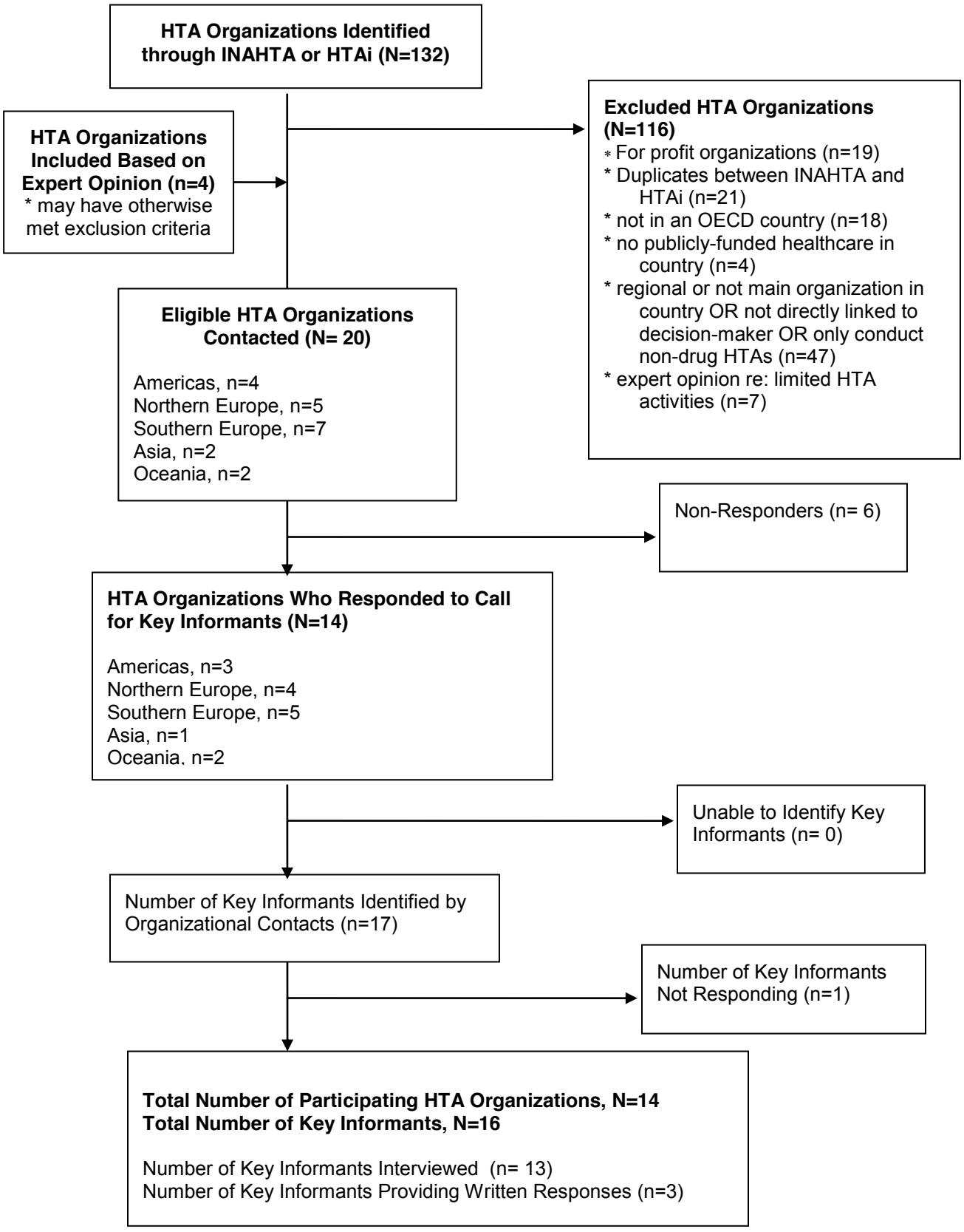
The qualitative nature of this study provides a rich evidence-base upon which to explore how to best develop approaches that enhance the presentation of economic evaluations and NMAs for policymakers. The strong methodological approaches used in the current study are also notable, including the high response rates, theory-grounded development based on the KTA framework and the free-flowing approach to interviews, which allowed for comprehensive identification and analysis of themes across interview questions.

Some methodological limitations exist in conducting key informant interviews including the potential influence of interviewer bias, recall bias and self-report biases (37-39). However, the consistency of themes across respondents and with existing literature suggests these biases may be limited and enhances credibility of these results. Although key informant interviews may provide a more narrow perspective than broader recruitment in the sample populations, given the dearth of research in this area, key informant approaches allow one to obtain a purposeful sample of knowledgeable individuals who 'think in terms of the organization" when responding to questions (39). Also, the variety of roles key informants filled in the HTA and drug-funding processes provides multiple complementary perspectives (e.g. senior management or directors, expert committee members, scientific staff directly supporting policymakers, senior individuals in HTA organizations) and allowed for confirmation of some themes across population.

Supplemental Figure 2. Key Informant Identification and Recruitment – Canadian Policymakers



Supplemental Figure 3. Key Informant Identification and Recruitment – HTA Organizations



Supplemental Box 2. Additional Detailed Results

ADDITIONAL DETAILED RESULTS

How Do Canadian Policymakers Use Economic Evaluations and NMAs?

Economic Evaluations: In discussing how the information was used, policymakers most frequently indicated that economic information was used to make funding decisions and to inform price negotiations. Policymakers frequently emphasized the value of having such information to inform consistent high quality decisions that allow them to "put your budget into where you're going to be most cost-effective, and be able to help the most people with the limited resources you have in order to provide the best outcomes for the patient" (I5). In addition, policymakers indicated that information provided at a national level required adaptation to the local/provincial context, at times, as "it's very difficult to design a study where you can…look at all different scenarios across the country" (I4). As one participant noted, "there [is] a level of contextualization that [is] done from the [provincial] perspective" (I6). This contextualization was frequently referred to by participants and was usually related to as population characteristics, funded comparators, drug utilization and available budgets.

NMAs: When considering how policymakers used NMAs to inform health policy decisions, adapting to the local context was identified similar to economic evaluations. However, a novel theme emerged in how NMAs are used. Although respondents acknowledged that head-to-head trials would be preferred, they are willing to accept NMAs in the absence of any other information. As one respondent noted, "what I see is people saying I wish there was…[but] in the absence of that, we need to look at what other information we have available" (I10) and another indicated that "we're pretty pragmatic" (I12).

What are Policymakers' Perceptions of Knowledge About Economic Evaluations and NMAs?

Economic Evaluations: It was frequently noted that many policymakers had a broad understanding of complex economic issues rather than detailed knowledge. This was sufficient for informing policy decisions, e.g. making funding recommendations or implementing funding decisions), when accompanied by access to expert resources. However, policymakers expressed an interest in more health economics training. Expert committee members noted this would bring "strength" to the committee. For example, one respondent noted that "one of the unknowns to me is the inside of the pharmacoeconomic review" (I14). In general, individuals noted that some aspects of the economic evaluations were challenging to understand and that there were always more questions on the economic aspects of an HTA than the clinical aspects. As one respondent noted, "they're not easy for non-economists to understand" (I15). Suggestions were made that "demystifying them in some way" (I21) would be beneficial. It was also discussed that "there's huge variability" (I17) in economic knowledge across individuals involved in the decision-making process but that having access to economic experts who can respond to questions was valuable.

NMAs: Similar themes emerged with respect to policymakers' knowledge of NMAs in that there was variable expertise but interest in learning more about NMAs. One of the unique themes that emerged related to NMAs was that, although the analyses are relatively novel and complex, there is emerging comfort and a growing knowledge about NMAs among policymakers. As one participant noted, "I think it's actually very much an evolution" (I19).

What are Policymakers' Barriers and Facilitators to Using and Understanding Economic Evaluations and NMAs?

Economic Evaluations: Common barriers included issues related to limited quality of evidence and models, lack of economic expertise and training for policymakers, health system challenges such as drug pricing and budgetary constraints and challenges adapting information to the local

context. For example, it was noted that "oftentimes their models don't take into account practical things that they need to be taken into account" (I9) or that "its hard to know where to even start in a negotiation when I don't even really know what the ICER was or I can't use their model" (I11).Common facilitators were having access to good quality information, economic experts and training opportunities. One aspect frequently raised regarding the quality of information was transparency of information and models: "the more transparent they are the better it is" (I13). Or, as one participant expressed it "the ability to deconstruct the model and see what it's sensitive to is what's most important to us" (I20).

NMAs: Similar themes emerged with respect to policymakers' barriers and facilitators to using and understanding NMAs in that quality, expertise and real-world relevance were frequently raised. An additional barrier, often cited related to NMAs was their lack of clarity. One participant noted, "they should also be transparent" (I21).

PATTERN ANALYSIS

Canadian Policymakers' Patterns of Response by Provincial Size and Policymakers' Role
Canadian policymakers patterns of response based on provincial/plan size and by policy-makers role were also considered. It was noted that there was agreement, regardless of size and role, on the importance of content aspects such as identifying the evidence-base, potential limitations and uncertainty in the results. While expert committee members focused on wanting specific details related to the evidence, policy-makers and senior support staff in government emphasized needing details related to implementation and adaption to the local or real-world context. When size was considered, it was noted that smaller programmes more frequently raised issues regarding lack of expertise and resources related to economics and NMAs.

International Patterns of Response by Geographic Region or Years of HTA Organization's Experience
Despite variability in geographic location of the 14 responding organizations, similar patterns existed across geographic regions. The amount of experiences most organizations had with HTA and drug funding policies was similar, i.e. between 10-20 years, therefore, no patterns existed.

Supplemental Table 1. Content Elements of Economic Evaluations and NMAs that are Valued by Canadian Policymakers

Economic Elements	NMA Elements
• Identifying comparators and their costs • Knowing place in therapy of treatment • Both industry and independent estimates of cost-effectiveness with reasons why estimates differ • Identifying limitations of analysis • What the assumptions were made in the analysis and how realistic they are • Identifying ethical context such as if unmet patient needs, end-of-life therapy, pediatric population, severity of disease, availability of alternative therapies • Understanding natural history of disease • Knowing the long-term treatment effects • Identifying priority of having the therapy available • ICER and most important factors influencing it • Conclusion on if its cost-effective or not with clear rationale • Different prices at which cost-effectiveness is achieved • ICERs in the best and worst case scenarios • Identifying factors that can change cost-effectiveness upon implementation • Results of sensitivity analyses • Knowing what costs are included in analysis • Knowing if one-time cost or if long-term costs • Understanding the potential cost savings • General descriptions (vs. detailed explanations) of technical aspects including modeling approach, perspective and time horizon	• Knowing when/if appropriate to conduct • Knowing what was compared and if all appropriate comparators included • Knowing level of evidence supporting comparators and historical context of therapies • What studies were included/excluded • Why specific studies were excluded • If all appropriate trials were captured/included • Critically evaluating if all assumptions were met • Heterogeneity of trials and if comparable, e.g. time when trials conducted, patient populations, other confounders • How/if data are applicable to real-world context, e.g. dosing, populations, treatment sequencing • If all influential factors incorporated • Sensitivity analyses for uncertain factors • Threats to internal/external validity – quality checklists • If conclusions are firm or hypothesis-generating • The power of the NMA to show a difference • How different types of evidence are valued • General descriptions of methodology (vs. detailed explanations) • Clarity on what is the reference treatment in the analysis • Understanding rankings • Understanding credible intervals • Understanding weightings
Abbreviations: ICER=incremental cost-effectiveness ratio; NMA=network meta-analysis	

Supplemental Table 2. Barriers and Facilitators to Canadian Policymakers' Use and Understanding of Economic Evaluations and NMAs

ECONOMIC	
Barriers	**Facilitators**
• Relevant comparators not included • Not knowing why trials are included or excluded from evidence base used in the model • No harms data • Don't know actual drugs prices or price point of cost-effectiveness • Lengthy • Non-drug comparator costs unknown • Assumptions questionable • Doesn't reflect clinical practice • Populations not generalizable • Needs local adaptation/context • Only theoretical cost savings, not actual savings	• Clarify standards or concepts in addition to providing ICER • Glossary or list of terms presented in layman's terms • Indication of priority of the drug • Tools that allow for consistent decision-making • Understanding context of the analysis • Providing price point in implementation tools • Understanding analysis steps leading to the ICER
NMA	
Barriers	**Facilitators**
• Correct comparators not included • Inconsistent approaches for different drugs • Knowing how to interpret NMAs • Finding relevant info in long reports • Uncertainty in analysis • Poor quality of included studies • Time differences in when trials conducted • Heterogeneity of studies • Doesn't answer the question • Doesn't include relevant real-world issues • Not relevant to local factors/real-world context • Terminology hard to understand	• Checklists/quality assessment tools • Understanding real world context of NMA • Methods guideline for NMA • Clear, concise critical appraisal • Transparent analysis/report • Understanding clinical context • When done thoughtfully • Done with methodological rigour • Can detect/validate assay sensitivity of NMA • Clear key messages • Details available • Limitations identified • Comparability of studies considered • Quality of studies considered • Assumptions explicit • Assumptions justified • Identifying trials even if not in statistical analysis • Concrete information, not theoretical

Supplemental Table 3. Detailed Approaches to the Knowledge Translation of Economic Evaluations and NMAs Identified by International Key Informants

General KT Approaches That Could be Applied to Economic Evaluations and NMAs
Types of Knowledge Tools Developed • Decision Aids • Infographics • PowerPoint Presentations • Mini-Quizzes for Clinicians • Social Media Tools • Clinical Practice Guidelines • Patient Leaflets • Clinician-Patient Guides
Targeted Approaches for Policymakers and Clinicians • Sections of documents targeted to clinicians and the public • Key messages for policymakers and for clinicians • Briefing notes for policymakers • Implication briefs for policymakers • Policy recommendations and synthesis documents • Policy scenario analysis based on evidence including course of actions and their possible effects (e.g. stakeholder responses, resources required, etc.)
Summary Documents • Executive summaries • Abstracts • One-page project briefs by researchers • Summaries Written by non-research staff • Public summaries of expert committee deliberations • Consumer-focused executive summaries • Summaries by professional writers • Press releases • Summary of evidence highlighting key issues that could make a difference in the decision
Approaches to Writing Reports • Writing guides and templates • Writing transparent scientific reports • Specific chapters in reports describe methodology

- Benefits and harms tables explicitly identifying trade-offs
- Type and amount of detail based on knowledge of target audience (e.g. researchers vs. lay audiences)
- Clear rationale for decisions
- Table summarizing crucial factors influencing the decision

Editorial and Review Processes

- Review by staff not involved in project
- Documents reviewed by a professional journalist
- External peer-review processes
- In-house quality assurance processes
- Use of communications staff

Stakeholder Interactions

- Involving key stakeholders in the beginning
- Engaging stakeholders in defining the project scope
- Dialogues with policymakers and people who will use the results
- Targeted meetings or presentations for different stakeholder groups after project
- Opportunity/teleconferences to discuss evidence with researchers
- Publicly available reports for all stakeholders

Other KT Strategies

- Assessing feasibility and potential impact of tools at the beginning of project
- Integrating researchers in tool development
- Multi-component or multi-targeted strategies
- Learning from larger, more established HTA organizations
- Hiring all staff based on abilities to write clearly
- Staff dedicated to implementation methods and research

Economic-Specific KT Approaches*

Summaries of Economic Evaluations

- Executive summaries for reports
- Short summary with recommendations based on ICERs and QALYs
- One-page public summary with basic methodology, data sources, overview of results and key points in the critique
- Summary of other HTA organizations economic evaluation and conclusion that it should be interpreted with caution b/c not always transferable between countries
- Media summaries including a bullet in laymen's terms on the health economic analysis
- Completed economic critical appraisal checklists
- Key messages from economic report
- Plain-language versions of reports

Approaches to Writing Economic Evaluations

- Standardized report templates
- Suggested wording in report templates
- Standard reports (i.e. not a plain language version) written in plain-language for policymakers without economic training
- Writing as simply as possible to keep people interested in the question
- Use simple figures accompanied by brief, simple text
- Providing more integration with the clinical review to demonstrate linkages
- Identifying willingness to pay based on the ICERs of already funded treatments in the same field and framing results based on the price point that would be need to achieve that same ICER with the new therapy
- Explaining ethical dimensions to economic modeling choices, e.g. selecting time horizon
- Clarifying that value judgments are for policymakers to decided and outside economists scope
- Focusing on the 'value for money' interventions provide, not just their costs
- Discussing opportunity costs that could be lost
- Explaining the impact of using relevant or less relevant comparators in economic evaluations
- Making comparisons with published economic evaluations and explaining why results differ
- Discussing model inputs, not just the final ICER
- Communicate strength of the underlying evidence
- Expressing cost of intervention in terms of numbers needed to treat from clinical evaluation
- Understanding how much treatment for one patient would cost

Implementation tools

- Planning tools for policymakers related to budget impact (to distinguish from cost-effectiveness)
- Graphics customized to different policymakers explaining what the economic evaluation meant and how to apply to their local setting

Stakeholder Interactions

- Discuss with and present reports to policymakers
- Being available to policymakers for questions
- Multi-stakeholder presentations/discussion of the project
- Provide verbal explanation before policymakers read reports in detail
- Communicating clear expectations for analyses submitted by pharmaceutical industry
- Conduct local economic evaluation only if clinical benefit suggests its useful
- Getting feedback from policymakers on economic report templates

Available Resources

- Economic evaluation guidelines

- Glossaries to explain economic terminology
- Report for policymakers and stakeholders on how to interpret an ICER in the decision-making context
- Report for policymakers and stakeholders on why a single willingness to pay threshold isn't used
- Report for policymakers on how to use ICER values in MCDA framework
- Training workshops
- Access to economic experts
- Gaining familiarity over time with economic issues and reports

NMA-Specific KT Approaches*

Summaries of NMAs

- Summary of existing NMAs with key Messages, strengths and weaknesses
- Summary paragraph on strengths and weaknesses of NMA
- Bottom-line key messages and recommendations
- Completed indirect comparison checklists

Approaches to Writing NMAs

- Figures explaining what evidence is available
- Explain treatment outcomes, similar to how done for traditional meta-analyses
- Provide cautionary statements on using results based on indirect evidence

Stakeholder Interactions

- Identifying local champions
- Engaging opinion leaders

Available Resources

- Guidelines or sections of guidelines on NMAs
- Detailed methods support documents
- Recommend NMA literature to policymakers for better understanding
- Policymakers have access to statisticians or NMA experts
- Internal researcher training in NMAs

*Some general KT approaches are listed under Economic or NMA approaches if specific examples were provided by key informants or if they specifically identified that the approach was used only for economic evaluations or NMAs

CHAPTER 8.

ENHANCING THE ACCESSIBILITY OF ECONOMIC EVALUATIONS AND NETWORK META-ANALYSES FOR POLICYMAKERS: DEVELOPMENT OF A TOOLKIT

CHAPTER OVERVIEW

Purpose and Relationship to Other Manuscripts:

One of the goals of this thesis is to propose a toolkit, guided by the Knowledge-to-Action framework, which enhances the relevance, transparency and accessibility of economic evaluations and NMAs for policymakers. Therefore, the objective of this sixth manuscript is to outline the development of a toolkit that supports the use of economic evaluations and network meta-analyses (NMAs) to inform policy decisions. This work integrates findings from:

- literature on knowledge, barriers, facilitators and approaches to communicating economic evaluations and NMAs (Chapters 3 and 5);
- researchers' standards for presenting economic evaluations and NMAs as identified in systematic reviews of guidelines (Chapters 4 and 6); and
- Canadian policymakers' informational needs and international best practices for the knowledge translation of economic evaluations and NMAs (Chapter 7).

The first component of the toolkit, a guide for researchers on writing economic evaluations and a guide for researchers on writing NMAs for policymakers, is developed and illustrated through a case study of an economic evaluation and NMA of osteoporosis therapies.

Supporting Material:

Appendix D. Osteoporosis Economic Evaluation (Coyle et al., 2014)

Appendix E. Osteoporosis Network Meta-Analysis (Maher et al., 2014)

Manuscript Status:

Draft, not submitted. The usability of tools will be further evaluated with end-users prior to submitting the manuscript for publication in order to provide a robust publication. However, assessing the usability of tools was outside the scope of this thesis' objectives. All information in this chapter, including the supplementary materials, is intended for journal submission.

Authors: Shannon M Sullivan (SMS), George Wells (GW), Doug Coyle (DC), Elaine MacPhail (EM), Barbara Davies (BD)

Author Roles (SMS, GW, DC, EM, BD) and Contributions:

SMS, GW, DC were responsible for the study conception.

SMS, GW, DC, EM, BD were responsible for the identification and selection of toolkit components

SMS, GW, DC, EM, BD were responsible for drafting the tool templates

SMS, GW and DC were responsible for applying the tools to the osteoporosis case study

SMS was responsible for drafting the manuscript.

All authors (SMS, GW, DC, EM, BD) provided comments on the draft manuscript, agree with the draft manuscript's results and conclusions, and read and approved the latest version of the manuscript.

NOTE: SMS was not the primary author on the two manuscripts reporting on the economic evaluation (Appendix D, Coyle et al, 2014) and the NMA of osteoporosis therapies (Appendix E, Maher et al., 2014). These manuscripts are complementary to this thesis and SMS was highly involved in their development. However, they do not represent work done specifically as part of this thesis' objectives.

Title: Development of a Toolkit to Enhance Accessibility of Economic Evaluations and Network Meta-Analyses for Policymakers

Short Title: Economic and NMA Toolkit

Authors:

Shannon M Sullivan,[1] George Wells,[1,2] Doug Coyle,[2] Elaine MacPhail,[3] Barbara Davies[4]

Affiliations:
1. University of Ottawa Heart Institute, Ottawa, Ontario
2. Department of Epidemiology & Community Medicine, University of Ottawa, Ottawa, Ontario
3. Canadian Agency for Drugs and Technologies in Health, Ottawa, Ontario
4. School of Nursing, University of Ottawa, Ottawa, Ontario

ABSTRACT

Background: Health policy decisions may be informed, in part, by health technology assessments (HTAs) that include economic evaluations or network meta-analyses (NMAs). Transparency is a key principle underlying fair and legitimate HTA processes and related policy decisions. This is especially relevant when complex methodologies, such as economic evaluations and NMAs are used to support evidence-informed policymaking. The objective of this work is to propose and develop an evidence-based toolkit that enhances the relevance, transparency and accessibility of economic evaluations and NMAs.

Methods: Development of the toolkit was guided by the Knowledge-to-Action framework. Findings were integrated from four systematic reviews and two sets of key informant interviews related to communicating and using economic evaluations and NMAs. Based on these integrated findings, a multi-disciplinary advisory panel identified, selected and prioritized development of a toolkit. A case study of an economic evaluation and NMA of osteoporosis therapies was applied to drafts of toolkit components.

Results: Three areas to be addressed by the toolkit were identified from systematic reviews and interviews: communication, education and accessibility. Six tools or approaches were proposed that could be applied to both economic evaluations and NMAs, i.e. 12 toolkit components, and targeted to either policymakers or researchers. These included: (1) guidance for researchers; (2) policymaker question and answer documents; (3) analysis templates; (4) learning modules; (5) tool repositories; and (6) lay peer-review processes. Drafts of the Guidance for Researchers tool were developed and illustrated using a case study of an economic evaluation and NMA of osteoporosis therapies.

Conclusions: Development of a multi-targeted, multi-strategy toolkit to enhance the relevance, transparency and accessibility of economic evaluations and NMAs is in progress. Engagement of key stakeholders and evaluation of the toolkit, including usability testing, are important next steps.

INTRODUCTION

Health policy decisions may be informed, in part, by health technology assessments (HTAs) that include economic evaluations and/or network meta-analyses (NMAs). Transparency is a key principle underlying fair and legitimate HTA processes and related policy decisions (Daniels & Sabin, 1997). Transparency requires not only providing sufficient information, but also providing information in an accessible and understandable format for end-users. This is especially relevant when complex methodologies, such as economic evaluations and NMAs, form the basis of an HTA and are used to support evidence-informed policymaking. Economic evaluations are conducted to evaluate the incremental costs and effects of alternative treatment strategies for the same patient population. When head-to-head trials with relevant comparators are not available, the comparative effectiveness of treatment strategies may be estimated through NMAs (Cooper et al., 2011). Economic evaluations and NMAs can be complex analytical techniques and audiences, i.e. knowledge users, have sometimes expressed challenges in understanding and applying these analyses to treatment and policy decisions (Cipriani et al., 2013). A systematic approach to knowledge translation (KT) and the development of tools and approaches targeted to policymakers in these areas could offer opportunities to enhance the relevance, transparency and accessibility of economic evaluations and NMAs.

KT can be defined as “a dynamic and iterative process that includes synthesis, dissemination, exchange and ethically-sound application of knowledge to improve health, provide more effective health services and products and strengthen the health care system.” (CIHR 2012) Knowledge that is created by researchers can sometimes be distilled into forms or tools that are more useful and allow the end-user to more easily apply research findings to improve health outcomes or health systems (Graham et al., 2006). Examples of tools that have been used to enhance the use of health-related evidence include clinical practice guidelines, patient decision aids and plain-language summaries (Brouwers et al., 2013). If more than one tool is developed to address the same problem, these tools can be assembled into a toolkit. Although, not commonly identified as knowledge tools, supportive tools exist in the field of economic evaluations and NMAs, including guideline documents, critical appraisal checklists, and training workshops (CADTH 2006; Chase et al., 2009; Craven et al., 2012; Evers et al., 2005). Although some tools have been developed for policymakers to guide them in their use of health-related evidence, many of these tools focus on systematic reviews and meta-analyses and have not expanded into more complex evidence syntheses such as economic evaluations and NMAs (Lavis et al., 2009; DECIDE Collaboration 2013). Other approaches have been developed to enhance uptake of economic evaluations and NMAs in decision-making (Caro et al., 2014; DePourville et al., 2005; ISPOR Guidelines Around the World 2014); however, routine and explicit integration of KT principles in these fields could

enhance the accessibility of economic evaluations and NMAs to stakeholders and increase the value of these analyses in health policy decision-making.

The objective of this work is to propose and develop an evidence-based toolkit that enhances the relevance, transparency and accessibility of economic evaluations and NMAs, guided by the Knowledge-to-Action framework (KTA). The KTA framework is based on a comprehensive review of planned-action models (Graham et al., 2006, Straus et al., 2013). In the KTA framework, after an assessment of barriers and facilitators to knowledge use has been conducted, tools are developed that are both tailored to the target audience and adaptable to the local setting. Systematic reviews of the literature on economic evaluations and NMAs have been conducted, along with qualitative key informant interviews with Canadian policymakers and HTA organizations around the world, to inform the development of this toolkit (Sullivan et al., 2014a; Sullivan et al., 2014b; Sullivan et al., 2014c; Sullivan et al., 2014d; Sullivan et al., 2014e). Integrating findings from this research, a toolkit is proposed with the goals of (1) assisting policymakers involved with drug funding decisions in understanding and identifying key elements of complex methodologies such as economic evaluations and NMAs; (2) supporting policymakers in determining how to adapt complex analyses to their local drug funding policy context; and (3) enhancing the interface between researchers, i.e. knowledge producers, and policymakers, i.e. knowledge users, when sharing information on economic evaluations and NMAs.

METHODS

KTA Framework

The KTA framework guided the toolkit development. Key steps in the framework include: identifying the problem; adapting knowledge to the local context; assessing barriers and facilitators; selecting interventions; tailoring interventions to target audiences; and implementing, monitoring, evaluating and sustaining the intervention (Graham et al., 2006; Straus et al., 2013).

Identifying Toolkit Components

Data from four systematic reviews and two sets of key informant interviews were analyzed and results were previously reported (Sullivan et al., 2014a; Sullivan et al., 2014b; Sullivan et al., 2014c; Sullivan et al., 2014d; Sullivan et al., 2014e). Findings were integrated across these six studies and synthesized to identify key issues to address in the toolkit and potential tools to develop.

Selecting Toolkit Components

A multi-disciplinary advisory panel that included a statistician, a health economist, a KT specialist and a health policy advisor guided the development of the toolkit, i.e. the Toolkit Advisory Panel. Based on a review of the integrated findings and detailed descriptions of potential tools the panel evaluated each of the proposed toolkit components. Evaluation criteria were adapted from criteria developed by the Agency for Healthcare Research and Quality (AHRQ) for the selection of Effective Health Care program topics (Whitlock et al., 2009). The five criteria applied to inform the selection of toolkit components and prioritize their development were: feasibility; appropriateness to the drug funding context; potential value; importance; novelty and lack of duplication with existing tools. Detailed definitions of the adapted criteria used to develop this toolkit are outlined in Table 1.

INSERT TABLE 1.

Final selection of tools was also considered in relation to the specific objectives of the toolkit, which were:

1) To assist policymakers involved with drug funding decisions in understanding and identifying key elements of complex methodologies such as economic evaluations and NMAs;
2) To support policymakers in determining how to adapt complex analyses to their local drug funding policy context and;
3) To enhance the interface between researchers, i.e. knowledge producers, and policymakers, i.e. knowledge users, when sharing information on economic evaluations and NMAs.

A judgment of high, medium or low was assigned to each criterion, which is similar to other standardized evaluation and assessment processes (e.g. Cochrane Risk of Bias) that use a three category ranking system (Higgins et al., 2010). The advisory panel discussed the proposed rankings until consensus was gained.

Tool Development Case Study: Osteoporosis Therapies

Templates were developed for two tools prioritized by the Toolkit Advisory Panel. The advisory panel reviewed drafts of the templates and tools to ensure they aligned with the toolkit objectives, key criteria and core elements of the proposed tools. The technical content of the tools as it related to the economic evaluation and NMA of osteoporosis therapies was validated with content

experts. Feedback from the Panel and content experts was incorporated to ensure that the tools were appropriate for further evaluation by end-users.

An economic evaluation and NMA of osteoporosis therapies (Coyle et al., 2014; Maher et al., 2014) was applied to two components of the toolkit as an illustrative case study. The topic of the case study (i.e. osteoporosis therapies) was selected because it was considered to have sufficient complexity to demonstrate key economic and NMA concepts that were identified in systematic reviews and key informant interviews. In addition, a case study in which access was available to the original economic model and source NMA data was considered essential. This provided the flexibility to manipulate the economic and NMA data when creating new presentation approaches for the knowledge tools. This would not have been possible if the case study relied on an example in the published literature.

Osteoporosis NMA

A Bayesian NMA based on a random effects model was conducted that includes 34 randomized trials (two active-controlled, one placebo-controlled three-arm trial and 31 placebo-controlled two-arm trials) evaluating six osteoporosis therapies (alendronate, risedronate, etidronate, ibandronate, zoledronic acid, denosumab) and placebo. It is based on five outcomes: vertebral fracture, non-vertebral fracture, wrist fracture, hip fracture and withdrawals due to adverse events. Denosumab and zoledronic acid were the most effective therapies in preventing vertebral fracture in postmenopausal osteoporotic women when compared with placebo. However, alendronate was the only therapy effective in preventing all of vertebral, non-vertebral and hip fractures, the latter of which is a more clinically relevant outcome. Conclusions on comparative safety are limited by relying only on WDAEs and the challenges synthesizing data on rare but relevant serious adverse events in NMAs.

Osteoporosis Economic Evaluation

A cost utility analysis was performed of six different treatment strategies for osteoporosis (alendronate, etidronate, risedronate, zoledronic acid, denosumab and no pharmacotherapy) using a Markov model incorporating data on the fracture risk, and the associated costs, mortality and disutility and the effect of treatment. The base case was a 75 year old osteoporotic women with no previous fracture and a bone mineral density (BMD) 2.5 standard deviations below the young female mean. Uncertainty was assessed through detailed probabilistic and deterministic sensitivity analysis. For osteoporotic women with a previous fracture and/or aged over 70 and/or with a BMD at least 3.5 standard deviations below the young adult mean, alendronate is the optimal pharmacotherapy based on a threshold value of $50,000 per QALY. For all other women, no pharmacotherapy is cost effective.

RESULTS

KTA Framework

The application of the KTA framework to the development of the economic evaluation and NMA toolkit is outlined in Figure 1; and key steps completed in knowledge creation and in the action cycle are described in more detail below.

INSERT FIGURE 1

Knowledge Creation

- *Knowledge Inquiry*: Four systematic reviews (communicating economic evaluations; economic evaluation guidelines; communicating NMAs; NMA guidelines) and two sets of key informant interviews (Canadian policymakers, international HTA organizations) were conducted.
- *Knowledge Synthesis*: Findings from all six studies were integrated to inform tool identification. These findings were also aligned with KT frameworks and considered in the context of existing literature and resources to propose 12 possible tools.
- *Knowledge Tools*: A toolkit containing tools for policymakers and/or researchers to enhance the relevance, transparency and accessibility of economic evaluations and NMAs was proposed. Drafts of two tools were developed and illustrated using a case study of osteoporosis therapies and are ready for further evaluation.

Action Cycle

- *Identification of Problem:* The original hypothesis was that economic evaluations and NMAs can be used to inform health policy decisions but their transparency and accessibility to policymakers and other non-technical audiences may be limited due to their complexity. This was confirmed, in part, by systematic reviews of the literature and by Canadian and international key informant interviews.
- *Adapt Knowledge to Local Context:* Key informant interviews with policymakers identified areas relevant to local context (e.g., available comparators, populations, drug utilization, drug pricing, implementation issues) and real-world context (e.g., evidence gaps, differences in clinical practice) that could be incorporated into approaches for enhancing the transparency and accessibility of economic evaluations and NMAs.
- *Assess Barriers and Facilitators*: Systematic reviews and key informant interviews identified potential barriers and facilitators to the use and understanding of economic evaluations and

NMAs that could be addressed through knowledge tools, including factors related to content of analyses, contextual factors, clarity of presentation and available expertise.

- *Select Tools and Tailor to Target Audience*: Systematic reviews and key informant interviews identified how economic evaluations and NMAs could be tailored to policymakers to enhance use and accessibility. Twelve potential tools targeting either policymakers or researchers were identified as potential toolkit components and two were prioritized for development.
- The implementation, monitoring, evaluation and sustainability of the toolkit are not addressed in this paper.

Identifying Toolkit Components

An integrated review of key findings from the systematic reviews (Sullivan et al., 2014a; Sullivan et al., 2014b; Sullivan et al., 2014c; Sullivan et al., 2014d) and key informant interviews (Sullivan et al., 2014e), identified three key areas to be addressed by the toolkit: (1) improving the communication of economic evaluations and NMAs; (2) developing educational initiatives; and (3) increasing access to existing resources related to economic evaluations and NMAs.

A more detailed assessment of individual results from each of the systematic reviews and qualitative interview sets was supportive of the overarching findings described in Tables 2 and 3 (see also Supplementary Materials 1, Supplementary Tables 1 and 2). Other areas were identified based on individual study results, including methodological developments, stakeholder interactions and health systems pressures. However, these issues were considered outside the scope of this toolkit.

During the integrated review of findings, it was noted that the themes of content, context and clarity that had emerged during key informant interviews with Canadian decision-makers (Sullivan et al., 2014e) were also relevant to the systematic review findings (Sullivan et al., 2014a; Sullivan et al., 2014b; Sullivan et al., 2014c; Sullivan et al., 2014d). How these themes map onto results from the four systematic reviews and which of the themes the strategy and tool would most address is also noted in Tables 2 and 3, as well as Supplementary Tables 1 and 2. In this regard:

- **Content** referred to any technical issues or quality issues that could be addressed in tools;
- **Clarity** referred to any issues related to formatting of information, approaches to explaining information or different ways of focusing information for end-users; and
- **Context** referred to any issues related to adapting analyses to the real-world application of decisions or the local setting of the end-user.

Six different tools or approaches were identified that could be applied to both economic evaluations and NMAs, for a total of 12 different toolkit components. A general description of

each of the tools is provided below. More detailed descriptions of each of the tools, outlining key criteria and core elements specific to either economic evaluations or NMAs are provided in Supplementary Materials 1 (Supplementary Table 3).

1) Guidance for Researchers: This tool addresses issues around communications and interactions between policymakers and researchers. The target audience is researchers. It provides guidance to researchers on how to write for policymakers about economic evaluations or NMAs, emphasizing overarching principles identified through this work and providing examples of how principles can be actioned. This is one way that these findings can be disseminated to the target audience of researchers. This tool is distinct from documents such as reporting guidelines that identify specific elements to be included in a technical analysis report on economic evaluations. Rather, this tool focuses more on how to explain what matters most to policymakers in an economic evaluation or NMA. This could be a resource that academic journals or others make available to researchers when publishing economic evaluations or NMAs that may subsequently be used to inform policy and treatment decisions.

2) ***Question and Answer (Q&A) Document for Policymakers:*** This tool addresses issues around communications and interactions between policymakers and researchers. The target audience for the completed Q&A document is policymakers. Instructions on how to complete the Q&A document will be provided to researchers and will include a list of standard questions that require answers. The Q&A document presents information in a format that is easily understood by non-technical audiences such as policymakers and highlights implications for policy and decision-making. The information in the answers would not be novel and would likely already be provided by a technical analysis. However, in that format, the information may be either lost in technical details or not framed in a way that the relevance to a policy decision is clear. Questions in the Q&A document are tailored to what policymakers want and/or need to know most to inform a policy decision. The answer to each question is structured around three components relating to the policy context: (1) identifying what is the issue; (2) explaining why it matters; and (3) suggesting how the issue can be addressed. This format will focus the answers created by researchers and keep responses brief. This tool would not replace the usual technical analysis researchers provide to journals or to others, but would complement it. This could be included as supplementary material to a journal article or made available through the author.

3) Technical Analysis Template: This tool addresses issues around communication. The primary target audience is researchers, as the template provides them with more guidance on

how to focus the technical analysis. It incorporates economic evaluation and NMA presentation formats that were identified in systematic reviews (Sullivan et al., 2014a; Sullivan et al., 2014b; Sullivan et al., 2014c; Sullivan et al., 2014d). The completed technical analysis is directed toward a technical audience and so has a greater technical focus than the Q&A document. Therefore, it will include details that are not directly relevant to a policy decision. However, this template ensures that, as part of the technical analysis, the most policy-relevant details are not overlooked and are included in the written analysis. This allows policymakers to extract and reference these technical details when they need a more in-depth resource than the Q&A document.

4) ***Tool Repository:*** This tool addresses issues around accessibility of information on economic evaluations and NMAs. The target audience is policymakers. Tools and approaches that are helpful to policymakers in these areas may already exist but are not easily accessible. Collecting existing tools and making them available in a single accessible location such as an online repository would be valuable.

5) ***Learning Modules:*** This tool addresses issues around education (i.e., increasing policymakers' knowledge of technical analyses) and communication (i.e., allowing policymakers to understand how researchers communicate and think). The target audience is policymakers. Learning modules developed as PowerPoint slide decks (or more interactive technological approaches) that provide a step-by-step walk through of decisions made at each step of the analysis to understand how researchers think when conducting an economic evaluation or NMA. Researchers unfamiliar with NMAs who have uncertainty around how to conduct and/or interpret NMAs may also be a target audience. Versions of the NMA modules could also be developed and tailored towards them.

6) ***Lay Peer-Review Process:*** This tool addresses issues around communication. The primary target audience is HTA organizations, journals or researchers wanting to incorporate a lay peer-review process. End-users that could participate as peer reviewers are also a target audience. Indirectly, policymakers would benefit from the non-technical peer-review of economic evaluation and NMA analyses that they use to inform policy decisions. Using a lay reader to provide a non-technical peer review of analyses produced by technical writers (i.e., researchers), could help identify issues that may be unclear. Lay readers will require guidance on how to review an analysis: what to look for and what questions to ask.

Selecting Toolkit Components

Proposed rankings for each of the 12 potential toolkit components are outlined in Table 4 with the detailed rationale for each assessment is described in Supplementary Materials 1 (Supplementary Table 4). In general, NMA tools were generally ranked lower on domains related to appropriateness and importance due to their still low, although growing, use to inform policy decisions. Novelty was considered high for most tools with the exception of learning modules and analysis templates as versions of these tools likely already exist. Guidance for Researchers and the Policymaker Q&A document received the highest rankings overall. However, the policymaker documents were considered slightly less feasible to develop, considering some of the contextual issues that would require involvement of an identified end-user during development of the tool in order to optimally incorporate these issues.

INSERT TABLE 3

Based on the assessments of the advisory panel, all proposed tools were considered valuable components of a toolkit and could be developed. However, the Guidance for Researchers, as applied to both economic evaluations and NMAs, was identified as top priority for development.

Case Study: Guidance for Researchers Applied to an Economic Evaluation and NMA of Osteoporosis Therapies

Based on the recommendations of the Toolkit Advisory Panel, templates for the economic and NMA Guidance for Researchers were developed and are described below. The tools and the illustrative examples from the economic evaluation and NMA of osteoporosis therapies are provided in Supplementary Materials 2 and Supplementary Materials 3 due to their length.

Economic Guidance for Researchers: Guidance was developed that directs economists to focus on seven key principles when writing an economic evaluation that may inform policy decisions. Two principles relate to the content of analyses: explaining uncertainty in the results; and describing limitations in the analysis. Two principles relate to contextual factors: identifying gaps between the evidence and real-world practice; and identifying local issues that could influence analysis results and conclusions. Three principles relate to clarity and presentation of information: disaggregating information when possible; emphasizing key messages; and focusing on policymakers' needs. Specific examples of how this guidance could apply to an economic evaluation of osteoporosis therapies are provided in the tool to illustrate concrete examples of the principles in action (Supplementary Materials 2).

NMA Guidance for Researchers: Guidance was developed that directs researchers to focus on seven key principles when writing a NMA that may inform policy decisions. Two principles relate to the content of analyses: describing the available evidence; and describing limitations of the analysis. Two principles relate to contextual factors: identifying gaps between the evidence and real-world practice; and identifying local issues that could influence analysis results and conclusions. Three principles relate to clarity and presentation of information: judiciously selecting information to present when there is a large volume of data; emphasizing key messages; and focusing on policymakers' needs. Specific examples of how this guidance could apply to a NMA of osteoporosis therapies are provided in the tool to illustrate concrete examples of these principles in action (Supplementary Materials 3).

DISCUSSION

Summary of Findings

Integrated findings from systematic reviews (Sullivan et al., 2014a; Sullivan et al., 2014b; Sullivan et al., 2014c; Sullivan et al., 2014d) and key informant interviews (Sullivan et al., 2014e) suggested that a toolkit to enhance the accessibility of economic evaluations and NMAs should address three areas: improving communication, developing educational initiatives and increasing access to existing resources. Six tools or approaches were identified, for a total of 12 toolkit components when applied to both economic evaluations and NMAs: (1) guidance for researchers when writing analyses, (2) policymaker Q&A documents, (3) analysis templates that incorporate that incorporate standard presentation formats but focus analyses to policymakers' needs, (4) learning modules for policymakers, and researchers in the case of NMAs, (5) a tool repository for existing resources and (6) a lay peer-review process. One of the tools, Guidance for Researchers, was drafted and applied to an economic evaluation and NMA of osteoporosis therapies. The tools describe seven principles related to the content, context or clarity of an economic evaluation or NMA. Five principles were common to the economic tool and the NMA tool: explaining limitations, identifying potentially relevant local issues, relating evidence to real-world applications, emphasizing key messages and focusing on policymakers' needs. The economic tool also included principles related to uncertainty in the results and disaggregating information while the NMA tool included principles related to describing the available evidence and selecting and layering large volumes of information.

Comparison with Other Literature

When comparing the development of this toolkit with other literature, three specific areas were considered: existing knowledge or support tools; differences and similarities between tools for

economic evaluations and those for NMAs and consideration of the use of KT tools to inform health policy.

Comparison with Existing Tools and Resources

Based on the integrated review of findings from systematic reviews and key informant interviews, the toolkit was designed to enhance communication, education and accessibility of economic evaluations and NMAs. The proposed toolkit provides some ways in which these goals could be achieved. However, other tools and approaches may exist that could be leveraged during the development of this toolkit. For example, exploring general approaches in these three areas, which are not specific to economic evaluations and NMAs, could be considered (e.g. best practices in educational initiatives). It is also possible that supports in the areas of communication, education and accessibility of economic evaluations and NMAs were developed previously but not identified in systematic reviews we conducted (Sullivan et al., 2014a; Sullivan et al., 2014c). This may have occurred if tools were not formally evaluated or described in the published literature. For example, the International Society for Pharmacoeconomics and Outcomes Research (ISPOR) has a list of tools for healthcare decision-makers on its website (ISPOR Tools for Healthcare Decision-Makers, 2014), many of which would not have met the inclusion criteria of our systematic reviews. Also, tools focusing on communicating systematic reviews would not have been included in our systematic reviews but could be adapted to economic evaluations and NMAs (Lavis et al., 2009; DECIDE Collaboration 2013). Finally, the focus of the systematic reviews we conducted was on approaches to communicating economic evaluations and NMAs but tools related to education, accessibility and other approaches to supporting the use of economic evaluations and NMAs in policy decisions were not comprehensively searched and identified (Caro et al., 2014; DePourville et al., 2005). More novel approaches to sharing information such as infographics and social media approaches (e.g., Twitter, journal blogs and Facebook) could also be studied in more depth to determine how they could apply to economic evaluations and NMAs (Grande et al., 2014; Speigelhalter et al., 2011).

During further development of the toolkit, consideration around how existing resources can be built upon or incorporated into the toolkit should be considered. For example, although learning modules related to economic evaluations and NMAs exist, tailoring their content to real-world applications and scenarios for policymakers could enhance their impact. Other resources, such as critical appraisal checklists, may exist but not be accessible to policymakers. Therefore, a tool repository that organizes, collates and comments on these resources could be of value. Integrating literature on policymakers' preferences and systematic reviews of guidelines also led to proposals for tools that enhance the interface between researchers and policymakers.

Developing tools such as the Guidance for Researchers is novel as many communication and educational approaches target policymakers and not those who produce the evidence, i.e., the researchers. Although reporting guidelines exist that researchers may use (Moher et al., 2011), which are intended to improve transparency of analyses, findings from our systematic reviews suggest that this is insufficient for good communication to non-technical audiences or to inform specific policy decisions. Discovering approaches that meet both researchers' and policymakers' needs in sharing information on economic evaluations and NMAs would contribute to capacity building initiatives in these fields and further exploration of this interface through deliberative dialogues could be valuable (Boyko et al., 2012). While tailoring knowledge tools requires an understanding of end-users and their needs, some of the findings in literature related to communicating technical information to lay audiences suggests there is variation in individual preferences for information formats and in individuals' numeracy levels (Trevena et al., 2006). Therefore, presenting information in multiple formats can be beneficial. Although this can be challenging, especially when trying to judiciously select information, ensuring a mix of tables, figures and text, as has been done in some published NMAs (Tan et al., 2013; Bafeta et al., 2014), should be further emphasized when writing analyses.

Comparison of Findings Related to Economic Evaluations and Related to NMAs

Although differences were identified between economic evaluation and NMA findings, similar tools and approaches were selected for both economic evaluations and NMAs. However, the content of the tools can be tailored to the issues in each subject area. For example, content issues were more frequently identified with NMAs, reflecting the novelty of the methodology and challenges researchers experience when conducting and interpreting NMAs (Salanti et al., 2012). In addition, policymakers had less experience using NMAs than economic evaluations to inform policy decisions, so it was understandable that contextual factors were less frequently identified for NMAs. The longer experience using economic evaluations to inform health policy decisions has allowed a greater body of literature to be established in this area and more specific consideration of some of the barriers and facilitators to using economic evaluations (Sullivan et al., 2014a). Although the NMA literature is not as well-developed (Sullivan et al., 2014c), application of many of the principles in economic evaluations could likely be applied to NMAs since the introduction of different types of complex evidence-based analyses to inform health policy share similar features. Because NMA is a novel area where methodology is still emerging, some NMA studies suggested that greater knowledge and understanding among researchers about NMA approaches is still needed (Abdelhamid et al., 2012; Coleman et al., 2012). Therefore, some of the tools being developed for policymakers in this toolkit, such as the tool repository and learning modules, could also have versions targeted towards researchers.

Although similar concepts were sometimes identified for economic evaluations and NMAs (e.g. describing limitations, emphasizing key messages), a few concepts were prioritized differently when incorporated into the Guidance for Researchers tool. For example, layering of details may be even more important in NMAs than economic evaluations given the large volume of data and studies that sometimes must be synthesized. For example, one published osteoporosis NMA contains 116 RCTs (Murad et al., 2012). Explaining uncertainty associated with results, while applicable to NMAs, was considered particularly important for economic evaluations, given the extent of modeling that occurs in cost-effectiveness analyses (Roberts et al., 2012) and the need to understand how that uncertainty applies when making concrete policy decisions.

Finally, when considering themes of content, context and clarity, findings related to economic evaluations were distributed evenly across these three areas. For NMAs, there was a much larger focus on content, followed by some attention to clarity but little mention of context. This distribution for NMAs could be due to the current emphasis on standardizing methodologies in the evolving field of NMAs as well as the slow but emerging uptake of NMAs to inform policy decisions (Salanti et al., 2012). This is supported by key informant interviews with Canadian policymakers, which indicated that despite the complexity of NMAS, acceptance and comfort with the method is growing. Although the primary target audience for the toolkit was policymakers, the review of knowledge, barriers and facilitators to using NMAs revealed that there is considerable uncertainty among researchers related to conducting and interpreting NMAs, which was not observed with economic evaluations (Abdelhamid et al., 2012; Coleman et al., 2012). This was supported by key informant interviews with HTA organizations around the world, suggesting that their current focus is on educating internal resources and researchers on NMAs methodologies before focusing on how to best share information with policymakers and other knowledge users.

Considerations on the use of Knowledge Tools to Inform Health Policy

Qualitative studies exploring policymakers' knowledge found that although individuals had some conceptual knowledge of economic evaluations, there was a desire for increased knowledge of how to best implement and apply these concepts and/or economic evaluation results to actual policy decisions. Therefore, toolkit components that incorporate contextual information or that create links between the evidence and how policy-related factors will influence the results of economic evaluations would be of value in enhancing applied knowledge of health economics (Bombard et al., 2012; Chambers & Wilson, 2012; Lavis et al., 2009; Lehoux et al., 2007). Although this issue was not identified for NMAs, it is likely to be similar and did not emerge only because these are early days for the use of NMAs in health policy.

The importance of content, context and clarity in approaches to communicating economic evaluations and NMAs first emerged in key informant interviews with policymakers. During the integrated review, these themes were also found to be applicable to findings from systematic reviews, suggesting that these are important cornerstones in approaches to enhancing the communication of economic evaluations and NMAs. Of note, these themes are similar to those identified in the Promoting Action on Research Implementation in Health Services (PARIHS) framework for successfully implementing evidence in practice. However, in PARIHS, the three concepts are more specifically referred to as evidence (i.e. content), context and facilitation (i.e. clarity) (Rycroft-Malone & Bucknall, 2012). In-depth consideration of how our findings compare and contrast with PARIHS and other frameworks for implementing evidence in practice or policy could further elucidate systematic principles for the knowledge translation of economic evaluations and NMAs.

A number of health systems issues were noted that are outside the scope of issues that can be addressed by researchers and in this toolkit (e.g. drug pricing policies, budgetary constraints, market authorization requirements for evidence). However, tools such as the policymaker Q&A document that would allow for these issues to be identified and their potential impact on the analysis noted, can help policymakers better place economic evaluations and NMAs into their local context. Although contextual issues may be challenging for researchers to identify when there is not an identified receptor for the analysis, transferability checklists have previously been developed when attempting to adapt economic evaluations to alternate settings and could be used to guide what issues are identified in a policymaker Q&A document (Chase et al., 2009). Incorporating ethical and social issues into HTAs has been considered at a systems level and researchers may require greater support from end-users in considering these issues (Chambers & Wilson, 2012; Lehoux 2007). As well, greater engagement and interactions between researchers and policymakers could lead to a better, shared understanding of these issues and how best to approach them in economic evaluations and NMAs that inform policy decisions (CIHR 2012). The use of deliberative dialogues is one approach that could be applied to obtain this shared understanding (Boyko et al., 2012; Bombard et al., 2011).

Strengths and Limitations of Tool Development

While different tools and approaches may be used in practice to enhance accessibility of economic evaluations or NMAs, few tools and toolkits have such a robust evidence base to support their use and development. Furthermore, no such toolkit focusing on these topic areas has been developed to date and the evidence from systematic reviews and key informant interviews suggests this toolkit could have an important impact. In the current resource-constrained environment of drug funding policy, enhanced use and accessibility of economic

evaluations and NMAs has considerable value (Sullivan et al., 2011; Niezen et al., 2006). Identifying approaches that target not only policymakers but also the researchers who produce economic evaluations and NMAs is a strength as it recognizes the important role and necessary engagement of multiple stakeholders to enhance the accessibility of economic evaluations and NMAs. Although only one of the six tools was developed in this paper, the strength of proposing a multicomponent toolkit, and not simply one tool is that different components and strategies may be more successful than others at enhancing transparency, which can be assessed during evaluation of the toolkit (Straus et al., 2013). Development of additional components of the toolkit, evaluation of the usability of tools with researchers and policymakers and applying the tools to more than one case study will allow incorporation of broader experiences and stakeholder perspectives on the proposed toolkit. Other tool development processes have used larger and more established working groups than ours to guide tool development, e.g., guidelines groups, ISPOR, SUPPORT, Grading of Recommendations Assessment, Development and Evaluation (GRADE). However, we anticipate broader engagement of stakeholders in future steps of the toolkit evaluation. We also had the strength of a theory-based approach to toolkit development by following the KTA framework as well as an existing process from AHRQ on topic prioritization. (Straus et al., 2013; Whitlock et al., 2009)

Policy Implications and Future Research

In light of the toolkit that is being proposed to enhance the accessibility of economic evaluations and NMAs, some policy implications and future research that should be considered including:

1) **Engaging Stakeholders in Toolkit Development**: While policymakers participated in initial interviews to explore barriers and facilitators to using economic evaluations and NMAs, they should be further engaged in the toolkit development by participating in evaluations of the usability and acceptability of draft tools. In addition, some tools were developed that target researchers. Researchers are key stakeholders that were not widely engaged in the qualitative studies informing this toolkit (Sullivan et al., 2014e), with the exception of some members of the Toolkit Advisory Panel and some key informants from HTA organizations. Broadly soliciting the perspectives of researchers on interactions with policymakers about economic evaluations and NMAs would strengthen development of the toolkit.
2) **Exploring Partnerships to Ensure Sustainability of Toolkit**: The toolkit was developed as a research undertaking, however, identifying partnerships that could support the toolkit and interested end-users such as journals, academic collaborations and/or HTA organizations would contribute to the sustainability and success of the toolkit

in enhancing the transparency and accessibility of economic evaluations and NMAs. Implementation of the toolkit into various organizations will require development of a long-term sustainability plan that considers the adaptability of the toolkit over time and across settings as well as the degree of organizational and stakeholder support for greater transparency of economic evaluations and NMAs.

3) **Evaluating Impact of the Toolkit**: Systematic reviews that were conducted to identify approaches to communicating economic evaluations and NMAs (Sullivan et al., 2014a; Sullivan et al., 2014c) revealed that even when such approaches exist, their impact is rarely evaluated. Therefore, upon development, ensuring that the impact of the toolkit and/or its components is robustly assessed will be an important contribution both to policy environments and the academic literature. Ongoing monitoring of the toolkit following its implementation will also provide an opportunity for feedback, revisions and continuous quality improvement of the toolkit.

Conclusions

Numerous approaches to enhancing the accessibility of economic evaluations and NMAs were identified that will target strategies to improve communication, education and access to existing resources. Development of a toolkit is in progress and engagement of key stakeholders and evaluation of the toolkit will be important next steps.

REFERENCES

Abdelhamid A, Loke Y, Parekh-Bhurke S, Chen Y, Sutton A, Eastwood A, Hollanda R, Song F. Use of indirect comparison methods in systematic reviews: a survey of Cochrane review authors. Research Synthesis Methods 2012, 3: 71-79.

Bafeta A, Trinquart L, Seror R, Ravaud P. Reporting of results from network meta-analyses: methodological systematic review. BMJ. 2014 Mar 11;348:g1741.

Bombard Y, Abelson J, Simeonov D, Gauvin FP. Eliciting ethical and social values in health technology assessment: A participatory approach. Soc Sci Med. 2011, 73:135-44.

Boyko J, Lavis J, Abelson J, Dobbins M, Carter N. Deliberative dialogues as a mechanism for knowledge translation and exchange in health systems decision-making. Social Science & Medicine 75 (2012) 1938-1945.

Brouwers, M. C., Stacey, D., & O'Connor, A. M. (2013). Knowledge translation tools. In S. Straus, J. Tetroe & I. D. Graham (Eds.), *Knowledge translation in health care: Moving from evidence to practice* (pp. 50-62). Chichester: Wiley Blackwell BMJ Books.

CADTH. Guidelines for the economic evaluation of health technologies: Canada [3rd Edition]. Ottawa: Canadian Agency for Drugs and Technologies in Health; 2006.

Canadian Institutes of Health Research (2012). Guide to Knowledge Translation Planning at CIHR: Integrated and End-of-Grant Approaches.

Canadian Institutes of Health Research. Guide to Researcher and Knowledge User Collaboration in Health Research. Ottawa: CIHR.

Caro J, Eddy DM, Kan H et al. Questionnaire to assess relevance and credibility of modeling studies for informing health care decision making: an ISPOR-AMCP-NCP Good Practice Task Force report. Value in Health 2014, 17: 174-182.

Chambers D, Wilson P. A framework for production of systematic review based briefings to support evidence-informed decision-making. Syst Rev 2012; 1: 32.

Chase D, Rosten C, Turner S, Hicks N, Milne R. Development of a toolkit and glossary to aid in the adaptation of health technology assessment (HTA) reports for use in different contexts. *Health Technol Assess* 2009; **13**(59).

Cipriani A, Higgins JP, Geddes JR, Salanti G. Conceptual and technical challenges in network meta-analysis. Ann Intern Med 2013, 159 (2): 130-137.

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.

Cooper NJ, Peters J, Lai MC, Juni P, Wandel S, Palmer S, Paulden M, Conti S, Welton NJ, Abrams KR, Bujkiewicz S, Spiegelhalter D, Sutton AJ. How valuable are multiple treatment comparison methods in evidence-based health-care evaluation? Value Health. 2011 Mar-Apr;14(2):371-80.

Coyle DC, Sullivan SM, Wells GA. Cost-Effectiveness of Pharmacological Treatments for Osteoporosis in Canada. Submitted, 2014.

Craven M, Allsop M, Morgan S, Martin J. Engaging with economic evaluation methods: insights from small and medium enterprises in the UK medical devices industry after training workshops. Health Research Policy and Systems 2012, 10:29.

Daniels N & Sabin J. Limits to health care: Fair procedures, democratic deliberations and the legitimacy problem for insurers. Philosophy & Public Affairs 1997; 303-350

DECIDE Collaboration, 2013.Available at: http://www.decide-collaboration.eu. [Last accessed: August 2014].

de Pouvourville G, Ulmann, P, Nixon J, Boulenger S, Glanville J, Drummond M. The Diffusion of Health Economics Knowledge in Europe. The EURONHEED (European Network of Health Economics Evaluation Database) Project. Pharmacoeconomics 2005; 23 (2): 113-120.

Evers S, Goossens M, de Vet H, van Tulder M, Ament A. Criteria list for assessment of methodological quality of economic evaluations: Consensus on Health Economic Criteria. Int J Technol Assess Health Care. 2005 Spring;21(2):240-5.

Graham ID, Logan J, Harrison MB, Straus SE, Tetroe J, Casell W et al. Lost in knowledge translation: time for a map: J Contin Ed Health Prof 2006; 26(1):13-24.

Grande D, Gollust SE, Pany M, Seymour J, Goss A, Kilaru A et al. Translating Research For Health Policy: Researchers' Perceptions And Use Of Social Media *Health Affairs*, 2014, 33:1278-1285.

Higgins J, Altman D, Gotzsche P, Juni P, Moher D, Oxman A, Savovic J, Schulz K, Weeks L, Sterne J, Cochrane Bias Methods Group, Cochrane Statistical Methods Group. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. BMJ 2011;343:d5928

International Society for Pharmacoeconomics and Outcomes Research. Pharmacoeconomic (ISPOR) guidelines around the world. Available from: http://www.ispor.org/PEguidelines/index.asp [Last accessed May 2014]

International Society for Pharmacoeconomics and Outcomes Research (ISPOR). ISPOR Tools for Healthcare Decision-Makers. Available from: http://www.ispor.org/workpaper/UseOfResearch_index.asp [Last accessed August 2014]

Lavis JN, Oxman AD, Lewin S, Fretheim A: SUPPORT Tools for evidence-informed health Policymaking (STP). Introduction. Health Research Policy and Systems; 2009, **7**(Suppl 1):I1.

Lehoux P & Williams-Jones B. Mapping the integration of social and ethical issues in health technology assessment. International Journal of Technology Assessment in Health Care 2007, 23: 9-16.

Maher M, Sullivan SM, Hsieh S, Peterson J, Wells GA. The effectiveness and safety of denosumab compared with bisphosphonates in the prevention of osteoporosis in post-menopausal women: a systematic review and network meta-analysis. Submitted to: Annals of Rheumatic Disease, 2014.

Moher D, Weeks L, Ocampo M, Seely D, Sampson M, Altman DG et al. Describing reporting guidelines for health research: a systematic review. J Clin Epidemiol 2011, 64: 718-742.

Murad MH, Drake MT, Mullan RJ, Mauck KF, Stuart LM, Lane MA et al. Comparative effectiveness of drug treatmetns to prevent fragility fractures: a systematic review and network meta-analysis. J Clin Endocrinol Metab, 2012, 97: 1871-1880.

Niezen M, Stolk E, Steenhoek A, Uyl-De Groot C. Inequalities in oncology care: Economic consequences of high cost drugs. Eur J of Cancer, 2006, 42: 2887–2892.

Roberts M, Russell LB, Paltiel AD, Chambers M, McEwan P, Krahn M; ISPOR-SMDM Modeling Good Research Practices Task Force. Conceptualizing a model: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force--2. Value Health, 2012, 15: 804-11.

Rycroft-Malone J & Bucknall T. Models and Frameworks for Implementing Evidence-Based Practice: Linking Evidence to Action. Wiley-Blackwell: West Sussex, UK. 2010.

Salanti G. Indirect and mixed-treatment comparison, network, or multiple-treatments meta-analysis. Many names, many benefits, many concerns for the next generation evidence synthesis tool. Research Synthesis Methods 2012, 3: 80-97.

Spiegelhalter D, Pearson M, Short I. Visualizing Uncertainty About the Future. Science 2011, 333: 1393-1400.

Straus S, Tetroe J & Graham ID (Eds.), Knowledge translation in health care: Moving from evidence to practice, 2nd edition. Oxford: Blackwell Publishing, 2013.

Sullivan SM, Akbar Z, Sabharwal M, Bell C, Coyle D (a). Presenting and Communicating Health Economic Information: A Systematic Review. Submitted to: Int J Technol Assess Health Care, 2014.

Sullivan SM, Wells GA, Coyle D (b). What Guidance are Economists Given on How to Present Economic Evaluations? A Systematic Review. Submitted to: Value in Health, 2014.

Sullivan SM, Coyle D, Wells GA (c). Knowledge, Barriers, Facilitators and Approaches to Communicating Network Meta-Analyses: A Systematic Review. Submitted to: Research Synthesis Methods, 2014.

Sullivan SM, Coyle D, Wells GA (d). What Guidance are Researchers Given on How to Present Network Meta-Analyses? A Systematic Review. Submitted to: PLOS One, 2014.

Sullivan SM, Wells GA, Coyle D, MacPhail E, Davies B (e). Canadian and International Perspectives on the Knowledge Translation of Economic Evaluations and Network Meta-Analyses for Policymakers. A Qualitative Analysis. Submitted to: Int J Technol Assess Health Care, 2014.

Sullivan R, Peppercorn J, Sikora K, Zalcberg J, Meropol N, Amir E. Delivering affordable cancer care in high-income countries. Lancet Oncology, 2011,12: 933-980.

Tan SH, Bujkiewicz S, Sutton A, Dequen P, Cooper N. Presentational approaches used in the UK for reporting evidence synthesis using indirect and mixed treatment comparisons. J Health Serv Res Policy 2013, 18: 224-232.

Trevena L, Davey H, Barratt A, Butow P, Caldwell P. A systematic review on communicating with patients about evidence. J Evalu Clin Practice 2006, 12: 13-23.

Whitlock EP, Lopez SA, Chang S, et al. Identifying, selecting, and refining topics. In: Agency for Healthcare Research and Quality. Methods Guide for Comparative Effectiveness Reviews [posted April 2009]. Rockville, MD.

TABLES AND FIGURES

Table 1. Tool Selection Criteria Adapted from the Agency for Healthcare Research and Quality (AHRQ) Criteria for the Selection of Effective Health Care Program Topics*

Criteria	Description
Feasibility	- Effectively uses existing research and knowledge - Requires minimal resources - Can be developed in a timely way
Appropriateness (to drug funding context)	- Could be reasonably used in existing structures of drug funding context - Could support or address current challenges or priorities in drug funding context (e.g. Pricing negotiations, budget silos, etc.
Importance	- Strongly supported by one or more stakeholder groups or of high public interest - Represents important uncertainty for decision makers - Incorporates issues around both clinical benefits and potential clinical harms - Can address issues around important variations, controversies or inconsistencies
Potential Value	- For significant health impact: - For significant economic impact: - Proposed topic exists within a clinical, consumer, or policy-making context that is amenable to evidence-based change - Tool could be an appropriate vehicle for change - Risks or harms from lack of prioritization or development of tool - Topic that has clear implications for resolving important dilemmas in health and healthcare decisions made by one or more stakeholder groups
Novelty	- No existing tools - Existing tools address only some components or perspectives - Novelty of topic or issue

***** Used with permission from the Agency for Healthcare Research and Quality (AHRQ), Whitlock et al., 2009.

Figure 1. Knowledge-to-Action Framework: Action Cycle and Knowledge Creation Activities Applied to the Development of an Economic Evaluation and NMA Toolkit. Shaded boxes represent work completed in this paper; unshaded boxes represent plans for future research.

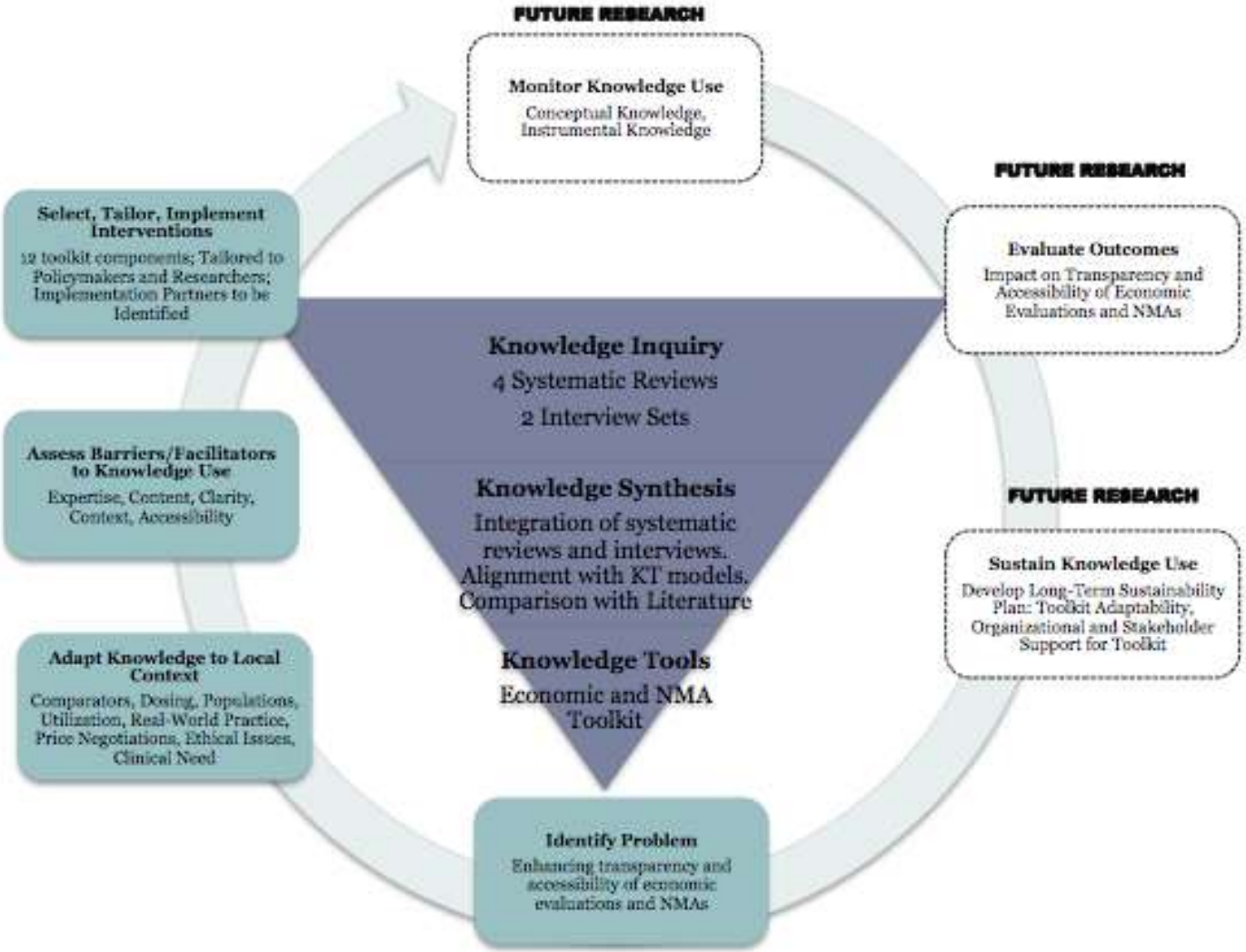


Table 2. Integrated Economic Findings to Inform Identification of Potential Economic Tools

FINDING	HOW FINDING COULD BE ADDRESSED	STRATEGY	POTENTIAL TOOL	MOST RELEVANT TO EITHER CONTENT, CONTEXT OR CLARITY
Economic Communication SR				
Disaggregate information better understood	Improve Communication	Encourage use of disaggregate information in current reports	Analysis Templates	Clarity
	Enhance Education	Incorporate into teaching approaches	Learning Modules	
Knowing practical, real-world applications of the evidence important	Improve Communication	Create formats that frame issues from a real-world perspective	Policymaker Q&A	Context
Optimal formatting of the evidence and tailoring details to end-user helps	Improve Communication	Get economists to think more about how to best present information and what audience needs	Researcher Guidance	Clarity
		Get input from end-users on readability and understandability of information	Lay Peer-Review	
Economic Guidelines SR				
Presentation formats currently targeted to technical audiences	Improve Communication	Emphasize thinking about end-users needs	Researcher Guidance	Clarity
Focus on disaggregating information for transparency purposes	Improve Communication	Encourage use of disaggregated details	Analysis Templates	Clarity
		Emphasize other merits of disaggregating information	Researcher Guidance	
Issues related to uncertainty and limitations can be complex to explain	Improve Communication	Enhance presentation of uncertainty and limitations	Analysis Templates	Content
	Enhance Education	Increase understanding of key limitations for end-users	Learning Modules	

FINDING	HOW FINDING COULD BE ADDRESSED	STRATEGY	POTENTIAL TOOL	MOST RELEVANT TO EITHER CONTENT, CONTEXT OR CLARITY
Canadian and International Interviews – Economic Findings				
Understanding key content including assumptions, model inputs and critical appraisal points	Improve Communication	Emphasizing key content details in technical reports	Analysis Templates	Content
	Enhance Education	Increase knowledge of critical appraisal points and impact of limitations	Learning Modules	
	Increase Accessibility	Facilitate access to existing critical appraisal resources and educational materials	Tool Repository	
Relating analysis to real-world and local situations	Improve Communication	Identifying links between analyses and practical implications	Policymaker Q&A	Context
Clarity and consistency of information valued	Improve Communication	Encourage economists to think about presentation of information, not just content	Researcher Guidance	Clarity
Few specific economic KT approaches but strong need for better communication with policymakers and other stakeholders	Improve Communication	Create toolkit addressing communication issues	Identified Toolkit Components	Content, Clarity, Context
	Increase Accessibility	Ensure dissemination of toolkit to international HTA agencies	Tool Repository	

Table 3. Integrated Network Meta-Analysis (NMA) Findings to Inform Identification of Potential NMA Tools

FINDING	HOW FINDING COULD BE ADDRESSED	STRATEGY	POTENTIAL TOOL	MOST RELEVANT TO EITHER CONTENT, CONTEXT OR CLARITY
NMA Communication SR				
Tailor to end-user using established formatting approaches and policymaker preferences from other fields	Improve Communication	Get researchers to think more about how to best present information and what audience needs	Researcher Guidance	Clarity
		Get input from end-users on readability and understandability of information	Lay Peer-Review	
Explain uncertainty including issues related to emerging methods, NMA assumptions and variability in results	Improve Communication	Enhance presentation of uncertainty	Analysis Templates	Content
	Enhance Education	Incorporate into educational initiatives	Learning Modules	
Build on similarities with traditional MAs	Improve Communication	Incorporate presentation formats that emphasize similarities or contrast key differences between NMAs and MAs	Analysis Templates	Content
	Enhance Education	Educate about similarities and differences between NMAs and traditional meta-analyses	Learning Modules	
NMA Guidelines SR				
Both researchers and non-technical audiences need more support in understanding NMAs	Enhance Education	Increase educational initiatives targeted at both researchers and end-users	Learning Modules	Content
Details important to emphasize include trial characteristics, network diagrams, key assumptions and differences between direct and indirect results	Improve Communication	Enhance presentation of key NMA details	Analysis Templates	Content
	Enhance Education	Ensure researchers and end-users understand key NMA details	Learning Modules	

FINDING	HOW FINDING COULD BE ADDRESSED	STRATEGY	POTENTIAL TOOL	MOST RELEVANT TO EITHER CONTENT, CONTEXT OR CLARITY
Presentation formats align with policymakers' needs and are tailored to non-technical audiences	Improve Communication	Emphasize thinking about end-users' needs	Researcher Guidance	Clarity
Canadian and International Interviews - NMAs				
Understanding key content such as included and excluded evidence and critical appraisal points	Improve Communication	Emphasize key content in analysis	Analysis Templates	Content
	Enhance Education	Increase knowledge of critical appraisal and impact of limitations	Learning Modules	
	Increase Accessibility	Facilitate access to existing critical appraisal resources and NMA educational materials	Tool Repository	
Relating analyses to real-world and local situations	Improve Communication	Identifying links between analyses and practical implications	Policymaker Q&A	Context
Clarity and consistency of information valued	Improve Communication	Encourage researchers to think about presentation of information, not just content	Researcher Guidance	Clarity
NMA-specific KT approaches needed but primary focus is ensuring researchers understand appropriate methodologies	Improve Communication	Create toolkit to enhance understanding and communication of NMA approaches	Identified Toolkit Components	Content, Clarity, Context
	Increase Accessibility	Ensure dissemination of toolkit to international HTA agencies	Tool Repository	
Abbreviations: KT=knowledge translation; MA=meta-analysis; NMA=network meta-analysis; Q&A=Question and Answer				

Table 4. Rankings to Inform Selection of Toolkit Components and Priority for Tool Development.

Potential Tool	Feasibility	Appropriateness to Drug Funding Context	Potential Value	Importance	Novelty and Lack of Duplication
ECONOMIC					
Researcher Guidance *	High	High	High	High	High
Policymaker Q&As	Medium	High	High	High	High
Technical Analysis	Medium	High	Medium	Medium	Low
Tool Repository	Medium	Medium	High	Medium	Medium
Learning Modules	Medium	Medium	High	Medium	Low
Lay Peer-Review	Medium	High	Medium	Medium	Medium
NMA					
Researcher Guidance *	High	High	Medium	High	High
Policymaker Q&As	Medium	High	Medium	High	High
Technical Analysis	Low	High	Medium	Medium	High
Tool Repository	Medium	Medium	High	Medium	High
Learning Modules	Medium	Medium	High	Medium	Medium
Lay Peer-Review	Medium	High	High	Medium	Medium

* Tools prioritized for further development and provided in this paper.

SUPPLEMENTARY MATERIAL 1.

Supplementary Table 1. Identification of Potential Tools Based on Individual Economic Study Findings

Individual Economic Study Finding	How Finding Could Be Addressed	Related to Either Content, Context or Clarity*	Potential Tool
ECONOMIC COMMUNICATIONS SR (See Sullivan et al., 2014a for Details)			
Economic Knowledge			
Interest to learn more	Education	Content	Learning Modules
Variable economic training	Education	Content	Learning Modules
Focus on general concepts	Education Communication	Content Clarity	Learning Modules
Difficulties applying concepts	Education Communication	Context	Learning Modules Policymaker Q&A
Low knowledge economic modeling	Education	Content	Learning Modules
Want to understand how different factors impact economic model	Education Communication	Content Context Clarity	Learning Modules Technical Analyses Policymaker Q&A Researcher Guidance
Low numeracy skills in general population	Education Communication	Content Clarity	Learning Modules Researcher Guidance
Clinicians - less favorable attitudes to economics	Communication	Context	Policymaker Q&A
Barriers			
No graphical presentation of models	Communication	Content	Technical Analyses
Lack of clarity in presentation of economic information	Communication	Clarity	Researcher Guidance
Better metrics than QALY and ICER	Methods	NA	Out-of-Scope
Too much focus on methodological rigour and not enough on key principles	Communication	Content Context Clarity	Technical Analyses Policymaker Q&A Researcher Guidance
Results not well disseminated	Accessibility	NA	Tool Repository
Everyday language has different meaning in economics	Education Communication	Content Clarity	Learning Modules Lay Peer Review Researcher Guidance
Technical language not understood	Education Communication	Content Clarity	Learning Modules Researcher Guidance Lay Peer Review
Oversimplification of complexities	Education Communication	Content Clarity	Learning Modules Technical Analyses Policymaker Q&A Researcher Guidance
Can't communicate results to policymakers to help understand value	Education Communication	Content Clarity	Learning Modules Policymaker Q&A

Individual Economic Study Finding	How Finding Could Be Addressed	Related to Either Content, Context or Clarity*	Potential Tool
			Researcher Guidance
Can't explain when queried about results	Communication	Clarity	Researcher Guidance
Analyses complicated and hard to understand	Education	Content	Learning Modules
Industry bias	Communication Stakeholder Interactions	Content	Technical Analyses Policymaker Q&A Out of Scope
Time required	System Pressures	NA	Out of Scope
Lack of expertise	Education	Content	Learning Modules
Political, ethical or organizational issues	System Pressures Communication	Context	Out of Scope Policymaker Q&A
Decision-making process issues	System Pressures	NA	Out of Scope
Limited relevance	Communication	Context Clarity	Policymaker Q&A Researcher Guidance
Limited availability or accessibility	Accessibility	NA	Tool Repository
Credibility of methods	Education Communication Methods	Content	Learning Modules Technical Analyses Out of Scope
Not timely	System Pressures	NA	Out of Scope
Challenges communicating	Communication	Clarity	Researcher Guidance
Budgetary constraints	Communication System Pressures	Context Clarity	Policymaker Q&A Researcher Guidance Out of Scope
Conflicts with clinical values	Communication Values Stakeholder Interactions	Context Clarity	Policymaker Q&A Researcher Guidance Decision Aid Out of Scope
Unaware of drug costs	Communication Systems Pressures	Content Clarity	Technical Analyses Policymaker Q&A Out of Scope
Facilitators			
Reporting of results is easy to understand	Communication	Clarity	Technical Analyses Policymaker Q&A Researcher Guidance Lay Peer Review
Model assumptions are clear	Communication	Content Clarity	Technical Analyses Policymaker Q&A Lay Peer Review
Standardized presentation format	Communication	Clarity	Technical Analyses Policymaker Q&A Researcher Guidance Lay Peer Review

Individual Economic Study Finding	How Finding Could Be Addressed	Related to Either Content, Context or Clarity*	Potential Tool
Disaggregated list of costs and benefits	Communication	Clarity	Technical Analyses Researcher Guidance
Explaining practical relevance of results	Communication	Context	Policymaker Q&A Researcher Guidance
Dissemination by media	Accessibility Communication	NA	Out of Scope? Tool Repository? Lay Peer Review?
Greater awareness of results by decision makers	Communication Accessibility	Context	Policymaker Q&A Tool Repository
Strong Collaborations	Stakeholder Interactions	NA	Out of Scope
Economic resources available	Systems Pressures	NA	Out of Scope
Good communication	Communication	Clarity	Researcher Guidance
Political awareness and ethical factors	Communication Systems Pressures	Context	Policymaker Q&A Out of Scope
Timely	Systems Pressures	NA	Out of Scope
High quality information	Communication	Content	Technical Analyses Policymaker Q&A
Relevant	Communication	Context Clarity	Policymaker Q&A Researcher Guidance
Training available	Education Accessibility	Content	Learning Modules Tool Repository
Trusted or credible information and sources	Education Communication Stakeholder Interactions	Content	Learning Modules Technical Analyses Policymaker Q&A
Supportive policies	Systems Pressures	NA	Out of Scope
Flexible budgets	Systems Pressures	NA	Out of Scope
ECONOMIC GUIDELINES SR (See Sullivan et al., 2014b for Details)			
Recommendations on How to Present Information			
Cross-referencing and citations of sources	Communication	Content	Technical Analyses
Detailed table of data, variables, costs, resource use	Communication	Content Clarity	Technical Analyses Researcher Guidance
Transparent model description	Communication	Content Clarity	Technical Analyses Researcher Guidance
Clear methods and assumptions	Communication	Content Clarity	Technical Analyses Researcher Guidance
Model or decision tree diagram	Communication	Content Clarity	Technical Analyses Researcher Guidance
Disaggregated and aggregated results	Communication	Content Clarity	Technical Analyses Researcher

Individual Economic Study Finding	How Finding Could Be Addressed	Related to Either Content, Context or Clarity*	Potential Tool
			Guidance
Incremental costs and benefits	Communication	Content	Technical Analyses
Total costs and benefits	Communication	Content	Technical Analyses
Tables of results	Communication	Content Clarity	Technical Analyses Researcher Guidance
Graphical displays encouraged	Communication	Content Clarity	Technical Analyses Researcher Guidance
Sensitivity analyses tables	Communication	Content Clarity	Technical Analyses Researcher Guidance
Tornado diagrams to highlight drivers	Communication	Content Clarity	Technical Analyses Researcher Guidance
Identifying influential parameters impacting results	Communication	Content Context	Technical Analyses Policymaker Q&A
Cost-effectiveness acceptability curve	Communication	Content	Technical Analyses
Scatter plot, cost-effectiveness plane	Communication	Content Clarity	Technical Analyses Researcher Guidance
Most Frequent Reporting Recommendations	Communication	Content	Technical Analyses
Primary perspective: healthcare payer or societal	Communication	Content Context	Technical Analyses Policymaker Q&A
Population: approved indication or clearly specified target population	Communication	Content	Technical Analyses
Relevant subgroup analyses	Communication	Content Context	Technical Analyses Policymaker Q&A
Time horizon considering disease characteristics and period of consequences and benefits	Communication	Content	Technical Analyses
Details on modeling and assumptions	Communication	Content	Technical Analyses
Outcomes of QALYs, natural events, LYs or quality of life	Communication	Content Context	Technical Analyses Policymaker Q&A
Sensitivity analyses on parameters with greatest uncertainty based on univariate, multivariate or probabilistic approaches	Communication	Content	Technical Analyses Policymaker Q&A
Sensitivity analyses conducted over a range of confidence intervals, credible range of variability or extreme best/worst case values	Communication	Content Context	Technical Analyses Policymaker Q&A
Advice on Presenting to Non-Economists			
Calculations explained in a straightforward manner	Communication	Clarity	Researcher Guidance Policymaker Q&A Technical Analyses
Disaggregated results more easily understood	Communication	Clarity	Researcher Guidance

Individual Economic Study Finding	How Finding Could Be Addressed	Related to Either Content, Context or Clarity*	Potential Tool
			Policymaker Q&A Technical Analyses
Transparently presented so can generalize results	Communication	Clarity	Researcher Guidance Policymaker Q&A Technical Analyses
2 to 3 page summary written in an accessible way	Communication	Clarity	Researcher Guidance Policymaker Q&A Technical Analyses
Conclusions written to be understood by a non-specialist and should not use technical terms	Communication	Clarity	Researcher Guidance Policymaker Q&A Technical Analyses
Graphical presentations help readers	Communication	Clarity	Researcher Guidance Policymaker Q&A Technical Analyses
INTERVIEW RESULTS (See Sullivan et al., 2014e for Details)			
Content Elements Valued by Policymakers			
Identifying comparators and their costs	Communication	Content Context	Policymaker Q&A Researcher Guidance
Knowing place in therapy of treatment	Communication	Content Context	Policymaker Q&A Researcher Guidance
Both industry and independent estimates of cost-effectiveness with reasons why estimates differ	Communication	Content Context	Policymaker Q&A Researcher Guidance
Identifying limitations of analysis	Communication	Content	Policymaker Q&A Researcher Guidance
What the assumptions were made in the analysis and how realistic they are	Communication	Content Context	Policymaker Q&A Researcher Guidance
Identifying ethical context such as if unmet patient needs, end-of-life therapy, pediatric population, severity of disease, availability of alternative therapies	Communication	Content Context	Policymaker Q&A Researcher Guidance
Understanding natural history of disease	Communication	Content	Policymaker Q&A Researcher Guidance
Knowing the long-term treatment effects	Communication	Content	Policymaker Q&A Researcher Guidance
Identifying priority of having the therapy available	Communication	Content Context	Policymaker Q&A Researcher Guidance
ICER and most important factors	Communication	Content	Policymaker Q&A

Individual Economic Study Finding	How Finding Could Be Addressed	Related to Either Content, Context or Clarity*	Potential Tool
influencing it			Researcher Guidance
Conclusion on if its cost-effective or not with clear rationale	Communication	Content Clarity	Policymaker Q&A Researcher Guidance
Different prices at which cost-effectiveness is achieved	Communication	Content Context	Policymaker Q&A Researcher Guidance
ICERs in the best and worst case scenarios	Communication	Content Context	Policymaker Q&A Researcher Guidance
Identifying factors that can change cost-effectiveness upon implementation	Communication	Content Context	Policymaker Q&A Researcher Guidance
Results of sensitivity analyses	Communication	Content	Policymaker Q&A Researcher Guidance
Knowing what costs are included in analysis	Communication	Content	Policymaker Q&A Researcher Guidance
Knowing if one-time cost or if long-term costs	Communication	Content Context	Policymaker Q&A Researcher Guidance
Understanding the potential cost savings	Communication	Content Context	Policymaker Q&A Researcher Guidance
General descriptions of technical aspects including modeling approach, perspective and time horizon	Communication	Content	Policymaker Q&A Researcher Guidance
Barriers			
Relevant comparators not included	Communication	Context	Policymaker Q&A Researcher Guidance
Not knowing why trials are included or excluded from evidence base used in the model	Communication	Context Clarity	Policymaker Q&A Researcher Guidance
No harms data	Communication	Content	Policymaker Q&A Researcher Guidance
Don't know actual drugs prices or price point of cost-effectiveness	Communication Systems Pressures	Content Context	Technical Analyses Policymaker Q&A Researcher Guidance
Lengthy	Communication	Clarity	Policymaker Q&A Researcher Guidance
Non-drug comparator costs unknown	Systems Pressures Communication	Context	Out of Scope Policymaker Q&A
Assumptions questionable	Communication	Content	Policymaker Q&A Researcher

Individual Economic Study Finding	How Finding Could Be Addressed	Related to Either Content, Context or Clarity*	Potential Tool
			Guidance
Doesn't reflect clinical practice	Communication	Context	Policymaker Q&A Researcher Guidance
Populations not generalizable	Communication	Context	Policymaker Q&A Researcher Guidance
Needs local adaptation/context	Communication	Context	Policymaker Q&A Researcher Guidance
Theoretical savings, not actual savings	Systems Pressures Communication	Context	Policymaker Q&A Researcher Guidance
Facilitators			
Clarify standards or concepts in addition to providing ICER	Communication	Content	Technical Report Policymaker Q&A
Glossary or list of terms presented in layman's terms	Education	Content	Learning Modules Tool Repository
Indication of priority of the drug	Communication	Context	Policymaker Q&A
Tools that allow for consistent decision-making	Communication	Context Clarity	Researcher Guidance Policymaker Q&A
Understanding context of the analysis	Communication	Context	Policymaker Q&A
Providing price point in implementation tools	Communication	Content Context	Technical Analyses Policymaker Q&A
Understanding analysis steps leading to the ICER	Education	Content	Learning Modules
International Approaches Specific to Economic Evaluations			
Executive summaries for reports	Communication	Clarity	Technical Analyses
Short summary with recommendations based on ICERs and QALYs	Communication	Clarity	Policymaker Q&A
One-page public summary with basic methodology, data sources, overview of results and key points in the critique	Communication	Clarity	Policymaker Q&A
Summary of other HTA organizations economic evaluation and conclusion that it should be interpreted with caution b/c not always transferable between countries	Communication	Context	Policymaker Q&A
Media summaries including a bullet in laymen's terms on the health economic analysis	Communication	Clarity	Policymaker Q&A
Completed economic critical appraisal checklists	Communication	Content	Technical Analyses
Key messages from economic report	Communication	Clarity	Policymaker Q&A
Plain-language versions of reports	Communication	Clarity	Technical Analyses
Standardized report templates	Communication	Clarity	Technical Analyses
Suggested wording to use is in report templates	Communication	Clarity	Technical Analyses

Individual Economic Study Finding	How Finding Could Be Addressed	Related to Either Content, Context or Clarity*	Potential Tool
All reports written in plain-language for policymakers without economic training	Communication	Clarity	Technical Analyses
Writing as simply as possible to keep people interested in the question	Communication	Clarity	Policymaker Q&A
Use simple figures accompanied by brief, simple text	Communication	Clarity	Policymaker Q&A
Providing more integration with the clinical review to demonstrate linkages	Communication	Content Clarity	Technical Analyses Policymaker Q&A
Identifying willingness to pay based on the ICERs of already funded treatments in the same field and framing results based on the price point that would be need to achieve that same ICER with the new therapy	Communication	Context	Technical Analyses Policymaker Q&A
Explaining ethical dimensions to economic modeling choices, e.g. selecting time horizon	Communication	Context	Policymaker Q&A
Clarifying that value judgments are for policymakers to decided and outside economists scope	Stakeholder Interactions	Context	Policymaker Q&A Out of Scope
Focusing on the 'value for money' interventions provide, not just their costs	Communication	Clarity	Policymaker Q&A Researcher Guidance
Discussing opportunity costs that could be lost	Communication	Context	Policymaker Q&A Researcher Guidance
Explaining the impact of choice of comparators in economic evaluations	Communication	Content	Technical Analyses Policymaker Q&A Researcher Guidance
Making comparisons with published economic evaluations and explaining why results differ	Communication	Content	Technical Analyses Policymaker Q&A Researcher Guidance
Discussing model inputs, not just the final ICER	Communication	Content	Technical Analyses Policymaker Q&A Researcher Guidance
Communicate strength of the underlying evidence	Communication	Content	Technical Analyses Policymaker Q&A Researcher Guidance
Expressing cost of intervention in terms of numbers needed to treat from clinical evaluation	Communication	Clarity	Technical Analyses Policymaker Q&A Researcher Guidance
Understanding how much treatment for one patient would cost	Communication	Clarity	Technical Analyses Policymaker Q&A Researcher

Individual Economic Study Finding	How Finding Could Be Addressed	Related to Either Content, Context or Clarity*	Potential Tool
			Guidance
Planning tools for policymakers related to budget impact (to distinguish from cost-effectiveness)	Communication	Context	Out of Scope
Graphics customized to different policymakers explaining what the economic evaluation meant and how to apply to their local setting	Communication	Context	Policymaker Q&A
Discuss with and present reports to policymakers	Communication Stakeholder Interactions	Clarity	Out of Scope
Being available to policymakers for questions	Communication Stakeholder Interactions	Clarity	Out of Scope
Multi-stakeholder presentations/discussion of the project	Communication Stakeholder Interactions	Clarity	Out of Scope
Provide verbal explanation before policymakers read reports in detail	Communication Stakeholder Interactions	Clarity	Out of Scope
Communicating clear expectations for analyses submitted by pharmaceutical industry	Communication Stakeholder Interactions	Clarity	Out of Scope
Conduct local economic evaluation only if clinical benefit suggests its useful	Communication	Context	Technical Analyses
Getting feedback from policymakers on economic report templates	Communication	Content Clarity	Technical Analyses
Economic evaluation guidelines	Education Communication	Content Clarity	Tool Repository
Glossaries to explain economic terminology	Education Communication	Content Clarity	Tool Repository
Report for policymakers and stakeholders on how to interpret an ICER in the decision-making context	Education	Context	Tool Repository Learning Modules
Report for policymakers and stakeholders on why a single willingness to pay threshold isn't used	Education	Context	Tool Repository Learning Modules
Report for policymakers on how to use ICER in multi-criteria decision analysis (MCDA) framework	Education	Context	Tool Repository Learning Modules
Training workshops	Education	Content	Learning Modules
Access to economic experts	Education	Content	Learning Modules
Gaining familiarity over time with economic issues and reports	Education	Content	Learning Modules

Note: Methodological issues, health systems pressures, stakeholder interactions and values clarifications were either infrequently identified or considered outside the current scope of the toolkit and potential tool development.

* These themes (content, context, clarity) were identified in interviews with Canadian decision-makers and an attempt was made to see if they also applied to findings from previously conducted systematic reviews and key informant interviews with HTA organizations. **Content** referred to any technical issues or quality issues that could be addressed in tools; **Clarity** referred to issues related to formatting of information, explaining information or targeting to end-users;

Context referred to issues related to identifying relevant local and real-world information for the end-user or the key messages to emphasize for the end-user.

Supplementary Table 2. Identification of Potential Tools Based on Individual NMA Study Findings

Individual NMA Study Finding	How Finding Could Be Addressed	Related to Either Content, Context or Clarity*	Potential Tool
NMA COMMUNICATIONS SR (See Sullivan et al., 2014b for More Details)			
NMA Communication Approaches			
Multiple formats used: tables, figures and text	Communication	Clarity	Technical Analyses Researcher Guidance
Supplementary materials option permits sufficient information	Communication	Clarity	Technical Analyses Researcher Guidance
Different audiences prefer different formats or info	Communication	Clarity	Policymaker Q&A Researcher Guidance
Graphics can have different purposes in NMAs	Communication	Content Clarity	Technical Analyses
Graphical design features can enhance readability	Communication	Clarity	Policymaker Q&A Technical Analyses Researcher Guidance
Evidence network formats	Communication	Content Clarity	Policymaker Q&A Technical Analyses
Trial characteristic formats	Communication	Content Clarity	Policymaker Q&A Technical Analyses
Critical appraisal formats	Communication	Content Clarity	Policymaker Q&A Technical Analyses
Presenting key assumptions, e.g. heterogeneity and inconsistency	Communication	Content Clarity	Policymaker Q&A Technical Analyses
Effect estimate formats	Communication	Content Clarity	Policymaker Q&A Technical Analyses
Ranking formats	Communication Education	Content Clarity	Policymaker Q&A Technical Analyses Learning Modules
Uncertainty formats	Communication	Content Clarity	Policymaker Q&A Technical Analyses
NMA Knowledge			
Inadequate use of methods in some NMAs	Education Accessibility	Content	Learning Modules Tool Repository
Underlying SR or MA principles not applied	Education	Content	Learning Modules
NMAs sometimes misinterpreted	Education Communication	Content	Learning Modules Policymaker Q&A
Statisticians certain of NMAs value	Education	Content	Learning Modules
NMA terminology not always used clearly	Communication Education	Content	Learning Modules Tool Repository
NMA reporting/conduct guidance only adequate for half statisticians	Education Accessibility	Content	Learning Modules Tool Repository
Include network diagram	Communication	Content	Technical Analyses Policymaker Q&A
NMA knowledge varied among researchers	Education	Content	Learning Modules
Researchers would consider indirect evidence, at least sometimes	Education	Content	Learning Modules
NMA Barriers			
Volume of data and studies	Communication	Content	Technical Analyses
Challenges interpreting NMAs	Communication	Content	Policymaker Q&A

Individual NMA Study Finding	How Finding Could Be Addressed	Related to Either Content, Context or Clarity*	Potential Tool
	Education		Learning Modules
Challenges presenting NMAs	Communication	Clarity	Technical Analyses Policymaker Q&A
Evolving methods	Education Accessibility	Content	Learning Modules Tool Repository
Lack of expertise	Education	Content	Learning Modules
Critical appraisal challenges	Education Accessibility	Content	Learning Modules Tool Repository
Lack of transparency	Communication	Clarity	Technical Analyses Policymaker Q&A Researcher Guidance
Inadequate conduct or acceptance of NMAs	Education Accessibility	Content	Learning Modules Tool Repository
Complexity of NMAs	Education Communication	Content Clarity	Learning Modules Technical Analyses Policymaker Q&A
Lack of reporting guidance	Education Accessibility	Content Clarity	Learning Modules Tool Repository
Time and resources required	Systems Pressures	NA	NA
Evidence base inadequate	Communication	Content	Technical Analyses Policymaker Q&A
Inconsistent language	Communication	Clarity	Researcher Guidance Technical Analyses Policymaker Q&A
NMA Facilitators			
Able to assess assumptions	Communication Education	Content	Policymaker Q&A Technical Analyses Learning Modules
Methodological strengths of NMA	Communication Education	Content	Policymaker Q&A Technical Analyses Learning Modules
Having reporting standards	Communication Accessibility	Content Clarity	Technical Analyses Tool Repository
Able to understand available evidence	Communication	Content	Technical Analyses Policymaker Q&A
Knowing rankings and their uncertainty	Communication Education	Content	Technical Analyses Policymaker Q&A Learning Modules
Having critical appraisal standards	Education Accessibility	Content	Technical Analyses Policymaker Q&A Learning Modules Tool Repository
Having guidance and tools for presenting NMAs	Communication Accessibility	Clarity	Technical Analyses Policymaker Q&A Tool Repository
Increased transparency of NMAs	Communication	Clarity	Technical Analyses Policymaker Q&A
Acceptance of NMA methods	Education	Content	Learning Modules

Individual NMA Study Finding	How Finding Could Be Addressed	Related to Either Content, Context or Clarity*	Potential Tool
More NMA training	Education	Content	Learning Modules
Better evidence base	Communication	Content	Technical Analyses Policymaker Q&A
NMA GUIDELINES SR (See Sullivan et al., 2014d for More Details)			
Recommendations on How to Present Information			
Graphics for presenting evidence networks	Communication	Content	Technical Analyses Policymaker Q&A
Flow diagrams of included/excluded studies	Communication	Content	Technical Analyses Policymaker Q&A
Tables or text discussion of trial characteristics	Communication	Content	Technical Analyses Policymaker Q&A
Methodological assumptions presented by text descriptions, graphics or numerical estimates	Communication	Content	Technical Analyses Policymaker Q&A
Present sensitivity analyses excluding trials	Communication	Content	Technical Analyses Policymaker Q&A
Results presented in tables or forest plots	Communication	Content	Technical Analyses
Many choices for presenting results (e.g. different comparisons, outcome measures, direct/indirect/mixed evidence)	Communication	Content	Technical Analyses Policymaker Q&A
Uncertainty through confidence/credible intervals	Communication	Content	Technical Analyses Policymaker Q&A
Rankings presented in tables or through graphics	Communication	Content	Technical Analyses
Rankings challenging to present without misinterpretation	Communication	Content	Technical Analyses Policymaker Q&A
Implications of findings related to: assumptions, validity of results, expectations vs. existing evidence, clinical or biological rationale, relevance to real-world clinical and policy decisions, extent of bias and impact of bias on direction of conclusions	Communication	Content	Technical Analyses Policymaker Q&A
Advice on Presenting to End-Users			
Converting outcomes to measures decision-makers might prefer (RR, ARR, NNT)	Communication	Content	Technical Analyses Policymaker Q&A
INTERVIEW RESULTS (See Sullivan et al., 2014e for More Details)			
Content Elements Valued by Policymakers			
Knowing when/if appropriate to conduct	Communication	Context	Policymaker Q&A
Knowing what was compared and if all appropriate comparators included	Communication	Context	Policymaker Q&A
Knowing level of evidence supporting comparators and historical context of therapies	Communication	Context	Policymaker Q&A

Individual NMA Study Finding	How Finding Could Be Addressed	Related to Either Content, Context or Clarity*	Potential Tool
What studies were included/excluded	Communication	Content	Technical Analyses Policymaker Q&A
Why specific studies were excluded	Communication	Context	Policymaker Q&A
If all appropriate trials were captured/included	Communication	Context	Policymaker Q&A
Critically evaluating if all assumptions were met	Communication	Content	Technical Analyses Policymaker Q&A
Heterogeneity of trials and if comparable, e.g. time when trials conducted, patient populations, other confounders	Communication	Content	Technical Analyses Policymaker Q&A
How/if data are applicable to real-world context, e.g. dosing, populations, treatment sequencing	Communication	Context	Policymaker Q&A
If all influential factors incorporated	Communication	Context	Policymaker Q&A
Sensitivity analyses for uncertain factors	Communication	Content	Technical Analyses Policymaker Q&A
Threats to internal/external validity – quality checklists	Communication	Content	Technical Analyses Policymaker Q&A
If conclusions are firm or hypothesis-generating	Communication	Context	Policymaker Q&A
The power of the NMA to show a difference	Communication	Content	Technical Analyses Policymaker Q&A
Underlying values	Communication	Context	Policymaker Q&A
General methodology descriptions (vs. detailed explanations)	Communication	Content	Policymaker Q&A
Clarity on what the reference drug is	Communication	Content	Policymaker Q&A
Understanding rankings	Communication	Content	Policymaker Q&A
Understanding credible intervals	Communication	Content	Policymaker Q&A
Understanding weightings	Communication	Content	Policymaker Q&A
NMA Barriers			
Correct comparators not included	Communication	Context	Policymaker Q&A
Inconsistent approaches for different drugs	Communication	Context Clarity	Researcher Guidance Policymaker Q&A
Knowing how to interpret NMAs	Communication Education	Content	Policymaker Q&A Learning Modules
Finding relevant info in long reports	Communication	Clarity	Researcher Guidance Policymaker Q&A
Uncertainty in analysis	Communication	Content	Technical Analyses Policymaker Q&A
Poor quality of included studies	Communication	Content	Technical Analyses Policymaker Q&A
Time differences in when trials conducted	Communication	Content	Technical Analyses Policymaker Q&A
Heterogeneity of studies	Communication	Content	Technical Analyses Policymaker Q&A
Doesn't answer the question	Communication	Context	Policymaker Q&A
Doesn't include relevant real-world issues	Communication	Context	Policymaker Q&A
Not relevant to local factors/real-world	Communication	Context	Policymaker Q&A

Individual NMA Study Finding	How Finding Could Be Addressed	Related to Either Content, Context or Clarity*	Potential Tool
context			
Terminology hard to understand	Communication Education	Content	Policymaker Q&A Learning Modules
NMA Facilitators			
Checklists/quality assessment tools	Communication Accessibility	Content	Tool Repository
Understanding real world context of NMA	Communication	Context	Researcher Guidance Policymaker Q&A
Methods guideline for NMA	Communication Accessibility	Content	Tool Repository
Clear, concise critical appraisal	Communication	Clarity Content	Researcher Guidance Technical Analyses Policymaker Q&A
Transparent analysis/report	Communication	Clarity	Technical Analyses
Understanding clinical context	Communication	Context	Researcher Guidance Policymaker Q&A
When done thoughtfully	Communication	Content	Technical Analyses
Done with methodological rigour	Communication	Content	Technical Analyses
Can detect/validate assay sensitivity of NMA	Communication	Content	Technical Analyses
Clear key messages	Communication	Clarity	Researcher Guidance Policymaker Q&A
Details available	Communication	Clarity	Researcher Guidance Technical Analyses
Limitations identified	Communication	Content	Technical Analyses Policymaker Q&A
Comparability of studies considered	Communication	Content	Technical Analyses Policymaker Q&A
Quality of studies considered	Communication	Content	Technical Analyses Policymaker Q&A
Assumptions explicit	Communication	Content	Technical Analyses Policymaker Q&A
Assumptions justified	Communication	Content	Technical Analyses Policymaker Q&A
Identifying trials even if not in statistical analysis	Communication	Content	Technical Analyses Policymaker Q&A
Concrete information, not theoretical	Communication	Content Clarity	Researcher Guidance Technical Analyses Policymaker Q&A
International Approaches			
Summary of existing NMAs with key messages, strengths and weaknesses	Communication	Clarity	Policymaker Q&A
Summary paragraph on strengths and weaknesses of NMA	Communication	Clarity	Policymaker Q&A
Bottom-line key messages and recommendations	Communication	Clarity	Policymaker Q&A
Completed indirect comparison checklists	Communication	Content	Technical Analyses
Figures explaining what evidence is available	Communication	Content	Technical Analyses Policymaker Q&A
Explain treatment outcomes, similar to	Communication	Content	Technical Analyses

Individual NMA Study Finding	How Finding Could Be Addressed	Related to Either Content, Context or Clarity*	Potential Tool
how done for traditional meta-analyses		Clarity	Policymaker Q&A
Provide cautionary statements on using results based on indirect evidence	Communication	Content	Technical Analyses Policymaker Q&A
Identifying local champions	Stakeholder Interactions	NA	Out-of-Scope
Engaging opinion leaders	Stakeholder Interactions	NA	Out-of-Scope
Guidelines or sections of guidelines on NMAs	Communication Education	Clarity Content	Tool Repository
Detailed methods support documents	Education	Content	Tool Repository
Recommend NMA literature to policymakers for better understanding	Education	Content	Tool Repository
Policymakers have access to statisticians or NMA experts	Education	Content	Learning Modules
Internal researcher training in NMAs	Education	Content	Learning Modules

Note: Health systems pressures and stakeholder interactions were either infrequently identified or considered outside the current scope of the toolkit and potential tool development.

* These themes (content, context, clarity) were identified in interviews with Canadian decision-makers and an attempt was made to see if they also applied to findings from previously conducted systematic reviews and key informant interviews with HTA organizations. **Content** referred to any technical issues or quality issues that could be addressed in tools; **Clarity** referred to issues related to formatting of information, explaining information or targeting to end-users; **Context** referred to issues related to identifying relevant local and real-world information for the end-user or the key messages to emphasize for the end-user.

Supplementary Table 3. Detailed Descriptions of Potential Tools

Detailed Descriptions of Potential Tools
Writing Guidance for Researchers
• *Key Criteria*: o A short two-page document. This will increase feasibility of its implementation and use in practice by researchers o Approximately five overarching principles will be emphasized to keep information concise and more likely to be remembered and used in practice o Each principle will be specifically related to one of the following: content, context or clarity. Each of these three areas will be addressed at least once in the principles that are identified. o Examples will be provided that relate the principle to specific situations in an economic evaluation or NMA. o The example will explain for a researcher audience how to identify technical issues and relate them to policy relevant issues. o Formatting approaches will be applied to the document to enhance clarity and readability for researchers (e.g. heading, spacing, key words emphasized, hyperlinks and layering of examples, etc.) o Targeted to researchers submitting manuscripts for journal publication but could be adapted or applicable to different researcher projects, e.g. analysis conducted for a health technology assessment organization o Tool will be non-specific, i.e., guidance/principles would not change with each analysis. Examples could be updated over time as needed or, if implemented by different journals or organizations, more relevant examples created. • *Core Economic Elements*: Will address how to write for policymakers: (1) About content issues such as uncertainty, variability, model assumptions, evidence gaps; (2) About contextual factors such as drug pricing, budget impact and resource constraints, variations in populations and comparators, ethical concerns and other implementation or real-world factors that policymakers will need to consider; and (3) With approaches that enhance clarity such as recognizing economic knowledge level of audience, providing appropriate levels of detail, creating documents with multiple layers (e.g. 1:3:25 approach; supplementary materials) and framing issues for policymakers. • *Core NMA Elements:* Will address how to write for policymakers: (1) About content issues such as explaining the available evidence, key assumptions, NMA limitations, and when its appropriate to conduct an NMA; (2) About contextual factors such as variations in dosing, availability of comparators, impact of treatment sequencing and other implementation or real-world factors that policymakers will need to consider, including how NMA methodologies are evolving; and (3) With approaches that enhance clarity such as recognizing NMA knowledge level of audience, providing appropriate levels of detail, creating documents with multiple layers (e.g. 1:3:25 approach; supplementary materials) and framing issues for policymakers.
Question and Answer (Q&A) Document for Policymakers
• *Key Criteria*: o Each question will address an issue related to either the content or context of the analysis. Clarity will be provided in the format of the tool. o Questions will be related to the overarching principles established in the Guidance for Researchers tool. o Generic questions would be provided in the template but in some cases, may want to adapt questions to unique economic evaluations. Answers would be unique to each economic evaluation. o Primary audience for the tool is policymakers but researchers may require some guidance on completing the Q&A template and on identifying standing issues that are relevant to the policy context. o Intended to be completed by individuals with technical proficiency, i.e. requires economic or NMA expertise and some judgment/interpretation in creating answers to questions.

- Answers to questions should integrate clinical knowledge and interpretation as it relates to assumptions and use of cost-effectiveness estimates.
- Maximum of 3 pages with approximately 10 questions when complete so that it is feasible to implement (i.e. not a lot of time needed for researchers to complete) and for policymakers to use (i.e. not a lot of time needed for policymakers to read).

- *Core Economic Elements*: Questions will address or identify:
  - Content issues such as uncertainty, robustness, model assumptions, evidence gaps, model validity.
  - Impact of sponsor bias on the economic evaluation.
  - Key drivers of the ICER and impact of variations in these inputs.
  - Factors that could change magnitude, direction, precision of economic estimates.
  - Sensitivity analyses and subgroup analyses of most policy-relevant issues.
  - Contextual issues that may influence cost-effectiveness at a local level such as: comparators and populations.
  - General relationships to opportunity costs and local budgetary constraints.
  - Identifying and labeling issues policymakers may have to address upon implementation and that may have an economic component: companion testing, adverse event monitoring, treatment sequencing, indication creep.
  - Information that is useful in price negotiations.
  - Identifying and labeling (but not exploring any further) ethical factors that could influence stakeholder perspectives on economic factors: uncommon diseases, unmet needs, patient values.
- *Core NMA Elements:* Questions will address or identify:
  - Content issues such as assumptions, rankings, critical appraisal, describing available evidence, clarity on characteristics of individual trials and results, identifying effect modifiers.
  - Clarity on if the evidence in the NMA is the best available or what could be reasonably generated to improve NMA
  - Providing rationale for when it is or isn't appropriate to conduct an NMA.
  - Contextual issues such as use in different patient populations, utilization and availability of comparators.

Technical Analyses

- *Key Criteria:*
  - Will build on current approaches to presenting technical analyses, e.g. report templates identified in guidelines, journal manuscript requirements.
  - Will provide clarity by incorporating formatting approaches that emphasize key issues, enhance readability and provide information in layers.
- *Core Economic Elements*: Content to include in technical analyses will be identified based on best practices identified in the review of economic guidelines. End-users' preferences for economic information will be incorporated, while maintaining technical focus of the analysis document.
- *Core NMA Elements:* Content to include in technical analyses will be identified based on best practices identified in the review of NMA guidelines. End-users' preferences for NMA information will be incorporated, while maintaining technical focus of the analysis document.

Tool Repository

- *Key Criteria*:
  - Could links to larger existing repository or resource that policymakers already use
  - Partner or collaborate with groups or organizations that this would be relevant to their purpose (e.g. HTA organizations, researcher, economist or policymaker collaborations)
  - Should be easily accessible, e.g. publicly available website
  - Repository is sufficiently flexible to update and easily add new materials
  - Tools provided in repository are organized such that its easy to find
  - Option to provide commentaries on tools (e.g. who can use, when to use, how to use, how to adapt, when its relevant)

- Mechanisms in place to maintain and sustain database, e.g. identifying potential new materials, ensuring materials are relevant to policymakers, minimal human resource commitment

- *Core Economic Elements***:** Will include quality assessment checklists, transferability checklists, checklists specifically for decision-makers, relevant tutorials or learning modules, relevant review articles or original studies relating to use of economic evaluations in policy settings
- *Core NMA Elements:* Similar content as economic repository but there may not be as many checklist tools available for NMAs due to their novelty and the repository may need a greater focus on NMA tutorials and learning modules.

Learning Modules

- *Key Criteria***:**
  - Focuses on how analyses are concretely applied to policy decisions vs. focusing on theoretical technical concepts.
  - Provides insights on decisions and assessments made at each step in an analysis to understand how the final product is developed and how researchers think along the way.
  - Use a real-world case study to walk audience through an analysis.
  - Considers how other modules and training workshops may have presented analyses to build on best practices and not duplicate approaches.
- *Core Economic Elements***:**
  - Will incorporate content important to economists when presenting technical analyses.
  - Provides policymakers with approaches and best practices when adapting technical analyses to their local context.
- *Core NMA Elements:*
  - Will incorporate content important to researchers when presenting technical analyses.
  - Modules for policymakers will help policymakers identify local and real-world factors that may alter how NMA results are implemented (e.g. dosing, populations) and any best practices that can help.
  - Modules for researchers may have more technical focus.

Lay Peer Review Process

- *Key Criteria***:**
  - Could be applicable to different materials that are produced related to analyses, e.g. manuscripts, policymaker Q&As, technical analyses, etc.
  - Multiple reviewers required so that individuals don't become too familiar with issues.
  - Public calls for reviewers may increase stakeholder engagement and awareness of analyses.
  - Provides outline of how process could integrate into current review systems (e.g. rapid response comments, journal peer review, HTA stakeholder engagement approaches).
  - Provides guidance for reviewers on how to review when no technical background and issues to identify.
  - Builds on or integrates with existing peer review systems (e.g. patient, public, citizen reviews).
- *Core Economic Elements*: Will identify economic concepts that are unclear, potential plain language that could be used, real-world applications of economic results that could be addressed or emphasized for end-users.
- *Core NMA Elements:* Will identify NMA concepts that are unclear, potential plain language that could be used, real-world applications of NMA results that could be addressed or emphasized for end-users.

Supplementary Table 4. Assessment of Selection Criteria and Rationale for Ranking Toolkit Components.

Assessment of Selection Criteria and Rationale for Toolkit Selection and Development
Guidance for Researchers - Economic
Economic guidance for researchers was ranked high on all five selection criteria: feasibility, appropriateness, potential value, importance and novelty. • The **feasibility** of developing guidance for economists was considered high because it would require minimal resources, could be developed in a timely way, and could easily incorporate existing evidence that was identified through this research and would provide a way to share these findings more broadly with economists. • The **appropriateness to the drug funding context** was considered high because economic reports are frequently written in the drug funding context and available economists interested and appropriately experienced in this work are in short supply. Therefore, increasing efforts to guide economists in how they write and present their work to policymakers would be appropriate. This could be used to foster capacity building among economists and help them better understand policymakers needs in the drug funding context. • **Potential value** was considered high because guide can assist economists in increasing the clarity of their economic evaluations – this could increase impact of economic evaluations, which has the potential to reduce costs while maximizing health benefits across the population. In the current resource-constrained environment, this has great potential value. • **Importance** was considered high because a writing guide has the potential to reduce important inconsistencies in how economists write and think about how economic evaluations are use in a policy environment to inform decisions. Having economists more clearly explain economic evaluations (and the uncertainty that exists within them) would reduce policymakers' uncertainty when interpreting and using economic evaluations. • **Novelty** of a writing guide was considered high because no such tool targeting economists was identified in literature searches, limited guidance was available in guidelines for economists on how to write for policymakers or other non-technical and would not duplicate other efforts.
Guidance for Researchers - NMA
NMA writing guidance for researchers was ranked high on four selection criteria (feasibility, appropriateness, importance and novelty) and medium on potential value. • The **feasibility** of developing guidance for researchers was considered high because it would require minimal resources, could be developed in a timely way, and could easily incorporate existing evidence that was identified through this research and would provide a way to share these findings more broadly with researchers. • The **appropriateness to the drug funding context** was considered high because NMAs are frequently conducted to inform the drug funding policies and decisions and NMA experts who are interested and appropriately experienced in this work are in short supply. Therefore, increasing efforts to guide researchers in how they write and present their work to policymakers would be appropriate. This could be used to foster capacity building among researchers and help them better understand policymakers needs in the drug funding context. • **Potential value** was considered medium because NMA is a relatively novel approach that is not frequently used, as yet. However, in time or in specific situations, it could have greater value because it could lead to researchers better understanding how to present NMAs to policymakers. This could increase the impact of NMAs, which has the potential to maximize health benefits across the population. • **Importance** was considered high because a writing guide has the potential to reduce important inconsistencies in how researchers write and think about how NMAs are use in a policy environment to inform decisions. Having researchers more clearly explain NMAs

(especially when methods are new and unfamiliar) would reduce policymakers' uncertainty when interpreting and using NMAs
- **Novelty** of a writing guide was considered high because no such tool targeting researchers was identified in literature searches, limited guidance was available in NMA guidelines on how researchers should write NMAs for policymakers or other non-technical and would not duplicate other efforts.

Q&A Document for Policymakers - Economic

The economic policymaker Q&A document was ranked high on four selection criteria (appropriateness, potential value, importance and novelty) and medium on feasibility.

- The **feasibility** of developing a policymaker Q&A document on economic evaluations was considered medium because identifying a target end-user/policymaker audience for the development of the tool was considered essential and some time may required to do this. However, once this is done, it could be developed in a timely way (short document) that could easily incorporate existing research (economic information already being generated within technical report) and integrates policymakers' preferences and needs.
- The **appropriateness to the drug funding context** was considered high because policymakers frequently have questions on the economic evaluation component of HTAs and must frequently determine how to adapt an economic evaluation to their local context before making and implementing a drug funding decision.
- **Potential value** was considered high because it could lead to policymakers having a better understanding of how to optimize the use of economic evaluations and which factors matter most for policy decisions. This could increase the impact of economic evaluations, which has the potential to reduce costs while maximizing health benefits across the population. In the current resource-constrained environment, this has high potential.
- **Importance** was considered high because interpreting economic evaluations is an important source of uncertainty for decision-makers. As well, Canadian decision-makers and individuals from HTA organizations around the world considered that providing clearer information on economic evaluations was an important issue they would like to see addressed. This can also lead to reduce inconsistencies in applying the results of economic evaluations to policy decisions and ensuring that variations occur when it makes sense for them to vary.
- **Novelty** of the Q&A document for policymakers was considered high as very few approaches to communicating economic evaluations to policymakers were identified in literature searches or through an environmental scan of HTA organizations around the world. Therefore, this would not duplicate existing tools.

Q&A Document for Policymakers - NMA

The NMA policymaker Q&A document was ranked high on three selection criteria (appropriateness, importance and novelty) and medium on two criteria (feasibility, potential value).

- The **feasibility** of developing a policymaker Q&A document on NMAs was considered medium because identifying a target end-user/policymaker audience for the development of this tool was considered essential and some time may required to do this. However, once this is done, it could be developed in a timely way (short document) that could easily incorporate existing research (economic information already being generated within technical report) and integrates policymakers' preferences and needs.
- The **appropriateness to the drug funding context** was considered high because head-to-head trials are frequently unavailable and policymakers must determine when and how to use NMAs to inform cost-effectiveness analyses and therapeutic or drug class reviews.
- **Potential value** was considered medium because NMA is a relatively novel approach that is not frequently used, as yet. However, in time or in specific situations, it could have greater value because it could lead to policymakers having a better understanding of how to use NMAs to support policy decisions. This could increase the impact of NMAs, which has the potential to maximize health benefits across the population.
- **Importance** was considered high because NMAs are often challenging for researchers

and others to interpret, which policymakers and others would like to address. This can also lead to reduced inconsistency in applying the NMAs to policy decisions and ensuring that variations occur when it makes sense for them to vary.
- **Novelty** of the Q&A document for policymakers was considered high as very few approaches to communicating NMAs to policymakers were identified in literature searches or through an environmental scan of HTA organizations around the world. Therefore, this would not duplicate existing tools.

Technical Analysis Presentations - Economic

The economic technical analysis was ranked high on one selection criteria (appropriateness), medium on three criteria (feasibility, potential value, importance) and low on one criterion (novelty).

- **Feasibility** was considered medium because additional methodological and technical review would be required before approaches to presenting technical analyses could be suggested. However, numerous manuscript guidelines and report templates exist so updating these approaches would be more feasible than developing new approaches to presenting technical analyses.
- The **appropriateness to the drug funding context** was considered high because policymakers frequently have questions on the economic evaluation component of HTAs based on presentations of technical analyses and must frequently determine how to adapt an economic evaluation to their local context before making and implementing a drug funding decision.
- **Potential Value** was considered medium because approaches to presenting technical analyses exist (e.g. report templates in guidelines, manuscript submission guidelines) and revisions may only offer small incremental improvements on actual policy decisions.
- **Importance** was considered medium because more effort on optimizing technical presentations would not likely have extensive uptake among multiple stakeholders and they would likely require more training and education before benefiting from the revised presentation approaches
- **Novelty and Lack of Duplication** was considered low because approaches to accurately presenting technical analyses exist.

Technical Analysis Presentations - NMA

NMA technical analyses were ranked high on two selection criteria (appropriateness, novelty), medium on two criteria (potential value, importance) and low on feasibility.

- **Feasibility** was considered low because methodological approaches are still evolving and more certainty on the appropriateness of methods would be required before approaches to presenting technical analyses could be suggested. Also, few manuscript guidelines and report templates exist so more work would be required than updating existing approaches to presenting technical analyses.
- **Appropriateness to Drug Funding Context** was considered high because head-to-head trials are frequently unavailable and policymakers must use presentations of technical analyses to inform cost-effectiveness analyses and therapeutic or drug class reviews.
- **Potential value** was considered medium because NMA is a relatively novel approach that is not frequently used, as yet. However, in time or in specific situations, it could have greater value because few standard approaches to presenting technical analyses of NMAs exist and it could lead to policymakers having a better understanding of how to use NMAs to support policy decisions. This could increase their impact and ability to maximize health outcomes in the health system.
- **Importance** was considered medium because more effort on optimizing technical presentations would not likely have extensive uptake with multiple stakeholders and they would likely require more training and education before benefiting from the revised presentation approaches
- **Novelty and Lack of Duplication** was considered high because no standard approaches for presenting technical analyses of were identified and methodological NMA approaches are still evolving.

Tool Repository - Economic

The economic tool repository was ranked high on potential value and medium on four criteria (feasibility, appropriateness, importance and novelty).

- **Feasibility** was considered medium because time would be required to develop an accessible and user-friendly platform for the repository as well as build partnerships that could host the repository.
- **Appropriateness to Drug Funding Context** was considered medium because, although, it would be an appropriate tool in the drug funding context, other tools being developed may be considered more appropriate.
- **Potential Value** was considered high because access to tools it could lead to policymakers having a better understanding of how to optimize the use of economic evaluations and which factors matter most for policy decisions. This could increase the impact of economic evaluations, which has the potential to reduce costs while maximizing health benefits across the population. In the current resource-constrained environment, this has high potential.
- **Importance** was considered medium because there was support from stakeholders for access to tools that would be housed in the repository.
- **Novelty** was considered medium because some tool repositories may exist but for policymakers but may not be well-publicized. Consideration would be needed to ensure efforts are not duplicated.

Tool Repository - NMA

The NMA Tool Repository was ranked high on two criteria (potential value, novelty) and medium on three criteria (feasibility, appropriateness, importance).

- **Feasibility** was considered medium because time would be required to develop an accessible and user-friendly platform for the repository as well as build partnerships that could host the repository.
- **Appropriateness to Drug Funding Context** was considered medium because, although, it would be an appropriate tool in the drug funding context, other tools being developed may be considered more appropriate.
- **Potential Value** was considered high because NMAs are often challenging for researchers and others to interpret. Providing access to tools and educational materials could increase understanding of NMAs and their use in supporting policy decisions. This could increase their impact and ability to maximize health outcomes in the population.
- **Importance** was considered medium because there was support from stakeholders for access to tools that would be housed in the repository.
- **Novelty and Lack of Duplication** was considered high because no NMA repositories for policymakers likely exist and NMA tools that would populate the repository are only starting to be developed

Learning Modules - Economic

The economic learning module was ranked high on two selection criteria (potential value, importance), medium on two selection criteria (feasibility, appropriateness) and low on novelty.

- **Feasibility** was considered medium because developing the modules would be time and resource intensive.
- **Appropriateness to Drug Funding Context** was considered medium because although it is an appropriate tool, some other toolkit components may be more appropriate and applicable to the drug-funding context.
- **Potential Value** was considered high because if learning modules have the right focus, they could lead policymakers to have a better understanding of how to optimize the use of economic evaluations when making policy decisions. This could increase the impact of economic evaluations, which has the potential to reduce costs while maximizing health benefits across the population. In the current resource-constrained environment, this has high potential.
- **Importance** was considered medium because there was some support from stakeholders in better understanding how to apply economic evaluations to local policy decisions. However, many policymakers considered that a broad understanding of

economics was sufficient for the work that was required in a policy setting.
- **Novelty and Lack of Duplication** was considered low because modules and workshops likely exist but may not be focused on applied aspects of economic evaluations, in which policymakers are most interested

Learning Modules - NMA

NMA Learning Modules was ranked high on potential value and medium on four selection criteria (feasibility, appropriateness, importance and novelty).
- **Feasibility** was considered medium because developing the modules would be time and resource intensive and because determining content would be more challenging as methods are still evolving and versions would be required for both policymakers and researchers.
- **Appropriateness to Drug Funding Context** was considered medium because although it is an appropriate tool, some other toolkit components may be more appropriate and applicable to the drug-funding context.
- **Potential Value** was considered high because NMAs are often challenging for researchers and others to interpret. Providing learning modules with the right focus for policymakers (and versions for researchers) could increase understanding of NMAs and their use in supporting policy decisions. This could increase their impact and ability to maximize health outcomes in the population.
- **Importance** was considered medium because there was some support from stakeholders in better understanding how to apply economic evaluations to local policy decisions. However, many policymakers considered that a broad understanding of economics was sufficient for the work that was required in a policy setting.
- **Novelty and Lack of Duplication** was considered medium because some modules and workshops may exist but may need to be more targeted to a policymaking setting.

Lay Peer Review Process - Economic

The economic lay peer review process was ranked high on two criteria (appropriateness, novelty) medium on three selection criteria (feasibility, potential value, importance) and low on feasibility.
- **Feasibility** was considered low because of the partnerships, human resources, training and process changes that would be required to implement a lay peer review process for economic evaluations.
- **Appropriateness to Drug Funding Context** was considered high because it could support and address challenges with communicating and presenting economic evaluations to policymakers who use them to inform policy decisions.
- **Potential Value** was considered medium because its uncertain what the potential of the lay peer-review process could be to impact change as plain-language approaches have previously been challenging to implement for economic evaluations.
- **Importance** was considered medium because a lay peer-review process may decrease variations across technical analyses in how economists present information to end-users.
- **Novelty and Lack of Duplication** was considered high because few processes or organizations (e.g. journals, collaborations, etc.) provide a lay peer-review of economic evaluations.

Lay Peer Review Process - NMA

NMA lay peer review process was ranked high on three selection criteria (appropriateness, potential value and novelty) and medium on importance and low on feasibility.
- **Feasibility** was considered low because of the partnerships, human resources, training and process changes that would be required to implement a lay peer review process for NMAs.
- **Appropriateness to Drug Funding Context** was considered high because it could support and address challenges with communicating and presenting NMAs to policymakers who use NMAs to inform policy decisions.
- **Potential Value** was considered high because NMAs are often challenging for researchers and others to interpret. Getting early feedback from lay peer reviewers could provide suggestions for increasing understanding of NMAs. This could increase their use in policy decisions, impact and ability to maximize health outcomes in the population.

- **Importance** was considered medium because a lay peer-review process may decrease variations across technical analyses in how researchers present NMAs to end-users.
- **Novelty and Lack of Duplication** was considered high because few processes or organizations (e.g. journals, collaborations, etc.) provide a lay peer-review of NMAs

SUPPLEMENTARY MATERIAL 2

Economic Guidance for Researchers Tool Illustrated with Examples from an Economic Evaluation of Osteoporosis Therapies.

This tool is presented in two parts. Part 1 explains principles economists could follow when presenting an economic evaluation to policymakers. Part 2 illustrates how these principles can be applied using examples from an economic evaluation of osteoporosis therapies (Coyle et al., 2014). In the final version of the tool, moving between these two sections could be facilitated by hyperlinks within the document.

PART 1. Economic Guidance for Researchers

HOW CAN I PRESENT AN ECONOMIC EVALUATION SO THAT IT IS MOST USEFUL TO POLICYMAKERS? Guidance for Economists.

This document provides an overview of some evidence-based principles that have been suggested to help policymakers better apply the results of an economic evaluation to a policy decision. Keeping these principles in mind when writing an economic evaluation could enhance the usefulness of your work for policymakers. Examples of how these principles can be applied to an economic evaluation are provided based on a published economic evaluation of osteoporosis therapies (Coyle et al., 2014), however, you may be able to identify other examples of how these principles could be applied. How explanations for policymakers are incorporated into a written economic evaluation could vary, depending on report or journal requirements (e.g. discussion, supplementary appendices). This guidance should be considered complementary to, but does not replace, reporting guidelines.

HOW CAN I MAKE THE *CONTENT* OF MY ECONOMIC EVALUATION MOST USEFUL TO POLICYMAKERS?

1) **Help Policymakers Understand the Uncertainty in the Results:**
   Policymakers want to understand how "believable" are the results of an economic evaluation. This helps them determine how much confidence to have in the ICER when applying it to a policy decision. Economists often explore different sources of uncertainty (parameter uncertainty, structural uncertainty, variability) through sensitivity analysis. However, explaining why these factors may be relevant to a policy decision, and which ones matter most, is as important as reporting the result itself. This can be done in different ways including:
   - Explaining how much of the estimate is derived from actual data and how much is reliant on extrapolated data.
   - Pointing out the most important assumptions as demonstrated by sensitivity analyses that have the greatest impact on the magnitude and direction of the ICER.
   - Noting what factors do not impact on the ICER when they have been explored in sensitivity analyses

See Example 1A. Uncertainty of NMA Effect Estimates.
See Example 1B. Impact of Adherence Assumptions.
See Example 1C. Extrapolating Fracture Risk and Treatment Benefits Over Time.

2) **Help Policymakers Understand What the <u>Limitations in How the Analysis was Conducted</u> Mean:**
Policymakers want to understand the limitations in an economic analysis so that they know how much they can 'trust' the results of the economic evaluation. This helps them determine how much confidence to have in the ICER when applying it to a policy decision. Although results produced by all economic models have some inherent uncertainty in them, modeling can be done well by following best practices and guideline standards or it can be done poorly. Clearly identifying and explaining these limitations, and the choices that have been made when conducting an analysis, is important. Economists often do this by identifying any potential limitations in their analysis and in any comparable published analyses, which helps policymakers understand which of the analyses are most valid and applicable to their local setting. Explaining which limitations are most important and framing the limitations of an analysis in ways that help policymakers determine how to apply the evidence to the context of their specific policy decision is important. This can be done in different ways, including:
   - Identifying which of the potential limitations are going to have the largest impact on the ICER, not just identifying all potential limitations. This helps policymakers realize that even if a limitation could be corrected, it may not have a meaningful impact on the ICER
   - Identifying when limitations are unavoidable, either because no better evidence or methodology currently exists. This helps policymakers realize that this is still the best available evidence, even though it has some limitations associated with it.
   - Acknowledging the relative importance of different limitations in an analysis. This helps policymakers realize that even if some of the minor limitations were corrected, results would still be overshadowed by major limitations in the analysis.
   - Acknowledging strengths of the analysis that may outweigh its limitations.

   See Example 2A. Comparison with Other Economic Evaluations.
   See Example 2B. **Structural Limitations in the Model**.

HOW CAN I MAKE MY ECONOMIC EVALUATION MORE RELEVANT TO A POLICYMAKERS DECISION-MAKING *CONTEXT*?

3) **Help Policymakers Identify How Evidence Differs from <u>Real-World Settings</u>:**
Policymakers want to know if the effect estimates and modeling assumptions informing the economic evaluation differ from how a patient experiences the disease (i.e. the natural disease history) and therapy (i.e. safety, effectiveness, quality of life) or how clinicians treat the disease in clinical practice (i.e. external validity of modeled populations and treatment regimens). This helps policymakers consider how the analysis applies to real-world patients and how clinical practice will be impacted by policy decisions. Economists can sometimes explore these issues through sensitivity analyses. At other times, the model structure or available clinical data don't permit sensitivity analyses. However, even if sensitivity analyses are not possible, identifying these discrepancies for policymakers, and noting the potential direction and magnitude of impact on the ICER

can provide helpful context to policymakers when applying economic results to a policy decision.

See Example 3A. **Patient Populations Treated in Practice**.
See Example 3B. Treatment Adherence in Practice.
See Example 3C. Side Effects of Osteoporosis Therapies.

4) **Help Policymakers Adapt Analyses to Their <u>Local Decision-Making Context</u>:** Policymakers want to know if the results of an economic analysis are applicable to their local setting. This helps policymakers make evidence-informed policy decisions that are relevant to the patients, clinical practices and existing policies in their local setting. Using one model and one data set, economists may be challenged to provide results for all the different scenarios that local policymakers must consider. However, different approaches can be taken in these cases that are helpful to policymakers:
   - Identifying those factors in the economic model that are likely to change across local settings (e.g. drug pricing, relevant comparators), even if the appropriate model or information is not available to be able to address these factors. This helps policymakers understand when an analysis is or isn't relevant to their local setting.
   - When sensitivity analyses can be conducted on local factors, determining which factors have an important impact on the magnitude and direction of the ICER and focusing on these sensitivity analyses when writing the analysis. This helps policymakers understand how the results of the economic analysis may change when applied to their local setting.
   - Noting that you explored a local factor through sensitivity analyses and that it did not have an important impact on the magnitude or direction of the ICER. This lets policymakers know that the issue was considered, not ignored, and therefore the results of the economic analysis are likely still applicable to their local setting.
   - When local factors cannot be addressed through sensitivity analyses, providing policymakers with a description of next steps or approaches they could take to address these factors locally. This helps policymakers as this may be required to conduct additional analyses locally but may not have access to the required economic expertise.

See Example 4A. Differences in Drug Pricing.
See Example 4B. Differences in Relevant Comparators.

HOW CAN I PROVIDE MORE *CLARITY* WHEN PRESENTING ECONOMIC EVALUATIONS TO POLICYMAKERS?

5) **Help Policymakers Better Understand the ICER By <u>Breaking Down the Analysis</u>:** Policymakers want to deconstruct the results and see the analysis in a disaggregated format. This helps policymakers determine how the analysis may need to be adapted to a local policy decision. It can also increase confidence in the analysis if disaggregated data provides a more transparent analysis that can be reconstructed. Economists can do this by identifying different components of the ICER (e.g. total costs and effects, incremental costs and effects, life-years, quality-adjusted life years and utility values) and providing a general description of the key inputs informing the ICER. Showing that an ICER may be different across settings, subpopulations or over time period of the analysis is another way to disaggregate results and make them most useful to policymakers.

See Example 5A. Disaggregating the ICER.
See Example 5B. ICERs in Different Subpopulations.

6) **Help Policymakers Identify the Key Messages from the Economic Evaluation:** Policymakers want to understand the key issues and bottom-line of the economic analysis as it relates to the policy decision to be made. This helps policymakers frame the details of the analysis and put them into the context of the main conclusion and messages from the analysis. This means not only providing the ICER, but also trying to emphasize the three or four most important ideas from the analysis that you want policymakers to remember when they are trying to make a decision based on the results. Providing this information in the beginning of the write-up and keeping these ideas concise, makes it more likely policymakers will remember them.

See Example 6A. Identifying Key Analysis Issues.
See Example 6B. Policy-Relevant Conclusions.

7) **Provide Policymakers With Clear, Relevant and Understandable Information:** Policymakers want to focus on the information that is most relevant to them, as policymakers trying to make evidence-informed policy decisions. This may differ from information that is important to other audiences, including economists, academics, clinicians, patients or the public. This helps policymakers quickly synthesize the economic analysis and apply the results to their local policy context. Providing a detailed write-up of the economic analyses is important for transparency but when a lot of information is provided, key details can get lost or buried in the write-up. However, providing policymakers with the most relevant information to them, can be done in different ways, including:

- Thinking about the policy questions that are of most interest to policymakers when selecting and presenting results.
- Emphasizing policy-relevant results but providing additional details in subsequent layers (e.g. in supplementary materials). This provides clarity on the most relevant findings but also gives policymakers confidence the analysis has been well-done and allows them to refer to the analysis more in-depth as needed.
- Making links between technical concepts that are often encountered and metrics or ideas with which policymakers are more familiar or explaining why the technical issue is important and how it impacts on the results of a subsequent policy decision.

See Example 7A. Identifying Policy-Relevant Questions.
See Example 7B. Linking Common Technical Ideas to Familiar Concepts.

PART 2. Illustrated Examples from an Economic Evaluation of Osteoporosis Therapies.

1) Help Policymakers Understand the Uncertainty in the Results:

Example 1A. Uncertainty of NMA Effect Estimates. Estimates comparing the effectiveness of different osteoporosis therapies were generated from a network meta-analysis for this economic evaluation and not direct head-to-head trials. Network meta-analyses can produce a range of different results, depending on which trials are included and which analytical approaches are used. When the effectiveness estimates are changed to represent results generated by a different NMA, in which denosumab was more effective in preventing hip fractures, the cost-effectiveness of denosumab becomes more favourable. However, we believe the estimates from our NMA are more accurate and reliable than the results from the other published NMA.

Example 1B. Impact of Adherence Assumptions. One of the proposed advantages of newer and more expensive therapies like zoledronic acid and denosumab is that treatment adherence will be improved (and subsequently effectiveness) because the treatments are administered less frequently, i.e. every six months or a year. For this reason, we did extensive sensitivity analyses on treatment adherence assumptions. However, because the cost differences between treatments are so large, differences in adherence have little impact on the cost-effectiveness results.

Example 1C. Extrapolating Fracture Risk and Treatment Benefits Over Time. Economic models have the advantage of being able to project the costs and effectiveness of therapies over time and not rely only on the length of time data were collected in clinical trials. Trials included in the NMA ranged from 6 months to 3 years in length and this economic model generates cost-effectiveness estimates over a person's lifetime. For post-menopausal women, a 'lifetime' could be an average of approximately 30 years in North America. In practice, treatment duration can vary. In this economic model, the length of treatment is 5 years. Extrapolating trial data from 3 to 5 years may be reasonable, which was confirmed in sensitivity analyses comparing treatment benefit based on actual trial duration with the extrapolated treatment benefit.

2) Help Policymakers Understand What the <u>Limitations in How the Analysis was Conducted</u> Mean:

> **Example 2A. Comparison with Other Economic Evaluations**. This economic evaluation was conducted in accordance with the CADTH Economic Evaluation Guidelines. The most important strength of the analysis is the transparency of the modeling approach. The most important limitation in this economic evaluation is that it is dependent on the quality and availability of data to support clinical assumptions in the model. Two published economic evaluations on this topic were reviewed for strengths and limitations in how the analyses were conducted.

> **Example 2B**. **Structural Limitations in the Model**. It was unclear if the two economic models used in two published economic evaluations were adequately calibrated. Calibration is important when data or specific input parameters are not available and can overestimate the cost-effectiveness of treatments. This model was appropriately calibrated and the ICER of alendronate versus no pharmacotherapy from this model was $10,895 per QALY. However, if the model had not been adequately calibrated, the ICER for treatments would be much lower and in the case of alendronate, treatment would dominate no pharmacotherapy, i.e. be more costly and less effective.

3) Help Policymakers Identify How Evidence Differs from Real-World Settings:

Example 3A. **Patient Populations Treated in Practice**. It has been suggested that more costly osteoporosis therapies (e.g., denosumab and zoledronic acid) should be considered for patients who are intolerant to bisphosphonates, Bisphosphonates, with the exception of zoledronic acid, are generic and less costly drugs. Most of the trials in the NMA used in the economic model excluded patients who were intolerant to bisphosphonates (otherwise they could not have been ethically randomized to treatment with a bisphosphonate or no bisphosphonate). Therefore, the patient population modeled in this economic evaluation does not provide any information on the cost-effectiveness of denosumab or zoledronic acid in patients intolerant to bisphosphonates. It is unclear how much of an issue this is in clinical practice. Therefore, policymakers may need to consider what proportion of patients is likely to be intolerant to bisphosphonates, the clinical need for therapeutic options in this subpopulation and if there is actual value in further research to determine the cost-effectiveness in this specific sub-population. Results comparing only zoledronic acid and denosumab suggest that denosumab is dominated by zoledronic acid, i.e. it is more costly and less effective.

Example 3B. Treatment Adherence in Practice. Data on patient adherence was obtained from published literature and previous analyses with the exception of zoledronic acid, for which no data were available. Assumptions were made about adherence in sensitivity analyses based on best and worst case scenarios. These suggested that the ICUR for zoledronic acid compared with alendronate could be as much as $2, 593, 864 per QALY if there is equal adherence with denosumab and zoledronic acid compared with and ICUR of $1,609,172 per QALY in the primary analysis. Policymakers should consider which of these scenarios are most likely to reflect real-world patient adherence to treatments.

Example 3C. Side Effects of Osteoporosis Therapies. Data on harms associated with osteoporosis therapies were not incorporated into the economic model. However, NMA results suggested that there may be no difference in withdrawals due to adverse events (WDAEs) between therapies, therefore modeled costs would also be similar between therapies. More uncommon and serious adverse events associated with bisphosphonates and denosumab (e.g., osteonecrosis of the jaw, esophageal cancer, atrial fibrillation, atypical fractures, malignancies, suppression of bone turnover) may have costs associated with their treatment that could be modeled. However, the impact of these costs in light of other cost drivers in the model (e.g. drug prices) is not expected to have a large impact on the magnitude or direction of the ICER. Disutilities, representing the lack of preference or aversion people have to fractures, were incorporated into the economic analysis, however, these are negative impacts of the osteoporosis and not side effects of treatment with bisphosphonates or denosumab.

4) Help Policymakers Adapt Analyses to Their Local Decision-Making Context:

Example 4A. Differences in Drug Pricing. Denosumab and zoledronic acid are newer drugs that still have patent protection in most countries while alendronate is a generic drug; generic drug prices are usually much lower than patent protected drugs. However, drug prices may change across local settings and this is a key driver in the model accounting for up to 51% of the total costs in the base case analysis. Scenario analyses explored how the cost-effectiveness of therapies would change, depending on the relative prices of denosumab, zoledronic acid and alendronate. Analyses found that drug prices of denosumab and zoledronic acid would need to be reduced substantially (by 82.6% and 77.3%, respectively) in order for either therapy to be considered cost effective when compared to alendronate.

Example 4B. Differences in Relevant Comparators. Availability of osteoporosis therapies varies with ibandronate not authorized in Canada and etidronate not authorized in the USA for postmenopausal osteoporosis In addition, in some local settings, clinicians may have concerns with adherence to a weekly oral medication such as alendronate and there may be greater utilization of therapies with less frequent administration, e.g. denosumab and zoledronic acid. If a policymaker only wants to fund one of these newer drugs, to provide options for clinicians and patients that prefer less frequent administration the question of interest may be how zoledronic acid and denosumab compare with each other. Focusing on the comparison between zoledronic acid and denosumab, denosumab is dominated by zoledronic acid, i.e. it is more costly and has a similar effect.

5) Help Policymakers Better Understand the ICER By Breaking Down the Analysis:

> **Example 5A. Disaggregating the ICER.** The key effects included in the economic model were fracture risk, mortality, disutilities associated with fractures and treatment efficacy. The key costs included in the economic model were costs associated with fractures including hospitalization, rehabilitation and institutionalization and drug costs. The utility values to inform the QALY estimate were based on values obtained from literature on healthy Canadian women. Based on these inputs, when comparing zoledronic acid with alendronate an ICER of $2,350 million per life year and an ICUR $1,175 million per QALY were calculated. This comparison was selected as a policy-relevant comparison because alendronate is frequently the standard of care and zoledronic acid is an effective therapy that has the advantage of reduced yearly administration.

> **Example 5B. ICERs in Different Subpopulations.** The final ICER is based on an NMA of 34 studies. While all of the studies included a similar population of postmenopausal women, some trials were primary prevention and some were secondary prevention trials. We considered using subgroup data from the NMA on primary and secondary prevention trials. However, the results from many of these subgroup analyses were not statistically significant in the NMA and so economic analyses might yield misleading results on the cost-effectiveness of osteoporosis therapies in a primary or secondary prevention setting. However, an estimate of the costs in these subgroups can be calculated.

6) Help Policymakers Identify the Key Messages from the Economic Evaluation:

Example 6A. Identifying Key Analysis Issues. The key points to remember in this analysis that have an important impact on the economic results:

1) The analysis is based on an NMA and not head-to-head trials but this is currently the best available evidence and head-to-head trials are not expected.
2) Drug prices are based on publicly available listed prices and cost-effectiveness estimates may change if other confidential prices of therapies exist
3) Patients intolerant to bisphosphonates were not modeled in this analysis

Example 6B. Policy-Relevant Conclusions. Important policy relevant-conclusions that can be drawn from this analysis include:

1) Alendronate is cost-effective in almost any scenario ($10 895 per QALY compared with no therapy).
2) Newer treatments such as denosumab and zoledronic acid, which are administered less frequently, are not cost-effective. Both therapies are more costly when compared with alendronate, and, in the case of denosumab, less effective than alendronate.
3) The reductions in drug price that would be required for newer therapies such as zoledronic acid and denosumab to be as cost-effective as alendronate are extremely large (approximately 40%) and may be difficult to obtain while these drugs are under patent protection and alendronate is available generically.
4) Zoledronic acid, which is administered yearly, may be an effective option for patients who cannot receive weekly alendronate, however, the size and clinical importance of this subpopulation is unclear and specific evidence for this subgroup was not available from trials and was not modeled in this analysis.
5) Many of the costs included in this analysis were outside the domain of drug budgets but would be gains received by the health care system as a whole due to the effective use of alendronate (e.g. hospitals, long-term care, rehabilitation)

7) Provide Policymakers With Clear, Relevant and Understandable Information:

Example 7A. Identifying Policy-Relevant Questions. When considering the current landscape of osteoporosis therapies, a number of policy-relevant questions can be identified. Ensuring that the provided results can help to inform some of these questions will increase the relevance of the analysis:

- What is the most cost-effective osteoporosis therapy?
- What is the cost-effectiveness of newer therapies with less frequent administration (e.g. zoledronic acid, denosumab) compared with older therapies (e.g. alendronate, etidronate, risedronate)
- If choosing between one of the newer osteoporosis therapies (i.e., zoledronic acid and denosumab), which one is cost-effective?

Example 7B. Linking Common Technical Ideas to Familiar Concepts. ICERs and ICURs are important measures that help people compare the value of different therapies across different diseases. Even though it shouldn't be the final metric that informs the decision, for reference, the ICER based on number of fractures avoided (rather than number of life years gained or number of quality-adjusted life years gained) could be provided to help people put this analysis in the context of its value and metrics specific to osteoporosis. Alternatively, the amount of actual time that each QALY or LY represents could be presented in days or weeks, which are more familiar to people than proportions. For example, 0.002 QALYs represents less than one day.

SUPPLEMENTARY MATERIALS 3.

Network Meta-Analysis (NMA) Guidance for Researchers Tool Illustrated with Examples from a NMA of Osteoporosis Therapies.

This tool is presented in two parts. Part 1 explains principles researchers could follow when presenting a NMA to policymakers. Part 2 illustrates how these principles can be applied using examples from a NMA of osteoporosis therapies (Maher et al., 2014). In the final version of the tool, moving between these two sections could be facilitated by hyperlinks within the document.

PART 1. NMA Guidance for Researchers

HOW CAN I PRESENT A NETWORK META-ANALYSIS SO THAT IT IS MOST USEFUL TO POLICYMAKERS? Guidance for Researchers.

This guide provides an overview of some evidence-based principles that have been suggested to help policymakers better apply the results of a network meta-analysis (NMA) to a policy decision. Keeping these principles in mind when writing a NMA could enhance the usefulness of your work for policymakers. Examples of how these principles can be applied to a NMA are provided based on a published NMA of osteoporosis therapies (Maher et al., 2014), however, you may be able to identify other examples of how these principles could be applied. How explanations for policymakers are incorporated into a written NMA could vary, depending on the report or journal requirements (e.g. discussion, supplementary material). This guidance should be considered complementary to, but does not replace, reporting guidelines.

HOW CAN I MAKE THE *CONTENT* OF MY NMA MOST USEFUL TO POLICYMAKERS?

1) **Help Policymakers Understand <u>What Evidence</u> an NMA is Based Upon:**
   Policymakers want to know what evidence underlies a NMA. This helps policymakers understand if the evidence will be applicable to their policy decision. Researchers often do this by providing flow diagrams, lists of included and excluded studies or network diagrams. These are important approaches but may require additional explanation such as:
   - Providing explanations of why specific studies (or bodies of evidence) were included or excluded from an NMA, especially if this differs from other published NMAs. This helps policymakers defend and communicate the relevance of the evidence-base to other stakeholders.
   - Analyzing how the exclusion of any key studies could impact the NMA results. When there is disagreement on the included and excluded studies, this can help policymakers understand if including the study would make a practical difference to the results.
   - Simplifying diagrams. This helps focus policymakers on key concepts in a diagram.
   - Identifying how much direct evidence underlies the analysis. This can help policymakers understand how much of the NMA results are based on direct evidence or indirect evidence.

- Identifying key characteristics of the direct evidence (e.g. important effect modifiers), underlies the analysis. This helps policymakers understand how much confidence they can have in the indirect or mixed estimates from an NMA.

See Example 1A. Included and Excluded Trials in an NMA.
See Example 1B. Characteristics of Included Trials.

2) **Help Policymakers Understand What the <u>Limitations</u> in the Analysis Mean:**
Policymakers want to understand the limitations in an NMA so that they know how much they can 'trust' the NMA results. This helps them determine how much confidence to have in the results when applying them to a policy decision. Researchers often do this by identifying any potential limitations in their analysis and in any comparable published analyses, which helps policymakers understand which of the available analyses are most valid and applicable to their policy context. Important limitations to consider in an NMA include whether or not the underlying assumptions of an NMA have been met. However, explaining why key limitations are important and framing the limitations of an analysis in a way that can help policymakers determine how to apply the evidence to the context of their specific policy decision is important. This can be done in different ways, including:
   - Identifying which of the potential limitations are going to have the largest impact on the NMA results, not just identifying all potential limitations. This helps policymakers realize that even if a limitation could be corrected, it may not have a meaningful impact on the results
   - Identifying when limitations are unavoidable, either because no better evidence or methodology currently exists. This helps policymakers realize that this is still the best available evidence, even though it has some limitations associated with it.
   - Acknowledging strengths of the analysis that may outweigh its limitations.

See Example 2A. Risk of Bias in NMA Evidence.
See Example 2B. Confirming NMA Assumptions.
See Example 2C. Understanding Limitations in Ranking Estimates.

HOW CAN I MAKE MY NMA MORE RELEVANT TO A POLICYMAKERS DECISION-MAKING *CONTEXT*?

3) **Help Policymakers Identify How Evidence Differs from <u>Real-World Settings</u>:**
Policymakers want to know if the clinical trials included in an NMA and the subsequent NMA results are generalizable to real-world patients who have the disease (e.g.. disease severity of patients is sometimes greater than clinical trial participants) and how clinicians treat the disease in clinical practice (e.g. treatment doses and protocols may change). This helps policymakers consider how patients and clinicians will be impacted by a policy decision based on the NMA results. If evidence on real-world populations and practices is not included in the data set, sensitivity analyses exploring these issues are not possible. However, even if sensitivity analyses are not possible, identifying these discrepancies for policymakers, and noting the potential direction and magnitude of impact on the NMA results can provide helpful context to policymakers when applying the NMA results to a policy decision.

See Example 3A. Treatment Adherence in Clinical Trials.
See Example 3B. Patient Populations Excluded from Clinical Trials.
See Example 3C. Limited Harms Data from Clinical Trials.

4) **Help Policymakers Adapt an NMA to the Local Decision-Making Context:**
Policymakers want to know if the results of an NMA are applicable to their local setting. This helps policymakers make evidence-informed policy decisions that are relevant to the patients, clinical practices and existing policies in their local setting. Using an NMA that has synthesized many different studies, policymakers may wonder if the results of the NMA are still relevant to their policy setting if some of the included trials in the NMA are not relevant (e.g. different comparators, different doses). Different approaches to presenting these issues can be helpful to policymakers:
   - Identifying those factors in the NMA that are likely to change across local settings (e.g. different doses, subpopulations treated), even if there is insufficient data or power in the NMA to address these factors. This helps policymakers understand challenges they may have applying the results of an NMA to their local setting.
   - When meta-regressions, subgroup analyses or sensitivity analyses can be conducted on local factors, determining which factors have an important impact on the magnitude and direction of the NMA estimates and focusing on these sensitivity analyses when writing the analysis. This helps policymakers understand how the results of the NMA may change when applied to their local setting.
   - Noting that you explored a local factor through additional analyses and that it did not have an important impact on the magnitude or direction of the NMA estimates. This lets policymakers know that the issue was considered, not ignored, and therefore the NMA results are likely still applicable to their local setting.
   - When local factors cannot be addressed through additional analyses, providing policymakers with a description of next steps or approaches they could take to address these factors locally. This helps policymakers as this may be required to conduct additional analyses locally but may not have access to the required statistical expertise.

See Example 4A. Comparators in the Network that Aren't Relevant
See Example 4B. Presenting Subgroup or Sensitivity Analysis.

HOW CAN I PROVIDE MORE *CLARITY* WHEN PRESENTING NMAs TO POLICYMAKERS?

5) **Help Policymakers Identify the Most Important Details by Layering NMA Information:**
Policymakers want to focus on the most important information in an NMA, when trying to make evidence-informed policy decisions. This helps them quickly synthesis the NMA results and determine how they apply to their local policy context. NMAs are sometimes based on a large number of studies and results can be presented in a wide variety of formats. However, when a lot of information is provided, key details can get lost in the write-up. One approach to addressing this is to provide information in a layered approach:
   - Provide the most important information up front in the write-up, and detailed information later in the document or in supplementary materials.
   - Select the best one or two figures or tables that are most relevant for policymakers to see upfront versus in supplementary materials. This helps emphasize the most important NMA data for policymakers to see.
   - Emphasize policy-relevant results but provide detailed results in additional layers. This gives policymakers confidence the analysis has been well-done and allows them to refer to the analysis more in-depth as needed.

See Example 5A. Selecting Most Important Tables or Figures.
See Example 5B. Optimizing Supplementary Information.

6) Help Policymakers Recognize the Key Messages from the NMA:
Policymakers want to understand the key issues and bottom-line of the NMA as it relates to the policy decision to be made. This helps policymakers frame the details of the analysis and put them into the context of the main conclusion and messages from the analysis. This means not only providing the NMA effect estimates, but also trying to emphasize the three or four most important ideas from the analysis that you want policymakers to remember when they are trying to make a decision based on the NMA results. Providing this information in the beginning of the write-up and keeping these ideas concise, makes it more readily accessible and more likely policymakers will remember them.

See Example 6A. Identifying Key Analysis Issues.
See Example 6B. Policy-Relevant Conclusions.

7) Provide Policymakers With the Most Relevant and Understandable Information:
Policymakers want to focus on the information that is most relevant to them, when making evidence-informed policy decisions. This may differ from information that is important for other audiences, including statisticians, researchers, academics, clinicians, patients or the public. This helps policymakers quickly synthesize the NMA information and apply the results to their local policy context. Providing policymakers with the most relevant information to them, can be done in different ways, including:

- Thinking about the policy questions that are of most interest to policymakers when selecting and presenting results.
- Making links between technical concepts that are often encountered and metrics or ideas with which policymakers are more familiar or explaining why the technical issue is important and how it impacts on the results of a subsequent policy decision, e.g. similarities between NMAs and traditional meta-analyses.

See Example 7A. Identifying Policy-Relevant Questions.
See Example 7B. Linking NMAs to Traditional Meta-Analysis Approaches.

PART 2. Illustrated Examples from a NMA of Osteoporosis Therapies.

1) Help Policymakers Understand <u>What Evidence</u> an NMA is Based Upon:

Example 1A. Included and Excluded Trials in an NMA. This NMA included 34 randomized controlled trials (RCTs) of six osteoporosis therapies plus placebo. In searching for evidence, some head-to-head RCTs that only reported changes in BMD, but did not report fractures, were excluded. Because there was no fracture data/fractures reported, we could not include them in the analysis. However, including more head-to-head trials would have added strength to the analysis.

Other NMAs have included therapies such as raxolifene, teriparitide, or strontium. We only included trials of bisphosphonates and denosumab because the question of interest was how does denosumab, a newer biologic, compare with the most commonly prescribed class of osteoporosis therapies, the bisphosphonates. Including more therapies would have increased the volume of data and complexity of the analysis. However, based on examining the published NMAs, it doesn't appear that many more head-to-head trials would have been included in the NMA and so the additional trials may not have given more precise NMA estimates. This is likely because most head-to-head trials measure only BMD (and so would be excluded), rather than fracture rates, which require trials with large sample sizes to show a difference between active therapies in a head-to-head trial. Other NMAs have excluded studies based on trial duration. Although longer trials provide more data on the effectiveness and safety of therapies, we did not limit our NMA by trial duration because this would have excluded some bisphosphonates from the network. However, we did subgroup analyses based on trial duration, which are outlined in Maher et al., 2014 (Supplemental Material 2; Tables 1 and 2).

Example 1B. Characteristics of Included Trials. This NMA included trials conducted in a primary prevention setting and in a secondary prevention setting. Osteoporosis therapies are expected to have different efficacy profiles in primary prevention trials compared with secondary prevention trials. Two network diagrams are provided showing the available evidence for primary and secondary prevention trials. There are substantially more secondary prevention trials (N=20) than primary prevention trials (N=8), however, none of the secondary prevention trials evaluate denosumab. All other therapies are represented in both evidence networks, although sometimes at different doses.

2) Help Policymakers Understand What the Limitations in the Analysis Mean:

Example 2A. Risk of Bias in NMA Evidence. An assessment of the risk of bias in individual studies included in the network was conducted and is described in detail in Maher et al., 2014 (Supplementary Table 2). Key limitations that could have an important impact on the NMA results are also noted. The risk of bias was frequently judged to be high or unclear. The main sources of risk of bias were incomplete outcome data reporting, especially safety data, and unclear bias due to a lack of detail reported on randomization and allocation concealment methods. When the risk of bias is unclear, it is challenging to determine the impact of the magnitude and direction of bias on the NMA results. The comparisons with the lowest risk of bias are those involving zoledronic acid; policymakers should have greater confidence that these pairwise comparisons have less risk of bias than other comparisons in the network.

Example 2B. Confirming NMA Assumptions. Three key assumptions underlying an NMA were assessed for the network of osteoporosis therapies. The assumptions of homogeneity and consistency were both met in this NMA. Therefore, policymakers can have confidence that the estimates in the NMA were obtained through valid methodological approaches. Conducting more subgroup analyses and meta-regressions to explore these assumptions would have been desirable. However, this was not possible with the available data.

Example 2C. Understanding Limitations in Ranking Estimates. Treatment rankings and the probability that a treatment is best were calculated for all therapies in the NMA. Four factors should be considered simultaneously when deciding if you can 'trust' these rankings: the rank of the treatment, the probability that a treatment is best, the difference in magnitude of effect estimates when treatments are ranked closely, and the probabilities that a treatment is second, third or fourth best, etc. When considering these factors among the osteoporosis therapies based on vertebral fractures, the effectiveness of both zoledronic acid and denosumab have a high probability of being the treatments of choice. If a different outcome, or combination of outcomes, were more important, this choice would need to be reconsidered.

3) Help Policymakers Identify How Evidence Differs from Real-World Settings:

Example 3A. Treatment Adherence in Clinical Trials. Treatment adherence in clinical trials is often more optimistic than what occurs when patients receive therapies outside of clinical trials. Adherence should be higher for treatments that are taken less frequently (e.g. zoledronic acid once per year, denosumab once every six months) than daily or weekly regimens (e.g., alendronate, risedronate, etidronate, ibandronate). Treatment adherence was not reported in the majority of trials included in the NMA. In the absence of this full data set, the magnitude of estimates from the NMA should be assumed to be optimistic for therapies such as alendronate; however, the direction of effect is not likely to change.

Example 3B. Patient Populations Excluded from Clinical Trials. It has been suggested that denosumab should be considered for patients who are intolerant to bisphosphonates because it would provide an option for patients who cannot tolerate them. Almost all trials included in the NMA excluded patients who were intolerant to bisphosphonates (otherwise they could not have been ethically randomized to treatment with a bisphosphonate or no bisphosphonate). The only exception to this was the denosumab trial as there was no requirement for a patient to be able to receive a bisphosphonate. However, the proportion of patients in the trial who were intolerant to bisphosphonates was not reported. Policymakers may need to consider how many patients in their setting are likely to be intolerant to bisphosphonates and, if it is an important proportion, how feasible it would be to collect data from them to inform a policy decision as the NMA is limited in what information it can provide on this population.

Example 3C. Limited Harms Data from Clinical Trials. In this NMA, the safety of osteoporosis therapies was considered, which has not been done in published NMAs to date. However, the outcome measured was withdrawals due to adverse events. This is a crude measure of safety and uncommon serious adverse events may be of more interest. These are often harder to synthesize and include in an NMA because safety data are not routinely and proactively collected in clinical trials. Rare or serious adverse events associated with osteoporosis therapies that have been noted in Canadian product monograph warnings and for which there is insufficient data to synthesize in a NMA are summarized in a table (Supplementary Table 5).

4) Help Policymakers Adapt an NMA to the Local Decision-Making Context:

Example 4A. Comparators in the Network that Aren't Relevant. Ibandronate was included in the NMA, although it is not a bisphosphonate that has regulatory approval in Canada for the prevention of osteoporosis. This does not invalidate the network and policymakers should still consider this over NMAs that exclude ibandronate. Because head-to-head trials of alendronate and ibandronate exist, including both bisphosphonates in the analysis allowed the additional evidence to provide strength and precision in the network. Policymakers can still focus on comparisons from the NMA that are of most interest (i.e., different comparisons of placebo, alendronate, zoledronic acid and denosumab), while gaining strength from including the ibandronate trials.

Example 4B. Presenting Subgroup or Sensitivity Analysis. This NMA included both primary and secondary prevention trials. Policymakers may want to understand how treatment effectiveness differs in these two categories. Subgroup analyses were conducted (Maher et al. 2014, Supplementary Materials 2, Tables 1 and 2). However not all therapies had primary and secondary prevention trials and the small patient numbers in each of the subgroups led to uncertain estimates that should be interpreted cautiously when trying to inform a policy decision. Nonetheless, the direction of effects was similar to those observed in the overall NMA analysis.

5) Help Policymakers Identify the Most Important Details by Layering NMA Information:

Example 5A. Selecting Most Important Tables or Figures. If selecting only one or two NMA figures or tables to provide to policymakers, the importance of different outcomes must be considered. Vertebral fractures have the greatest amount of evidence but, as measured in this NMA, also include asymptomatic radiographic fractures, which have limited clinical relevance. Hip fractures are the most serious type of osteoporotic fracture and are associated with permanent disability, institutionalization, and increased mortality. Therefore, it would be most relevant to emphasize hip fractures. A matrix table of hip fractures shows how each of the therapies compares with each other so that policymakers can determine which comparison is most relevant for their local setting. Results for the remaining outcomes and for NNTs could be explained in the text as not many statistically significant results were observed. Similarly, statements on risk of bias and critical appraisal of the evidence could be included in the text and detailed tables provided in supplementary materials. A network diagram that outlines the evidence included in the trial for hip fractures should also be selected. Different versions of the network diagram could be provided in supplementary materials for different outcomes.

Example 5B. Optimizing Supplementary Information. Many journals or reports allow for supplementary appendices to be created. While this is a useful feature, supplementary materials should be selected judiciously. For the osteoporosis NMA, important information to include in a supplementary appendix that policymakers might be particularly interested in having access to could include:

- A PRISMA flow diagram, clearly documenting included and excluded studies
- A table of key characteristics of included trials that influence the NMA results
- Figures or tables demonstrating that key NMA assumptions have been checked and verified
- Forest plots of direct and indirect estimates versus the reference standard so there is a side-by-side comparison of the direct and indirect estimates. This also allows people who prefer figures to tables to view these data rather than the matrix table. Providing the forest plot versus the reference standard limits the comparisons that can be viewed but is a pragmatic choice

Depending on the importance of findings and formatting requirements, some of these suggested supplementary tables and figures could be moved up into the next layer of a report or analysis.

6) Help Policymakers Recognize the Key Messages from the NMA:

Example 6A. Identifying Key Analysis Issues. The key points to remember in this NMA that have an important impact on the NMA results are:

1) The analysis includes 34 studies but only two of the studies (both comparing alendraonte with ibandraonte) are head-to-head trials so the majority of the NMA estimate are derived from indirect data
2) Vertebral fractures were heterogeneously defined across included studies and frequently included radiographic fractures, which are asymptomatic and have limited clinical relevance.
3) No differences in withdrawals due to adverse events WDAEs were detected but differences between therapies in rare serious adverse events are unclear and there were insufficient data to synthesize these in the NMA.
4) No differences in wrist fractures were found but this may be because of the low event rate in trials and the small number of trials in the analysis
5) No trials included patients intolerant to bisphosphonates

Example 6B. Policy-Relevant Conclusions. Important policy relevant-conclusions that can be drawn from this analysis include:

1) Only alendronate (NNT=188) and zoledronic acid (NNT=212) are effective at preventing hip fractures compared with placebo.
2) Alendronate is the only therapy effective at preventing vertebral fractures (NNT=28), non-vertebral fractures (NNT=80) and hip fractures (NNT=188) compared with placebo
3) The effectiveness of osteoporosis therapies in preventing clinically-relevant symptomatic vertebral fractures is unknown, given how vertebral fractures were measured in the clinical trials (i.e. asymptomatic radiographic fractures were included). Therefore, the numbers to treat to prevent a clinical vertebral fracture are unknown. However, based on the available data, all therapies with regulatory approval in Canada are effective at preventing vertebral fractures. Of these, the magnitude of effect was greatest for zoledronic acid (OR=0.28,95%CI:0.19, 0.39) and denosumab (OR=0.31,95%CI:0.22,0.44).Both were more effective than the current standard of care, alendronate (OR=0.53, .95%CI: 0.40, 0.70 for alendronate compared with placebo)
4) Only alendronate (NNT=80) and risedronate (NNT=70) are effective at preventing non-vertebral fractures.
5) No conclusions can be made about the effectiveness of the different osteoporosis therapies on wrist fractures due to their uncommon occurrence in clinical trials.
6) Withdrawals due to adverse events were similar across therapies but a more detailed analysis of rare serious adverse events would be a better indication of the comparative safety of osteoporosis therapies

7) Provide Policymakers With the <u>Most Relevant and Understandable Information</u>:

Example 7A. Identifying Policy-Relevant Questions. When considering the current landscape of osteoporosis therapies, a number of policy-relevant questions can be identified. Ensuring that the selected NMA results help inform these questions will increase the relevance of the analysis:

- What is the most effective osteoporosis therapy?
- What is the effectiveness of newer therapies with less frequent administration (e.g. zoledronic acid, denosumab) compared with older therapies (e.g. alendronate, etidronate, risedronate)
- If choosing between one of the newer osteoporosis therapies (i.e., zoledronic acid and denosumab), which one would be more effective?

Example 7B. Linking NMAs to Traditional Meta-Analysis Approaches. In the NMA of osteoporosis therapies, 34 trials were combined based on the assumption that the trials are similar enough on key clinical and methodological characteristics that could impact the effectiveness of osteoporosis therapies (e.g. age at menopause, previous fracture history, risk of bias), i.e. the homogeneity assumption was met. Trials that differ on these key characteristics should not be combined in either a traditional meta-analysis or an NMA of osteoporosis therapies. It was also assumed that the indirect estimates obtained from the NMA reflect estimates that could be obtained from head-to-head trials of therapies if they had actually been conducted, i.e. the consistency assumption was met. This assumption is unique to conducting NMAs but is an extension of the homogeneity assumptions. These assumptions were assessed on both a conceptual level for the osteoporosis NMA by comparing trial characteristics and using statistical approaches. If either of these assumptions were not met, the results of the NMA would not be reliable.

CHAPTER 9 - CONCLUSIONS

This chapter:

- Summarizes background on the thesis objectives and provides a general description of the toolkit and the evidence-base supporting its development
- Integrates findings from the thesis in the light of published research and considers implications for policy and practice, methodological research and theory development
- Considers the strengths and weakness of the research conducted in this thesis as well as the novelty of this thesis
- Identifies next steps in tool development and evaluation and proposes recommendations for future research
- Draws conclusions, considering the contribution of this thesis to the field and what it may help achieve in future.

9.1 BACKGROUND

Summary of Thesis Rationale

Transparency is a key principle underlying fair and legitimate health technology assessment (HTA) processes and related policy decisions (Daniels & Sabin, 1997). Transparency requires not only providing sufficient information, but also providing information in an accessible and understandable format for end-users. This is especially relevant when complex methodologies, such as economic evaluations and network meta-analyses (NMAs), form the basis of an HTA and are used to support evidence-informed policymaking. Approaches that present analyses and results in ways that are uncomplicated, but without altering their technical meaning, are challenging to obtain.

Knowledge translation (KT) tools and principles that have been applied in other areas of evidence-based medicine may offer opportunities to improve the communication and presentation of economic evaluations and NMAs. While reporting guidelines can contribute to enhancing the transparency of analyses (Moher et al., 2011), experience in the realm of clinical trials and evidence-based medicine has shown that they may be insufficient for meaningful communication to non-technical audiences and alternate tools and formats of sharing information such as decision-aids and clinical practice guidelines have been developed for patients and clinicians, respectively (Straus et al., 2013). Enhancing the accessibility of economic evaluations and NMAs may also increase their impact and value to policymakers. In the current resource-constrained environment of limited budgets and rising drug costs (Sullivan et al., 2011; Niezen et al., 2006), there is an even greater need for policymakers to understand the cost-effectiveness and comparative effectiveness of new drugs.

Thesis and Toolkit Objectives

The objective of this thesis was to explore the informational needs of policymakers with respect to economic evaluations and NMAs as well as researchers' approaches to communicating this information. The intent was to use these findings to develop a toolkit that could enhance the relevance, transparency, and accessibility of economic evaluations and NMAs for policymakers. The specific goals of the toolkit were to:

(1) Assist policymakers in understanding and identifying key elements of complex methodologies such as economic evaluations and NMAs;
(2) Support policymakers in determining how to adapt complex analyses to their local drug funding policy context; and
(3) Enhance the interface between researchers, i.e. knowledge producers, and policymakers, i.e. knowledge users, when sharing information on economic evaluations and NMAs.

Proposed Toolkit and Supporting Evidence

The proposed toolkit (Chapter 8) was developed by following the Knowledge-to-Action (KTA) framework (Straus et al., 2013). It addresses three key areas: communication, education and accessibility. It includes six different tools tailored to either policymakers or researchers, and each applied to economic evaluations and NMAs (i.e., 12 toolkit components), namely: guidance for researchers, policymaker question and answer (Q&A) documents, technical analysis formats, learning modules, a tool repository and a lay peer-review process. Three guiding principles, identified through this thesis research, were also applied, ensuring that the tools: (1) address key content issues in economic evaluations and NMAs, (2) identify the policy or decision-making context of the analyses and (3) clarify information by optimizing presentational approaches.

Four systematic reviews and two sets of qualitative interviews, provided the primary evidence-base for developing the toolkit:

- Two systematic reviews of the literature identifying approaches to communicating; barriers and facilitators to understanding and using; and general knowledge of economic evaluations (Chapter 3) and NMAs (Chapter 5).
- Two systematic reviews of guidelines on conducting and/or reporting economic evaluations (Chapter 4) and NMAs (Chapter 6) to identify what guidance researchers are provided on presenting analyses
- Key informant interviews with: (1) Canadian policymakers, i.e. knowledge users, on their experiences using information from economic evaluations and NMA to inform drug funding or other policy decisions; and (2) senior individuals in HTA organizations around the world, i.e.

knowledge producers, on what approaches they take when communicating information on economic evaluations and NMAs (Chapter 7).

These papers suggest that policymakers' information needs and researchers' approaches to presenting economic evaluations and NMAs overlap but are not always aligned. Researcher's presentation strategies are more often aligned with guideline standards or journal requirements than with end-users' needs to apply analyses to policy decisions. Economic evaluations are more established than NMAs, however common themes exist. Most barriers, facilitators and preferences relate to three key areas: robustness of content, adapting to the policy context and clarity of information. Few specific KT approaches have been developed or applied internationally to enhance the accessibility of economic evaluations and NMAs. In this thesis, these findings were also considered in the context of existing research on related topics such as methodological developments in health economics and NMAs, frameworks for knowledge translation and implementing evidence in practice, and policymakers' needs and preferences.

9.2 DISCUSSION OF FINDINGS

In this section, findings from different components of this thesis are compared and contrasted with each other as well as compared with existing research. This discussion is framed in terms of five key areas: implications for practice, implications for research, implications for theory, strengths and limitations of the thesis and novelty of the research.

Implications for Policy and Practice

A number of implications for policy and practice are identified as a result of the this thesis including: the role of clinicians, patients and the public as end-users of economic evaluations and NMAs; the importance of capacity building in the field of health economics and NMAs; and the impact of health systems pressures on economic evaluations and NMAs.

Clinicians, Patients and the Public as End-Users

The primary target audience for this toolkit is policymakers and their role as end-users, i.e., knowledge users, is the focus of this thesis and discussed throughout. However, other stakeholders, who are impacted by policy decisions, such as clinicians and patients, are indirect end-users of economic evaluations and NMAs who often influence policymakers' needs. Initial scoping searches when defining the objectives of this thesis, suggested that little primary research has been conducted on clinicians' and patients' needs with respect to understanding the

role of economic evaluations and NMAs in health policy. However, with rising drug costs impacting clinicians access to therapies (Sullivan et al., 2011; Niezen et al., 2006) and the growing engagement of patients and patient groups in HTA processes (Fattal & Lehoux, 2008; Gagnon et al., 2011), further consideration should be given to assessing the needs of these stakeholders and determine how to tailor economic evaluations and NMAs to them (Leonard et al., 2010; Rosenbaum & Lamas, 2012). Approaches to tailoring tools and information to patients and clinicians have been developed and refined in the field of evidence-based medicine through patient decision-aids and practice guideline summaries (Straus et al., 2013). The legitimacy of policy decisions rests on their transparency and the ability of individuals, such as patients and clinicians, to understand the rationale for a decision that impacts upon them (Daniels & Sabin, 1997). These were needs raised in key informant interviews with Canadian policymakers and international HTA organizations. Some general approaches such as developing key messages for clinicians and patients have been applied to HTAs and could be further tailored to economic evaluations and NMAs (see Chapter 7).

Capacity Building

One of the goals of the toolkit is to enhance the interface between researchers and policymakers. The limited available expertise to conduct economic evaluations and NMAs was frequently cited as a barrier both in the literature and in qualitative interviews with both Canadian and international participants. Capacity building efforts in these areas are of paramount importance (Abdelhamid et al., 2012; Madden et al., 2009; Thabane et al., 2011). However, capacity building requires not just recruiting and developing new economists and researchers capable of conducting economic evaluations and NMAs; it also merits optimizing existing analysis approaches and helping researchers better tailor their work to end-users' needs. Optimal approaches for enhancing this interface should be explored and will require greater engagement of researchers in the development of the toolkit. Additionally, engaging researchers will also be important in refining and optimizing presentational approaches for NMAs that could be used in technical analyses. This is because NMA methodology is evolving and there is a need to ensure that the identified presentational formats accurately represent key issues. While researchers' primary focus is conducting analyses, the more researchers, i.e., knowledge producers, are able to interact with end-users, i.e. knowledge users, and understand their needs when reporting analyses, the greater the impact such analyses can have on policy decisions. (CIHR 2012, Choi et al., 2005, Hyder et al., 2011). Although enhancing the interface between researchers and policymakers may, on its own, be insufficient to influence evidence-informed policy-making, it is still an important capacity building initiative, that could lead to more transparent and accessible economic evaluations and NMAs that meet policymakers' informational needs.

Health Systems Pressures

While barriers and facilitators outside the scope of the toolkit were identified during this thesis, they should not be dismissed. Canadian healthcare and health policy is part of a dynamic and interdependent system that influences the availability of evidence and how it is applied in practice. For example, market authorization and regulatory requirements for placebo-controlled trials impact the use of NMA (Mills et al., 2013). Similarly, policies on drug pricing impact the use of economic evaluations (Morgan et al., 2013). Variability across provincial healthcare contexts and disparities in available expertise, budgets and provincial mandates also have an important influence (Daw & Morgan, 2012). Enhancing communication of economic evaluations and NMAs is only one approach to increasing the legitimacy and confidence in health policy decisions. Other avenues or research should also be explored to address some of these health systems pressures that impact the use of economic evaluations and NMAs.

Implications for Methodological Research

A number of implications for methodological research are identified as a result of this thesis including: evolving methods for conducting NMAs, research on the interpretation and optimal presentation of NMAs for end-users; and better understanding researchers' perspectives on barriers and facilitators to producing economic evaluations and NMAs for policymakers,

Evolving Methods for Conducting NMAs

The emerging nature of NMAs makes it a unique topic with additional complexities that economic evaluations do not have. While considerable thought is currently being invested in how to optimize the presentation of NMAs given the challenges researchers have had interpreting them (Chaimani et al., 2013; Tan et al., 2014), little attention has been given to the needs of non-technical audiences such as policymakers, clinicians and patients. Despite the current challenges with an emerging methodology such as NMAs, this also offers an opportunity to learn from research in existing fields including economic evaluations (Galani et al., Williams et al., 2008), policymaker preferences (Lavis et al., 2005) and KT approaches used with meta-analyses (Straus et al., 2013). Building on the similarities of NMAs and meta-analyses offers avenues for enhancing the presentation and communication of NMAs (Salanti et al., 2012). Allowing for the parallel development of formats that methodologists value while considering formats that address policymakers' needs and preferences, rather than conducting this work sequentially, could lead to more rapid acceptance of NMAs and incorporation into regular policymaking, than would occur otherwise. Even if this is challenging to do while methodologies are evolving so rapidly, having researchers be more aware of these issues, as is attempted through the NMA Guidance for Researchers tool would be of value.

Research on the Optimal Presentation of NMAs and End-Users' Needs

Although reporting guidelines exist for economic evaluations and are in development for NMAs, the focus of reporting guidelines may not be communication to non-technical audiences. This premise is supported by findings from the systematic review of approaches to communicating economic evaluations, indicating that policymakers sometimes have challenges determining how to apply the results of economic evaluations to real-world decisions (Galani et al., 2008; Williams et al., 2008). Reporting templates were sometimes provided in economic guidelines and the importance of economic evaluations and/or NMAs to inform policymakers and clinicians' decisions was often noted (ISPOR Guidelines around the World, 2014). However, approaches to tailoring analyses to these audiences were rarely identified, despite the fact guidelines were usually developed by HTA organizations or related collaborations. Fewer NMA guidelines exist than economic evaluation guidelines, likely due to the nascent NMA methodology and slow, albeit growing, uptake in health policy decision-making. The current focus of NMA guidelines is defining methodological standards as numerous studies have noted the complexity of NMA methods, limitations in the methodologies used in published analyses and the need for better guidance (Cipriani et al., 2012; Mills et al., 2013). The systematic review of economic guidelines suggests that economists have standard formats for presenting economic evaluations that sometimes differ from policymakers' preferences. It may require more time before NMA technical approaches become standardized, but similar to economic evaluations, attempts should be made to better align and focus these technical presentation formats with policymakers' preferences.

Eliciting Researchers Perspectives on Creating Economic Evaluations and NMAs to Inform Policy

While the primary target audience for this toolkit is policymakers, researchers are a secondary target as they produce economic evaluations and NMAs for policymakers. Although there is existing literature in other fields on the challenges policymakers face when provided with research evidence (Hyder et al., 2011; Lavis et al., 2005; Murthy et al., 2012), there is little research documenting challenges researchers have in determining how to tailor and present evidence to policymakers. In this thesis, researchers' perspectives were determined indirectly from a review of methodological and reporting guidelines that they frequently use. While some researchers were part of the Toolkit Advisory Panel or participated as key informants in interviews with HTA organizations, capturing the perspectives of knowledge producers, not just knowledge users is important (CIHR 2012). This could include identifying researchers' perspectives on barriers and facilitators to developing economic evaluations and NMAs for policymakers and current approaches they have for engaging with policymakers. More primary research exploring researchers experiences would be a novel contribution to the field of economic evaluations and NMAs.

Implications for Theory

Implications for theory were identified as they relate to similarities and differences between economic evaluations and NMAs and how this could contribute to framework for optimizing the integration of future novel methodologies into health policy decision-making processes, as well as similarities and differences with existing frameworks for implementing evidence in practice or policy.

Contrast Between Findings Relevant to Economic Evaluations and Relevant to NMAs

This thesis addressed two areas with complex technical methodologies that are used to inform HTAs, economic evaluations and NMAs. A number of similarities were observed between the two fields, relating to policymakers' needs. These included understanding the evidence-base of the analysis, the uncertainty associated with results, limitations in the analysis; understanding how evidence relates to the real-world or local decision-making context; and presentational approaches including layering of details, emphasizing key messages. However, these two fields are also at different time points in their development and uptake in supporting health policy decision-making. While the use of economic evaluations has been established for the past 20 years, NMAs have only recently emerged (Drummond et al., 2013; Salanti et al., 2012). Differences observed between the two fields may be due, in part, to this timeline, as well as their differing theoretical bases. Successful approaches that enhance the communication of economic evaluations, such as disaggregating results, (Galani et al., 2008; Williams et al., 2008) were not specifically identified for NMAs. However, they could be generalizable to emerging methodologies such as NMAs. Evaluations of tools incorporating these approaches for NMAs could provide insight on generalizable concepts that could be applied in future to any new emerging methodology that begins to influence and inform health policy. Determining how to better integrate novel methods into practice and policy may be an area for future research, building on current experiences with NMAs and previous experiences with economic evaluations.

Building on Existing Frameworks for Implementing Evidence

The thesis was unique from most existing knowledge translation research initiatives in two important respects. One, it focused on economic evaluations and NMAs rather than other types of evidence syntheses. Two, it focused on communication approaches to enhancing the transparency and accessibility of evidence, not simply approaches that lead to greater use of evidence in practice or policy. However, successfully changing behaviours is a key principle underlying most frameworks for knowledge translation and implementing evidence, and is still a

relevant concept, despite these unique features. Effectively enhancing the transparency and accessibility of economic evaluations and NMAs will require behaviour change among key stakeholders. Our toolkit was developed following the KTA framework, which is a framework based on a theory analysis of 31 planned action theories. Planned action theories outline how planned changes occur and identify variables that influence those changes (Graham et al., 2006). However, the KTA framework is only one framework that has been developed to guide the implementation of evidence into practice or policy. This thesis' findings also have similarities with other frameworks, including the Promoting Action on Research Implementation in Health Services (PARIHS) and the Behaviour Change Wheel (Michie et al., 2011; Rhycroft-Malone & Bucknall, 2010). For example, in PARIHS, successful implementation of evidence in clinical practice is a function of the evidence, context and facilitation. All of these concepts are represented in some degree by themes identified in this thesis, which we have referred to as content, context and clarity. The Behaviour Change Wheel is a relatively recent framework that is based on a comprehensive review of frameworks for behaviour change interventions. It is based on the premise that the three essential requirements for behaviour change are capability, motivation and opportunity. The Behaviour Change Wheel identifies nine types of interventions and seven policy approaches for addressing these requirements and changing behaviour. Some of the policies and interventions in the Behaviour Change Wheel overlap with our proposed tools, including communications (e.g. policymaker Q&A documents), guidelines (e.g. guidance for researchers), education (e.g. learning modules), service provision (e.g. a lay peer-review process) and enabling interventions (e.g. a tool repository).

While general concepts overlap between different frameworks, how concepts are specifically defined in each framework can highlight important details and differences. For example, in the PARIHS framework, key features of context include organizational culture, leadership and evaluation measures. (McCormack et al., 2002) As it relates to HTAs, contextual elements that are frequently raised in the literature include legal, social and ethical dimensions that influence how evidence is implemented. (Potter et al., 2008) Economists have devoted efforts to determining how economic evaluations can be adapted to different local contexts, which is often discussed in terms of 'transferability' (Chase et al., 2009; Goeree et al., 2011). Local contextual elements specific to economic evaluations and NMAs that were identified in this thesis, and which may vary across Canadian provinces and jurisdictions included patient populations, drug utilization, available therapies and drug prices. These elements could be used to more comprehensively define 'context' as it applies to economic evaluations and NMAs and how context can be systematically addressed in relevant knowledge tools.

The KTA framework, the PARIHS framework and the Behaviour Change Wheel are a starting point to compare and contrast how key concepts are defined (e.g. context) and if other elements within these frameworks (e.g. legislation, environmental restructuring, facilitation) apply to enhancing the transparency and accessibility of economic evaluations and NMAs. This could further strengthen the toolkit and provide a stronger theoretical basis for the knowledge translation of economic evaluations and NMAs. Furthermore, many knowledge translation frameworks focus on implementing evidence in clinical practice and are not specific to implementing evidence in policy environments. Further exploring the development of a framework specific to knowledge translation of evidence-informed health policy could build upon the findings of this thesis.

Strengths and Limitations

The toolkit evolving from this thesis is necessarily influenced by the evidence-base underlying its development. Systematic reviews of economic and NMA literature were unable to identify strong study designs such as randomized controlled trials (RCTs), evaluating different communication approaches or presentation options tailored to policymakers, however such trials have been conducted with other types of knowledge tools (Akl et al., 2012; Brewer et al., 2012; Dobbins et al., 2009). In the absence of RCTs, studies on users' experiences and preferences with economic evaluations and NMAs were considered. Although economic literature provided numerous qualitative studies exploring policymakers' preferences (Galani et al., 2008; Hoffman et al., 2002; Thurston et al., 2008), the evidence-base for NMAs was limited even further due to its emerging nature and no published information was available on policymaker preferences for NMAs. Only studies of researchers' experiences with NMAs were identified (Abdelhamid et al., 2012; Coleman et al., 2012). When trying to assess knowledge and barriers or facilitators to the use of NMAs, few studies formally assessed this and indirect indicators were used as a proxy measures. These limitations should be recognized as the evidence-base in these fields evolves. However, this is currently the best available evidence in this field and should be considered a snapshot of potential issues to be further explored in appropriately designed studies. In the absence of these studies, the key informant interviews conducted as part of this thesis (Chapter 7), take on greater importance and the contribution they make to the NMA literature.

One of the greatest strengths of this thesis is the mixed-methods approach that integrated findings from systematic reviews of different bodies of literature and in different fields with qualitative interviews conducted with both Canadian policymakers and individuals from HTA organizations around the world. No other approaches that were identified in the literature have a similar or such a substantive evidence-base. The use of a mixed methods approach allows for

the integration and synthesis of data from multiple sources and methodological approaches and confirmation of findings leads to greater strength and confidence in the conclusions of this thesis (Greenhalgh et al., 2004; Miles & Huberman, 2013). For example, barriers and facilitators that policymakers reported in the literature such as access to experts, volume of information and understanding how analyses apply to policy decisions were often confirmed in key informant interviews. As well, similarities were identified between guideline standards for presenting economic evaluations and policymakers' preferences for disaggregated and transparent information. Qualitative interviews indicated that features of economic evaluations and NMAs that were important to policymakers centered on content, context and clarity; each of the findings from the systematic reviews could also be categorized into one of these three themes.

Exploring two related but unique topic areas, economic evaluations and NMAs, further strengthens the thesis. The two fields of evidence synthesis are at different time points in their development and uptake in supporting health policy decision-making. While economic evaluations have been used to support health policies over the last 20 years, NMAs have only recently emerged. The ability to identify challenges, successes or approaches from the field of health economics that could be applicable, not only to NMAs, but across any new and emerging methods can benefit other fields in the future. Integrating research and ideas across multiple content domains (i.e., economic evaluations and NMAs) as well as relating it to existing frameworks and approaches for implementing evidence (Graham et al., 2006; Rycroft-Malone & Bucknall, 2010), policymakers' needs and preferences (Lavis et al., 2005) and approaches to communicating technical information to lay audiences (Hildon et al., 2012; Trevena et al., 2006) to develop the toolkit, integrates evidence sources and creates a robust evidence-base for the toolkit.

This thesis is also strengthened by the broad nature of the systematic reviews that were conducted. While this led to a heterogeneous body of research that was sometimes challenging to synthesize and critically appraise, the novelty of research in this area required that a broad overview of existing evidence be summarized for consideration. Extensive searches of the grey literature were conducted in order to identify relevant guidelines and research produced by HTA organizations, which many academic papers have overlooked. The broad nature of the literature identified, especially as it relates to economic evaluations, should also be considered with respect to the period of time over which it has been published as well as the range of geographic regions. While many issues were identified in the early days of the use of economic evaluations to inform policy decisions, it is unclear how relevant these issues are in today's Canadian context where economic evaluations have been in play for a considerable period of time (Laupacis et al., 2005). Key informant interviews with policymakers suggest that they are comfortable with general

economic concepts but that their current focus is on applying concepts to actual policy decisions, taking into consideration the influence of local contextual factors on cost-effectiveness estimates. Further research on how to optimize the transferability of economic evaluations to different local contexts in the Canadian policy environment is warranted, especially with the advent of the Pan-Canadian Pricing Alliance (PCPA). Considerable research has explored transferability in international contexts and learnings could be applied to the challenges faced in the Canadian context (Goree et al., 2006; Chase et al., 2009).

This thesis was strengthened by its grounding in theory and the application of the Knowledge-to-Action (KTA) framework. Although different approaches to presenting NMAs and economic evaluations were identified in literature reviews, few of these approaches were grounded in theory. The KTA framework allowed for explicit consideration of barriers and facilitators to the use of economic evaluations and NMAs (e.g., expertise, complex content, formatting approaches), the incorporation of local factors that can alter or impact how results of economic evaluations and NMAs are used to inform decisions (e.g., drug pricing, prevalent populations, availability of drug therapies, doses used in clinical practice) and the tailoring of tools to different end-users' needs (e.g. policymakers, researchers) (Straus et al., 2013).

Although toolkit components have been proposed based on an integrated review of evidence from four systematic reviews and two interview sets, a comprehensive survey of existing tools was not conducted. The primary purpose of key informant interviews with HTA organizations around the world was to identify a sample of existing tools but few specific approaches were identified, suggesting that little formal use and development of KT approaches exists in this area. However, this was not a comprehensive search for tools and did not include other potential sources such as tools developed by professional groups and collaborations, e.g., International Society for Pharmacoeconomics and Outcomes Research (ISPOR), the Cochrane Collaboration; tools that exist in related health-policy fields that could be adapted to economic evaluations and NMAs (e.g. Lavis et al., 2009; Schunemann et al., 2003, DECIDE Collaboration, 2014); or tools summarized in academic literature that may not have met the inclusion criteria for the systematic reviews underlying this thesis because they were not developed using a defined methodology or because they were supportive tools and not actual approaches to communicating economic evaluations and NMAs (Caro et al., 2014; Jansen et al., 2014).

The proposed toolkit has the advantage of being a multi-component, multi-strategy intervention that targets different audiences. Multi-component interventions have the potential to be more effective than single interventions if they target different aspects of behaviour change . (Wensing et al., 2013). However, they also require greater resources and may be less feasible to

implement. Therefore, evaluation of the toolkit should determine which components are more effective to others so that knowledge translation resources can be efficiently allocated. No such toolkits to enhance the accessibility of economic evaluations and NMAs for policymakers exist currently, creating a highly original thesis that has real and timely implications in an environment of constrained resources, limited budgets, growing number of new drugs and rising drug costs (Sullivan et al., 2011; Niezen et al., 2006). A number of issues were raised in key informant interviews and were identified in the literature that are outside the scope of this thesis, however, aligning this thesis with other initiatives exploring health systems pressures, stakeholder interactions and capacity building initiatives, would strengthen the impact of the toolkit (Straus et al., 2013).

Another frequently raised issue was who is actually considered a policymaker. For the purpose of this thesis, policymakers included not solely the final decision-maker but also senior individuals within governments, on expert advisory committees or within HTA organizations who inform and influence policy decisions. While this is a broad definition and perspectives of each of these groups may be unique, there was sufficient similarity to lump them together at the early stages of research in this field. Further development and refinement of the toolkit components and additional primary research may further elicit distinctions between these different types of policymakers.

Novelty of this Thesis

This is the first and most comprehensive body of work in the field of health economics and NMAs to integrate evidence related to both researcher's, i.e., knowledge producer's, and policymaker's, i.e., knowledge user's, perspectives on how to best communicate these analyses to end-users. Its primary novelty rests on the fact that a comprehensive and integrated evidence-base has been created in a field where none previously existed, despite challenges and issues being raised over time and in different forums. While many individual studies have focused on either researchers or policymakers, few have identified possible solutions based on an integrated synthesis of the different, but complementary needs, of knowledge producers and knowledge users (CIHR 2012). Opinion papers and editorials have suggested there is a gap between these two groups but concrete solutions for addressing these gaps have not been identified specific to the areas of economic evaluations and NMAs (Choi et al., 2005; Hyder et al., 2011). However, it's also important to note that the integrated approach taken in this thesis to considering various stakeholders' needs also identified some common ground between researchers and policymakers, which should be emphasized going forward.

There are a number of characteristics of this thesis that underscore its originality, novelty and contribution to the literature. Although Canadian policymakers have been surveyed on a number of topics including informational needs and product listing agreements (Greyson et al., 2013; Morgan et al., 2013), this is the first published research to qualitatively interview Canadian drug policy decisions makers on their informational needs with respect to using and understanding economic evaluations and NMAs. With evolving Canadian policies on drug pricing and regulatory requirements, this thesis provides an important contribution to better understanding the application of these evidence syntheses and how they may adapt to and integrate with new and emerging policies. In addition, this is the first body of research in the world that has attempted to identify international KT approaches specific to economic evaluations and NMAs. The high response rates in both sets of interviews and data from the interviews suggested a strong interest in and desire for more research on the topic. Although some research has been done in this area as it relates to the use of economic evaluations in other geographic locations such as the UK and Australia, none exists in the Canadian context. (Baghbanian et al., 2011; Wiliams et al., 2008) Furthermore, no published studies exist that survey end-users of NMAs. Only two qualitative studies were identified in the literature, one that surveyed Cochrane reviewers and one that elicited views from NMA authors (Abdelhamid et al., 2012; Coleman et al., 2012). Therefore, this thesis is pioneering and much needed as it can contribute to the future development of NMA approaches that take into consideration knowledge users' needs.

Theory-based strategies to change behaviour and enhance the use of evidence in policy and practice have been applied to systematic reviews and meta-analyses for many years but have not yet been applied specifically to economic evaluations or the emerging field of NMAs. Learning from these strategies and the existing evidence on policymakers' needs and preferences when using evidence will be of value in the field of economic evaluations and NMAs. Key informant interviews with HTA organizations around the world supported that this is a novel approach to enhancing the use of economic evaluations and NMAs in health policy decision-making.

9.3 FUTURE RESEARCH AND CONCLUSIONS

In this section, both immediate next steps for further research and long-term recommendations for development of a research programme are proposed. Conclusions based on this thesis are then made, considering its contribution to the field and what it may help achieve in future.

Next Steps for the Economic Evaluation and NMA Toolkit

Two areas of further research and have been identified as key next steps following the work completed for the thesis, namely: further development of the toolkit components and evaluation of the toolkit.

Toolkit Development

When tools were ranked and prioritized in Chapter 8, the next toolkit component to be developed is the Policymaker Q&A document for policymakers. The goal of the policymaker Q&A will be to outline the most relevant issues in an economic evaluation and NMA for policymakers in a format that is easily understood and that highlights implications for decision-making. Although the target audience for the completed Q&A document is policymakers, a template on how to complete the Q&A document will also be provided for researchers. Researchers are key stakeholders that should be engaged in the further development of the toolkit to gain their perspectives on barriers and facilitators to interactions with policymakers. Soliciting the perspectives of researchers on interactions with policymakers about economic evaluations and NMAs would strengthen the development of the policymaker Q&A document and the toolkit, overall.

When soliciting researchers views, joint researcher-policy dialogues will offer a novel method of obtaining information on interactions between these two important stakeholders. Deliberative dialogues that engage participants from various fields on a topic of interest, such as increasing the accessibility of economic evaluations and NMAs and that ground dialogues in a strong evidence-base, as were created in this thesis through the conduct of systematic reviews and key informants, have been demonstrated to both explore key characteristics of relationships between participants and build capacity. (Boyoko et al. 2012; Bombard et al., 2011) Building capacity in both researchers and policymakers is an important goal that could enhance HTA processes and evidence-informed policy in the healthcare system (Thabane et al. Madden et al., 2009). Deliberative dialogues can enhance mutual understanding between participants, build relationships, enhance communication skills and identify new approaches and solutions to policy issues and strengthen organizational and system capacities, all of which could benefit the ongoing development of the economic evaluation and NMA toolkit (Boyoko et al. 2012)

Following the policymaker Q&A document, four tools (eight toolkit components) remain to be developed in the toolkit: technical analysis formats, learning modules, a tool repository and a lay peer-review process. Developing each of these may benefit from considering approaches developed outside the field of economic evaluations and NMAs. For example, reviews of web-

based continuing medical education programs could inform the development of these learning modules (Cook & Dupras, 2004; Cook et al., 2008). Considering the design characteristics and factors associated with the successful implementation of existing repositories such as the NHS economic evaluation database (NHS, 2014) the Rx for Change database (CADTH, 2014) and the Evidence-Informed Health Systems Renewal Portal (CIHR, 2014) would provide insight on the optimal development of a tool repository for policymakers and/or researchers as it relates to providing access to existing economic evaluation and NMA resources (DePourville et al., 2005). Many organizations are focusing more efforts on engaging patients and the public in evidence-generation and review processes (Leonard et al., 2010; Cavazza et al., 2012; Gagnon et al., 2011). For example, the British Medical Journal (BMJ) has recently launched a patient advisory panel to guide the implementation of partnerships between the journal and patients and which could intersect with a lay peer-review process of economic evaluations and NMAs informing policy decisions (BMJ, 2014). Other journals (e.g. Annals of Internal Medicine, Annals of Rheumatology) provide patient-targeted summaries of important research articles they publish.

In general, the long-term success and adoption of these tools will require the identification of partnerships that could support the toolkit such as journals, academic collaborations and/or HTA organizations. Early engagement with these groups and a focus on building these partnerships during the initial stages of toolkit development will contribute to greater buy-in later during the implementation and dissemination of the toolkit (Guise et al., 2013; Straus et al., 2013).

Toolkit Evaluation

The proposed toolkit was guided by initial steps of the KTA framework but additional research and steps in the action cycle remain. Key steps that remain to be completed include the implementation, monitoring, evaluation and sustainability of the toolkit. This will require broader engagement of relevant stakeholders such as policymakers, researchers and partner organizations (Guise et al., 2013; Straus et al., 2013). Evaluation of the toolkit has a number of components, however the immediate next step is to evaluate the usability of each tool as its development is completed.

Assessing the usability of tools prior to their large-scale evaluation and implementation is important, in order to identify any issues with the tools that can be resolved proactively (Straus et al., 2013). For example, concerns with clarity, completeness, readability and ease of use may be identified at this stage and can be addressed prior to widespread distribution. Specific questions may focus on feasibility, if and how participants would use the tool, and suggestions for improvement (Brewer et al., 2012) Both researchers, i.e., knowledge producers, and policymakers, i.e. knowledge users, should be engaged in the usability testing of the Guidance for

Researchers. Usability of the tools may require consultation on the acceptability of the content (Yost et al., 2014) the language that is used (Akl et al., 2012) and graphical design features (Tan et al., 2014). Usability studies can be conducted following a variety of approaches. Questionnaires have been developed to evaluate usability of decision-support tools and will be modified for the purpose of evaluating these toolkit components. In addition, analyses collecting observations on how tools are used in practice, semi-structured interviews with end-users and journal reflections are approaches to evaluating the usability and acceptability of tools that could be applied to this toolkit (Yost et al., 2014). Deliberative dialogues with researchers and policymakers could also be leveraged to gain feedback on the usability of the tool (Boyoko et al., 2012; Bombard et al., 2011). Engaging individuals who participated in key informant interviews as part of this thesis, in addition to broader research and policymaker involvement, will also provide some verification and confirmation of how findings from key informant interviews were interpreted. The Toolkit Advisory Panel will be consulted on the results of the usability evaluation. If changes to the tools or potential barriers to their use are identified, potential solutions will be proposed and discussed. Similar to the initial development of the draft tools, if needed, revisions will be prioritized by the advisory panel based on (1) feasibility (2) appropriateness to the drug funding context (3) potential value (4) importance (5) novelty and lack of duplication with existing tools (Whitlock et al., 2010). The tools will be revised based on recommendations made by the advisory panel. Depending on the amount of feedback received during usability testing and the extent of changes required, this process may be repeated until the usability of the toolkit is satisfactory enough for wide-scale evaluation.

Systematic reviews identifying approaches to communicating economic evaluations and NMAs (Chapter 3 and 5) revealed that even when such approaches exist, their impact was rarely evaluated. Therefore, upon development, ensuring that the impact of the toolkit and/or its components, is robustly assessed will be an important contribution both to policy environments and the academic literature (Grimshaw 2012). Randomized controlled trials evaluating the impact of knowledge tools have been conducted in other fields (e.g. Akl et al., 2012; Brewer et al., 2012, Dobbins et al., 2009). The wide-scale evaluation of the toolkit will assess whether or not the toolkit and its components contribute to greater transparency and accessibility of economic evaluations and NMAs. This may require the selection of different outcomes for different toolkit components. However, each tool's effect on conceptual knowledge use, instrumental knowledge use and knowledge impact should be evaluated (Straus et al., 2013). For example, when considering the Guidance for Researchers tool:

- conceptual knowledge could be measured as researchers' awareness of the principles for conducting economic evaluations and NMAs that are outlined in the Guidance for Researchers or as policymaker's understanding of economic and NMA concepts;

- instrumental knowledge could be measured by assessing whether or not economic evaluations and NMAs produced by researchers meet the principles outlined in the Guidance for Researchers tool (e.g. greater clarity, identifying relevant contextual factors) or as policymakers ability to apply the results of economic evaluations and NMAs to policy decisions
- impact on they health policy and health systems could be measured as perceived transparency and accessibility of economic evaluations and NMAs by key stakeholders such as healthcare providers, patients and the public as well as perceived legitimacy and confidence in health policy decisions informed by economic evaluations and NMAs.

Whether or not to evaluate the full toolkit compared with individual or customized components of the toolkit should also be considered. While there are 12 components to the toolkit, implementation approaches should allow for flexibility in adopting some or all of the components and customizing them to local settings and organizations as needed. While economic evaluations are widely used, NMAs are not used as frequently and priorities for implementing these components of the toolkit may not be as high, as was seen in this thesis when the importance and value of different toolkit components was ranked.

The optimal implementation of the toolkit in research and policy settings may require additional strategies such as the identification of opinion leaders and local champions in partner organizations as well as targeted dissemination approaches (Straus et al., 2013; McCormack et al., 2013). Research suggests that many well-developed tools fail at the implementation stage if the context of organizational processes and structures are not fully considered. Ensuring engagement and buy-in of partners at this stage will be critical to the successful implementation of the toolkit. While this toolkit was developed as an independent research undertaking, identifying partnerships that could support the toolkit and interested end-users such as journals, academic collaborations and/or HTA organizations would contribute to the sustainability and success of the toolkit. Ongoing monitoring of how the toolkit is used following its implementation will also provide an opportunity for feedback, revisions and continuous quality improvement of the toolkit. This is also essential for tools used in policy environments where governments, priorities and policies may rapidly change; ensuring the toolkit remains relevant over time and can flexibly adapt to these pressures will be critical to its success.

Monitoring how the tools are used and factors that influence their use in pilot projects with select journals or HTA organizations could lead to development of a long-term sustainability plan. Sustainability is defined as “the degree to which an innovation continues to be used after initial efforts to secure adoption is completed”. (Davies & Edwards, 2013). Ensuring the sustainability of

the economic evaluation and NMA toolkit over time will require a more detailed understanding of the organizations adopting the toolkit. A sustainability plan should consider six important factors including consensus on what key toolkit principles or elements should be sustained over time; identification of systems to monitor toolkit use over time; an assessment of the adaptability of the toolkit over time in a rapidly-changing policy environment and factors that could influence its adaptability to other organizations; and the required leadership, funding and stakeholder support to ensure toolkit sustainability (Davies & Edwards, 2013).

Overall Recommendations

Recommendations for development of a relevant future research program are outlined below, representing key issues that were raised by this thesis but that were outside the scope of what could be feasibly addressed at this time.

1) **Toolkit Development**: As described above, further development of the outstanding components of the toolkit, is required including the policymaker Q&A documents; technical analysis templates incorporating perspectives of non-technical end-users; learning modules with an applied focus for policymakers; tool repositories for existing resources related to NMAs and economic evaluations; and lay peer-review processes for analyses targeting policymakers. The policymaker Q&A documents were prioritized next in the development of the toolkit. Successful development of these remaining components will also require engaging perspectives of the relevant stakeholders and incorporating insights from other collaboratives (e.g. ISPOR, Cochrane Collaboration) and relevant fields such as continuing medical education.
2) **Evaluation and Implementation of the Toolkit**: As described above, usability testing of the Guidance for Researchers should be conducted prior to the widespread implementation of this tool. Also, a robust evaluation of the impact of the toolkit, or toolkit components, should be conducted in a randomized controlled trial. Many tools are often developed but never evaluated as was observed in this thesis, therefore, evaluating the toolkit will be an important contribution to the KT literature on economic evaluations and NMAs. Successful implementation of the toolkit also requires building or partnering with existing collaborations, as well as partnerships with relevant end-users.
3) **Adapting Tools to Clinicians, Patients and the Public**: The current focus of this toolkit was on policymakers, as they are primary end-users of HTAs and substantial evidence exists on their use of evidence and how it may apply in the fields of economic evaluations and NMAs. However, clinicians, patients and the public are also key stakeholders who are impacted by health policy decisions and access to treatments. In the current policy

environment and with the increasing engagement of patients, expanding this evidence base and toolkit to these populations would be valuable and innovative research. This will require substantive original research identifying barriers and facilitators these stakeholders experience with respect to using and understanding economic evaluations and NMAs. This will allow tailoring of tools to the needs of clinicians, patients and the public as well as consideration of the local context in which these stakeholders use evidence, which may be very different from the policymakers' context.

4) **Capacity Building in Health Economics and NMAs**: The limited expertise to conduct economic evaluations and NMAs was frequently raised during this thesis. Initiatives to increase knowledge of economics and NMAs among existing and invested stakeholders and developing new researchers in these areas are much needed. However, enhancing the interface between knowledge producers, i.e. researchers, and knowledge users, i.e. policymakers, is an important and sometimes-neglected approach that could increase capacity and that deserves attention. Joint researcher-policymaker deliberative dialogues on the use of economic evaluations and NMAs to inform health policy could provide insights on these interactions and build capacity.
5) **Exploring the Impact of Health Systems on Economic Evaluations and NMAs**: Economic evaluations and NMAs are used in a complex policy environment both in Canada and internationally. Pressures that are outside the scope of these fields often impact on the use of economic evaluations and NMAs (e.g., market authorization requirements, drug pricing transparency and practices, budget silos, provincial variability in healthcare and related policies) but should be considered and explored in parallel as part of the effort to increase accessibility, impact and legitimacy of cost-effectiveness and comparative effectiveness research (i.e. economic evaluations and NMAs) in health policy decision-making.
6) **Confirming NMA Standards**: NMA methods and research standards are still evolving and a comprehensive KT strategy, while it can be developed in parallel with this research, can only be finalized when the methodology becomes more widespread and researchers have greater comfort and familiarity with conducting NMAs. Incorporation of ongoing NMA research such as the development of reporting standards and critical appraisal standards will be essential to optimizing the communication of NMAs to policymakers and other relevant stakeholders.
7) **Framework for Integrating New Methods in Health Policy Decision-Making**. Developing a framework for how to integrate valuable but novel methodologies into health policy decision-making (e.g. novel approaches for evaluating rare diseases) would be of benefit, building on experiences with how economic evaluations were established, how they can still be optimized and refined for relevant stakeholders and current experiences with the integration of NMAs into policy decision-making. Some principles common to both economic evaluations and NMAs emerged in this thesis and applications to principles for the diffusion of

innovations and other frameworks for implementing evidence such as the PARIHS framework and the Behaviour Change Wheel could be explored (Michie et al., 2011; Rogers 2003; Rycroft-Malone & Bucknall, 2010).

Conclusions

Economic evaluations and NMAs are frequently included as components of HTAs that subsequently inform health policy decisions. Ensuring the transparency of policy decisions and the evidence that informs them is essential for the perceived legitimacy and obtaining confidence in health policy decision-making processes. However, transparency requires providing, not only access to information, but providing information in an accessible and understandable format to audiences unfamiliar with complex economic and statistical approaches. In the current resource-constrained and budget conscious environment, enhancing the use and impact of these analytic techniques for determining comparative effectiveness and cost-effectiveness of therapies will have a valuable impact.

This thesis makes an important contribution to the literature by providing a robust evidence-base for the development of tools and approaches to enhancing communication of economic evaluations and NMAs. After integrating findings from four systematic reviews and two sets of key informant interviews with Canadian policymakers and senior individuals form HTA organizations around the world, three key components of effective knowledge tools for economic evaluations and NMAs were identified: content that clearly describes the available evidence, potential limitations and uncertainty in results; identification and consideration of local and real-world contextual factors that inform the results and conclusions of analysis making them more relevant for implementation; and presentation approaches that provide clarity, emphasize key messages and layer information so that the most important information is first and details are provided later.

Based on these findings, a 12-component toolkit was proposed that contains six tools or approaches, each applied to economic evaluations and NMAs: guidance for researchers, policymaker Q&A documents, technical analysis templates, learning modules, a tool repository and a lay peer-review process. Two of these tools, Economic Guidance for Researchers and NMA Guidance for Researchers, were developed and applied to an economic evaluation and NMA of osteoporosis therapies. Usability and acceptability of the Guidance for Researchers will be evaluated with researchers and policymakers before its implementation. Development of additional toolkit components is ongoing. When developed this toolkit has to potential to enhance the relevance, transparency and accessibility of economic evaluations and NMAs, and thereby increase confidence in the legitimacy of health policies and decisions informed by these analyses.

Finally, in the complex environment of Canadian healthcare and policy, there are numerous systems interactions and parallel areas of research that were identified in recommendations stemming from this thesis. These included the further exploration of: toolkits adapted to clinician and patient audiences, capacity building initiatives, related health systems pressures, NMA standards, and a framework for the integration of novel methodologies in health policy. Therefore, this thesis has the potential to influence and inform future research in each of these areas.

9.4 REFERENCES

Abdelhamid A, Loke Y, Parekh-Bhurke S, Chen Y, Sutton A, Eastwood A, Hollanda R, Song F. Use of indirect comparison methods in systematic reviews: a survey of Cochrane review authors. Research Synthesis Methods 2012, 3: 71-79.

Akl EA, Guyatt GH, Irani J, Feldstein D, Wasi P, Shaw E, Shaneyfelt T, Levine M, Schünemann HJ. "Might" or "suggest"? No wording approach was clearly superior in conveying the strength of recommendation. J Clin Epidemiol 2012; 65: 268-275.

Baghbanian A., Hughes I., Khavarpour F.A. Resource allocation and economic evaluation in Australia's healthcare system. Australian Health Review 2011;35: 278-283.

British Medical Journal (BMJ). BMJ Patient-Partnership. Available at: http://www.bmj.com/campaign/patient-partnership and http://www.bmj.com/about-bmj/resources-reviewers/guidance-patient-reviewers. Last accessed: August 2014.

Bombard Y, Abelson J, Simeonov D, Gauvin FP. Eliciting ethical and social values in health technology assessment: A participatory approach. Soc Sci Med. 2011, 73:135-44.

Brewer N, Gilkey M, Lillie S, Hesse B, Sheridan S. Tables or Bar Graphs? Presenting Test Results in Electronic Medical Records. Med Decis Making 2012, 32:545–553

CADTH. Rx for Change database. Available at: http://www.cadth.ca/en/resources/rx-for-change/database. Last accessed: August 2014.

Caro J, Eddy DM, Kan H et al. Questionnaire to assess relevance and credibility of modeling studies for informing health care decision making: an ISPOR-AMCP-NCP Good Practice Task Force report. Value in Health 2014, 17: 174-182.

Cavazzaa M & Jommi C. Stakeholders involvement by HTA Organisations: Why is so different? Health Policy 105 (2012) 236–245.

Chaimani A, Higgins JPT, Mavridis D, Spyridonos P, Salanti G. Graphical Tools for Network Meta-Analysis in STATA. PLoS ONE, 2013, 8: e76654.

Chase D, Rosten C, Turner S, Hicks N, Milne R. Development of a toolkit and glossary to aid in the adaptation of health technology assessment (HTA) reports for use in different contexts. *Health Technol Assess* 2009;**13**(59).

CIHR. Evidence-Informed Health Systems Renewal Portal. Available at: www.healthsystemsevidence.org. Last accessed: August 2014.

CIHR Guide to Researcher and Knowledge User Collaboration in Health Research.

Choi B, Pang T, Lin V, Puska P, Sherman G, Goddard M et al. Can scientists and policy makers work together? J Epidemiol Community Health 2005; 59:632–637.

Cipriani A, Higgins JP, Geddes JR, Salanti G. Conceptual and technical challenges in network meta-analysis. Ann Intern Med 2013, 159 (2): 130-137.

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.

Cook DA & Dupras DM. A Practical Guide To Developing Effective Web-based Learning. J Gen Intern Med 2004, 19:698–707.

Cook DA, Levinson AJ, Garside S, Dupras DM, Erwin PJ, Montori VM. Internet-Based Learning in the Health Professions. A Meta-analysis. JAMA. 2008;300(10):1181-1196.

Daniels N & Sabin J. Limits to health care: Fair procedures, democratic deliberations and the legitimacy problem for insurers. Philosophy & Public Affairs 1997; 303-350.

Davies, B., & Edwards, N. (2013). Sustaining knowledge use. In S. Straus, J. Tetroe & I. D. Graham (Eds.), *Knowledge translation in health care: Moving from evidence to practice* (pp. 237-248). Chichester: Wiley Blackwell BMJ Books.

Daw JR & Morgan S. Stitching the gaps in the Canadian public drug coverage patchwork? A review of provincial pharmacare policy changes from 2000 to 2010. Health Policy 2012, 104: 19–26.

DECIDE Collaboration, 2013.Available at: http://www.decide-collaboration.eu. Last accessed: August 2014.

de Pouvourville G, Ulmann, P, Nixon J, Boulenger S, Glanville J, Drummond M. The Diffusion of Health Economics Knowledge in Europe. The EURONHEED (European Network of Health Economics Evaluation Database) Project. Pharmacoeconomics 2005; 23 (2): 113-120.

Dobbins M, Hanna S, Ciliska D, Manske S, Cameron R, Mercer S et al. A randomized controlled trial evaluating the impact of knowledge translation and exchange strategies. Implementation Science, 2009, 4:61.

Drummond M. Twenty years of using economic evaluations for drug reimbursement decisions: what has been achieved? J Health Polit Policy Law. 2013;38(6): 1081-102.

Fattal J & Lehoux P. Health technology assessment use and dissemination by patient and consumer groups: Why and how? International Journal of Technology Assessment in Health Care, 2008, 24: 473–480.

Gagnon MP, Desmartis M, Lepage-Savary D, Gagnon J, St-Pierre M, Rhainds M et al. Introducing patients' and the public's perspectives to health technology assessment: A systematic review of international experiences. International Journal of Technology Assessment in Health Care, 2011, 27: 31-42.

Galani C and Rutten FFH (2008). Self-reported health care decision makers' attitudes towards economic evaluations of medical technologies Cur Med Res Op 24: 3049-3058.

Goeree R, He J, O'Reilly D, Tarride J, Xie F, Lim M et al. Transferability of health technology assessments and economic evaluations: a systematic review of approaches for assessment and application. ClinicoEconomics and Outcomes Research 2011:3 89–104.

Graham ID, Logan J, Harrison MB, Straus SE, Tetroe J, Casell W et al. Lost in knowledge translation: time for a map: J Contin Ed Health Prof 2006; 26(1):13-24.

Greenhalgh T, Robert G, Bate P, Kyriakidou O, Macfarlane F, Peacock R. How to spread good ideas. A systematic review of the literature on diffusion, dissemination and sustainability of innovations in health service delivery and organisation. London: Report for the National Co-ordinating Centre for NHS Service Delivery and Organisation R&D (NCCSDO); 2004.

Greyson DL, Cunningham C & Morgan S. Information behaviour of Canadian pharmaceutical policy makers. Health Information and Libraries Journal, 2011, 29: 16–27.

Guise JM, O'Haired C, McPheeters M, Most C, LaBrant L, Lee K, Cottrella E, Graham E. A practice-based tool for engaging stakeholders in future research: a synthesis of current practices. Journal of Clinical Epidemiology 2013, 66, 666-674.

Hildon Z, Allwood D, Black N (2012). Impact of format and content of visual display of data on comprehension, choice and preference: a systematic review. Int J Qual Health Care, 24 (1): 55-64.

Hoffman C, Stoykova BA, Nixon J, et al. Do health care decision makes find economic evaluations useful? The findings of focus group research in UK health authorities. Value in Health 2002;5: 71-79.

Hyder A, Corluka A, Winch P, El-Shinnawy A, Ghassany H, Malekafzali H. National policy-makers speak out: are researchers giving them what they need? Health Policy and Planning 2011;26:73–82.

International Society for Pharmacoeconomics and Outcomes Research. Pharmacoeconomic (ISPOR) guidelines around the world. Available from: http://www.ispor.org/PEguidelines/index.asp [Accessed May 19, 2014]

Jansen JP, Trikalinos T, Cappelleri JC, Daw J, Andes S, Eldessouki R et al. Indirect treatment comparison/network meta-analysis study questionnaire to assess relevance and credibility to inform health care decision making: an ISPOR-AMCP-NPC Good Practice Task Force report. Value in Health 2014, 17(2): 157-173.

Laupacis A. Incorporating Economic Evaluations Into Decision-Making. The Ontario Experience. Med Care 2005;43: II-15–II-19.

Lavis, J., Davies, H., Oxman, A., Denis, J. L., Golden-Biddle, K., & Ferlie, E. Towards systematic reviews that inform health care management and policy-making. Journal of Health Services Research and Policy, 2005, 10 (Suppl 1), 35-48.

Lavis JN, Oxman AD, Lewin S, Fretheim A: SUPPORT Tools for evidence-informed health Policymaking (STP). Introduction. Health Research Policy and Systems; 2009, 7(Suppl 1):I1.

Leonard DT (2010). Translating and communicating patient centered research to the public. *Am J Pharm Benefits* **2:** 408-410.

Madden L, King L, Shiell A. How do government health departments in Australia access health economics advice to inform decisions for health? A survey. Aust New Zealand Health Policy. 2009, 6:6.

McCormack L, Sheridan S, Lewis M, Boudewyns V, Melvin CL, Kistler C, Lux LJ, Cullen K, Lohr KN. Communication and Dissemination Strategies To Facilitate the Use of Health-Related Evidence. Evidence Report/Technology Assessment No. 213. (Prepared by the RTI International–University of North Carolina Evidence-based Practice Center under Contract No. 290-2007-10056-I.) AHRQ Publication No. 13(14)-E003-EF. Rockville, MD: Agency for Healthcare Research and Quality; November 2013.

Michie S, van Stralen MM, West R. The behaviour change wheel: A new method for characterising and designing behaviour change interventions. Implementation Science 2011, 6:42.

Miles M & Huberman A. *Qualitative data analysis. An expanded sourcebook*, 3rd edn. Thousand Oaks, CA. Sage Publications Inc., 2013.

Mills E, Thorlund K, Ioannidis J. Demystifying trial networks and network meta-analysis. BMJ 2013, 346: f2914.

Moher D, Weeks L, Ocampo M, Seely D, Sampson M, Altman DG et al. Describing reporting guidelines for health research: a systematic review. J Clin Epidemiol 2011, 64: 718-742.

Morgan S, Thomson P, Daw J, Friesen M. Inter-jurisdictional cooperation on pharmaceutical product listing agreements: views from Canadian provinces. BMC Health Services Research 2013; 13:34

Murthy L, Shepperd S, Clarke MJ, Garner SE, Lavis JN, Perrier L, Roberts NW, Straus SE. Interventions to improve the use of systematic reviews in decision-making by health system managers, policy makers and clinicians. Cochrane Database of Systematic Reviews 2012, Issue 9. Art. No.: CD009401. DOI: 10.1002/14651858.CD009401.pub2

National Health Services. NHS economic evaluation database. Available at: http://www.crd.york.ac.uk/CRDWeb/. Last accessed: August 2014.

Niezen M, Stolk E, Steenhoek A, Uyl-De Groot C. Inequalities in oncology care: Economic consequences of high cost drugs. Eur J of Cancer, 2006, 42: 2887–2892.

Rogers EM. Diffusion of Innovations, 5th edition. New York: Free Press, 2003.

Rosenbaum L & Lamas D. Cents and Sensitivity. Teaching Physicians to Think about Costs. N Engl J Med, 2012, 367: 99-101.

Rycroft-Malone J &Bucknall T. Models and Frameworks for Implementing Evidence-Based Practice: Linking Evidence to Action. Wiley-Blackwell: West Sussex, UK. 2010.

Salanti G. Indirect and mixed-treatment comparison, network, or multiple-treatments meta-analysis. Many names, many benefits, many concerns for the next generation evidence synthesis tool. Research Synthesis Methods 2012, 3: 80-97.

Schünemann HJ, Best D, Vist G, Oxman AD, for the GRADE Working Group. Letters, numbers, symbols and words: how to communicate grades of evidence and recommendations. CMAJ, 2003, 169: 677-680.

Straus S, Tetroe J & Graham ID (Eds.), Knowledge translation in health care: Moving from evidence to practice, 2nd edition. Oxford: Blackwell Publishing, 2013.

Straus, S. E., Tetroe, J., Bhattacharyya, O., Zwarenstein, M., & Graham, I D., (2013). Monitoring knowledge use and evaluating outcomes. In S. Straus, J. Tetroe & I. D. Graham (Eds.), *Knowledge translation in health care: Moving from evidence to practice* (pp. 227-236). Chichester: Wiley Blackwell BMJ Books.

Sullivan R, Peppercorn J, Sikora K, Zalcberg J, Meropol N, Amir E. Delivering affordable cancer care in high-income countries. Lancet Oncology, 12: 933-980.

Tan SH, Cooper NJ, Bujkiewicz S, Welton NJ, Caldwell DM, Sutton AJ. Novel presentational approaches were developed for reporting network meta-analysis. J Clin Epidemol 2014, 67(6): 672-680.

Thabane L, Wells G, Cook R, Platt R, Pogue J, Pullenayegum E, et al. Canadian-led capacity-building in biostatistics and methodology in cardiovascular and diabetes trials: the CANNeCTIN Biostatistics and Methodological Innovation Working Group. Trials 2011, 12:48.

Thurston SJ, Craig D, Wilson P et al. Increasing decision makers' access to economic evaluations: alternative methods of communicating the information. Intl J Technol Assess Health Care 2008;24: 151-157.

Trevena L, Davey H, Barratt A, Butow P, Caldwell P. A systematic review on communicating with patients about evidence. J Evalu Clin Practice, 2006, 12: 13-23.

Wensing, M., Bosch, M., Grol R. (2013). Developing and selecting knowledge translation interventions. In S. Straus, J. Tetroe & I. D. Graham (Eds.), *Knowledge translation in health care: Moving from evidence to practice*. Chichester: Wiley Blackwell BMJ Books.

Williams I, McIver S, Moore D, Bryan S. The use of economic evaluations in NHS decision-making: a review and empirical investigation. Health Technol Assess 2008; 12: 1-175.

Yost J, Dobbins M, Traynor R, DeCorby K, Workentine S, Greco L et al. Tools to support evidence-informed public health decision making.

APPENDIX A.

ABSTRACT FOR ORAL PRESENTATION, COCHRANE COLLOQUIUM 2012

Understanding barriers and approaches to communicating health economic evidence: systematic review and needs assessment

Authors: Sullivan S[1 2], Akbar Z[3], Bell C[4], Sabharwal M[2]

[1] University of Ottawa, Canada [2] pan-Canadian Oncology Drug Review, Canada
[3] McMaster University, Canada [4] University of Toronto, Canada

Background: Economic evaluations are a key component of evidence-informed decision-making. However, gaps exist in understanding economic evaluations. Improved methods of communicating economic evaluations may increase understanding of health economic evidence and enhance its role in policy decision-making.

Objectives: To identify barriers and requirements for communicating health economic evidence to a non-technical audience.

Methods: (1) A systematic review was conducted identifying ways of using economic evidence; barriers to using economic evidence; comprehension of economic evidence and methods of communicating economic evidence. All studies with a defined methodology and reporting on methods, interventions or tools conveying economic evidence were included. Cochrane systematic review methodology was followed and risk of bias assessed. (2) A needs assessment was also conducted using a small purposive sample of the general public with exposure to economic evidence and health policy decision-making. Key informant interviews (N=7) identified challenges in understanding health economic evidence.

Results: Twenty-two studies were included. Most studies were surveys and interviews (semi-structured or key informant). Common study limitations were selection and detection bias. Despite heterogeneity of studies, conclusions were consistent across studies. Economic evidence was used to determine public health costs and resource allocation; support drug funding decisions; and inform cost-effectiveness analyses. Common barriers to using economic evidence were technical language and challenges transferring to the local context. Comprehension of economic evidence was low among all audiences. Improving communication should focus on clearly presenting cost-effectiveness analyses with the appropriate amount and level of detail. These findings were corroborated by the needs assessment. Key informants confirmed they have challenges understanding economic evaluations and indicated economic information should always be presented in the simplest, non-technical manner possible.

Conclusion: Effective communication of economic evidence requires concepts to be simplified and contextual examples provided. Findings will inform development of tools that enhance communication and understanding of economic evaluations.

Primary Theme: Knowledge Translation and Communicating Evidence

APPENDIX B.

ETHICS APPROVAL



Ottawa Hospital
Research Institute
Institut de recherche
de l'Hôpital d'Ottawa



uOttawa



UNIVERSITY OF OTTAWA
HEART INSTITUTE
INSTITUT DE CARDIOLOGIE
DE L'UNIVERSITÉ D'OTTAWA

***Ottawa Health Science Network Research Ethics Board*/ Réseau des sciences de la santé d'Ottawa Conseil d'éthique de la recherche**

Civic Box 411 725 Parkdale Avenue, Ottawa, Ontario K1Y 4E9 613-798-5555 ext. 14902 Fax: 613-761-4311
http://www.ohri.ca/ohsn-reb

Dear Dr. Wells:

Re: Protocol #	**20130714-01H**	**Survey of Drug Funding Decision-Makers to Determine Informational Needs for Communicating and Understanding Network Meta-Analyses and Economic Evaluations**
Protocol approval valid until -		**Wednesday, February 25, 2015**

I am pleased to inform you that this protocol underwent expedited review by the Ottawa Health Science Network Research Ethics Board (OHSN-REB) and is approved. No changes, amendments or addenda may be made to the protocol or the consent form without the OHSN-REB's review and approval.

Approval has been granted for the following:

- Exemption for French translation
- English Protocol, Version 1, dated November 11, 2013
- English Briefing Document Survey , Version 1, dated August 23, 2013
- Key Information Email, Version 1, dated December 30, 2013
- Email Recrutiment
- Decision- Maker Interview Questions, dated August 23, 2013
- Interview Script and Prompts for Decision Maker Survey, Version 1, dated August 23, 2013

If the study is to continue beyond the expiry date noted above, a Renewal Form should be submitted to the REB approximately six weeks prior to the current expiry date. If the study has been completed by this date, a Termination Report should be submitted.

The Ottawa Health Science Network Research Ethics Board (OHSN-REB) was created by the merger of both the Ottawa Hospital Research Ethics Board (OHREB) and the Human Research Ethics Board (HREB) for meetings held at the University of Ottawa Heart Institute.

OHSN-REB complies with the membership requirements and operates in compliance with the Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans; the International Conference on Harmonization - Good Clinical Practice: Consolidated Guideline; the provisions of the Personal Health Information Protection Act 2004.

APPENDIX C.

INTERVIEW GUIDES FOR KEY INFORMANT INTERVIEWS

INTERVIEW GUIDE FOR CANADIAN KEY INFORMANTS

Communication and Understanding of

Economic Evaluations and Indirect Comparisons

Economic evaluations and indirect comparisons are often used to inform drug funding decisions. These can be complex methodological approaches that may be **challenging to communicate and understand**. We would like to find out more about experiences with these approaches in your drug funding process in order to **develop better ways of sharing this information**.

As a key informant, you are one of the individuals selected to provide views **on behalf of the drug funding process in your province**. Please answer the following questions taking into consideration the needs of your drug funding process, **not just your own personal views and experiences**. There are no right or wrong answers. You may want to **consider recent example(s)** of drugs when providing your responses.

Thank you very much for your time. If you have any further comments or suggestions, please don't hesitate to contact me (Shannon Sullivan, email address)

PART 1 – BACKGROUND INFORMATION

Please provide some background information by checking the applicable category in the questions below:

1. What is your role in the drug funding process?

   ☐Management or Director
   ☐Expert Committee Member
   ☐Consultant or Other_____

2. How many years of experience do you have related to the drug funding context:

   ☐less than 5 years
   ☐5 to10 years
   ☐more than 10 years

3. If applicable, please describe your drug plan. Are there specific factors that influence the use of evidence in your drug funding process (e.g. organizational structure or process considerations, availability of resources, expertise, budgets)?

PART 2 - ECONOMIC EVALUATION QUESTIONS

1. When thinking about the information you receive on economic evaluations:

    a. What type of information related to economic evaluations of cost-effectiveness is used to inform a drug funding decision in your process?

    b. How or when do you use information from economic evaluations of cost-effectiveness to inform a drug funding decision in your process?

2. When thinking about the people in your drug funding process who use the economic evaluations:

    a. How well are individuals able to interpret the economic information they are provided?

    b. What key elements, if any, would individuals like to understand better?

3. Are there any barriers related to using and understanding economic evaluations you have observed in your drug funding process?

4. Are there any approaches that facilitate (or could facilitate) the use and understanding of economic evaluations in your drug funding process and make them more useful to you?

PART 3 - INDIRECT COMPARISON QUESTIONS

1. When thinking about the information you receive on indirect comparisons:

    a. What type of information related to indirect comparisons is used to inform a drug funding decision in your process?

    b. How and when do you use evidence from indirect comparisons to inform a drug funding decision in your process?

2. When thinking about the people in your drug funding process who use the indirect comparisons:

    a. How well are individuals able to interpret information on indirect comparisons they are provided?

    b. What key elements, if any, would individuals like to understand better?

3. Are there any barriers related to using and understanding indirect comparisons you have observed in your drug funding process?

4. Are there any approaches that facilitate (or could facilitate) the use and understanding of indirect comparisons in your drug funding process and make them more useful to you?

INTERVIEW GUIDE FOR INTERNATIONAL KEY INFORMANTS

Tools and Approaches to Improve Understanding and Enhance Communication of Economic Evaluations and Indirect Comparisons

Why is This Research Being Conducted?

- Economic evaluations and indirect comparisons are often used as part of a health technology assessment (HTA) to inform drug funding decisions.
- These can be complex methodological approaches that may be challenging for a general audience to understand.
- Increasingly, stakeholders such as the public, patients, policy makers and clinicians are seeking to better understand how drug funding decisions are made.
- We would like to find out more about any tools or strategies used to communicate economic evaluations and indirect comparisons to help stakeholders better understand them
- Understanding how HTA organizations around the world approach these issues will facilitate the development of new approaches that consider existing challenges and build on current best practices.

Who Will Participate and Be Interviewed?

- HTA organizations around the world that make drug funding decisions or provide advice to decision-makers on drug funding in a publicly-funded healthcare system
- Key contacts in the organizations are asked to identify participants to interview
    - Interviewees should be individuals who understand how indirect comparisons and economic evaluations are shared and communicated with others.
    - In some cases, identifying more than one person in the organization could be appropriate.
    - People with different responsibilities or experience can be included. For example, senior individuals in the organization who:
        - create and share this evidence with an expert decision-making committee; or
        - help implement decisions by working with external stakeholder groups; or
        - develop knowledge translation tools or strategies for your organization.

How Will the Interview Be Conducted?

- Approximately 30 minute interview via teleconference in May-June 2014
- A copy of interview questions provided in advance

- Includes background questions on organization and knowledge translation strategy
- Questions are about how indirect comparisons and economic evaluations are communicated and presented to key stakeholders including:
  - Tools and approaches that are used
  - Target audience and stakeholders for tools
  - How tools and approaches have been developed
  - Challenges and successes in using these tools and approaches

What Will Be the Outcome of This Research?

- Improved approaches for communicating economic evaluations and indirect comparisons to key stakeholders involved with HTA and drug funding
- High-level summary of results may be used in publications and presentations and shared with participating HTA organizations
- Upon request, findings will be shared with your organization for review and comment prior to being shared publicly

NOTE: Ethics approval has been obtained for this project through the Ottawa Health Science Network Research Ethics Board (Protocol # 20130714-01H)

PART 1 – KNOWLEDGE TRANSLATION APPROACHES

1. Which, if any, of the following approaches does your organization take to knowledge translation?

> **Note**: Knowledge translation can be defined as "a dynamic and iterative process that includes synthesis, dissemination, exchange and ethically-sound application of knowledge to improve health, provide more effective health services and products and strengthen the health care system." (Canadian Institutes of Health Research)

☐Part of the organization's mandate or strategic priorities
☐Have a unit or department dedicated to knowledge translation activities
☐Have dedicated staff for knowledge translation activities
☐Staff have specific responsibilities related to knowledge translation activities but are not dedicated to this alone
☐Use external groups or consultants to conduct knowledge translation activities
☐Knowledge translation activities target patients or public
☐Knowledge translation activities target policy makers
☐Knowledge translation activities target health care professionals
☐None of the above
☐
Other_____

PART 2 – APPROACHES TO COMMUNICATING ECONOMIC EVALUATIONS

2. (a) Does your organization use **economic evaluations** as part of a health technology assessment?

☐Yes
☐No

(b) If you responded **"YES"**, please describe how they are used and how long they have been used in your organization.

For example economic evaluations could be used to inform cost-effectiveness; to inform formulary listing or disinvestment decisions; by expert advisory committees; to support drug price negotiations; for education of stakeholders; outreach to stakeholders; to change policies or clinical practices; or in other ways.

3. (a) Does your organization have any **tools or approaches** to better communicate **economic evaluations** or to assist stakeholders in understanding how economic evaluations are conducted and how to use their findings?

☐Yes
☐No

Please Note:
- Examples of tools or approaches may include specific formats and templates for providing information, plain language versions of documents, implementation tools, educational sessions or workshops, providing access to expert resources, or other methods
- Stakeholders may include policy makers and administrators, expert committee members, the public, patients or patient groups, health care professionals, researchers or other groups with whom you interact.

(b) If you responded **"NO",** why not?

For example: because it is not part of your organization's mandate, your organization has limited expertise to develop tools, lack of financial resources, no interest or need for tools, etc.

(c) If you responded **"YES**", please describe some of these tools or approaches, including details such as:
- the type and content of the tool or approach
- how it is used
- the purpose
- the target audience
- how it was developed
- how long ago it was developed and amount of experience you have using it
- how it is shared with others

4. Please describe any **challenges** your organization may have had using these tools or approaches to improve understanding of **economic evaluations**.

5. Please describe any **successes** your organization has had using these tools or approaches to improve understanding of **economic evaluations**.

PART 3 – APPROACHES TO COMMUNICATING INDIRECT COMPARISONS

6. (a) Does your organization use **indirect comparisons** as part of a health technology assessment?

☐Yes

☐No

(b) If you responded **"YES"**, please describe how they are used and how long they have been used in your organization.

For example indirect comparisons could be used: to inform economic

evaluations; to inform drug class or therapeutic reviews; to inform clinical practice guidelines; to inform formulary listing or disinvestment decisions; by expert advisory committees; to support drug pricing negotiations; for education or outreach for stakeholders or to change clinical practices; or other purposes.

7. (a) Does your organization have any **tools or approaches** to better communicate **indirect comparisons** or to assist stakeholders in understanding how indirect comparisons are conducted and how to use their findings?

☐Yes

☐No

Please Note:

- Examples of tools or approaches may include: specific formats and templates for providing information, plain language versions of documents, implementation tools, educational sessions or workshops, providing access to expert resources, or other methods
- Stakeholders may include: policy makers and administrators, expert committee members, the public patients or patient groups, clinicians and other health practitioners, researchers or other groups with whom you interact.

(b) If you responded **"NO",** why not?

For example: because it is not part of your organization's mandate, your organization has limited expertise to develop tools, lack of financial resources, no interest or need for tools etc.

(c) If you responded **"YES**", please describe some of these tools or approaches, including details such as:

- the type and content of the tool or approach
- how it is used
- the purpose
- the target audience
- how it was developed
- how long ago it was developed and amount of experience you have using it
- how it was shared with others

8. Please describe any **challenges** your organization has had using these tools or approaches to improve understanding of **indirect comparisons**.

9. Please describe any **successes** your organization has had using these tools or approaches to improve understanding of **indirect comparisons**.

PART 4 – BACKGROUND INFORMATION

Please provide some background information in the questions below:

10. (a) What is your organization's **role in health technology assessment** in your country or jurisdiction?

For example the role could be to conduct health technology assessments for the purpose of informing: procurement decisions; funding decisions, disinvestment decisions; practice guidelines; optimal use; or other purposes.

(b) What is your organization's **role in the drug funding process** in your country or jurisdiction?

For example the role could be to: make recommendations to policy makers on drug funding decisions; make policy decisions on drug funding; provide evidence-based reports but no recommendations or decisions; or other roles.

(c) How long has your organization been doing this type of work?

☐less than 5 years
☐5 to10 years
☐more than 10 years

THANK YOU
Thank you very much for your time. If you have any further comments or suggestions please contact me (Shannon Sullivan, email address)

APPENDIX D.

OSTEOPOROSIS ECONOMIC EVALUATION

Cost-Effectiveness of Pharmacological Treatments for Osteoporosis in Canada

Doug Coyle [1]
Shannon M. Sullivan [2]
George A. Wells [1, 2]

1. Department of Epidemiology and Community Medicine, University of Ottawa, Ottawa, ON., Canada
2. University of Ottawa Heart Institute, Ottawa, ON, Canada

Address for Correspondence:

Doug Coyle
Department of Epidemiology and Community Medicine,

University of Ottawa,

451 Smyth Road,

Ottawa, Ontario,

Canada K1H 8M5

Keywords: Cost-effectiveness; osteoporosis; bisphosphonates

Statement

This manuscript was developed without direct funding from any source. It is original work which has not been published elsewhere. The manuscript has been read and approved by all authors. The opinions expressed in the manuscript are those of the authors alone and are not representative of any organisation to which we may be affiliated.

ABSTRACT

Introduction: There is a lack of independent analyses concurrently comparing numerous pharmacotherapies for osteoporosis. The primary objective of this study was to identify the optimal osteoporosis treatment based on the women's age, bone mineral density (BMD) and fracture history.

Methods: A cost utility analysis was performed of six different treatment strategies for osteoporosis (alendronate, etidronate, risedronate, zoledronic acid, denosumab and no pharmacotherapy) using a Markov model incorporating data on the fracture risk, and the associated costs, mortality and disutility and the effect of treatment. The base case was a 75 year old osteoporotic women with no previous fracture and a BMD 2.5 standard deviations below the young female mean. Uncertainty was assessed through detailed probabilistic and deterministic sensitivity analyses.

Results: For the base case, the incremental cost per QALY gained (ICUR) was $10,895 for alendronate versus no pharmacotherapy and $1,609,172 for zoledronic acid versus alendronate. All other therapies were dominated by alendronate. For older populations, those with a previous fracture and those with lower BMD, alendronate dominated no therapy. For younger populations, the ICUR increased up to $98,347, Across different populations, the ICUR for zoledronic acid versus alendronate ranged from $224,826 to $4,591,999. Sensitivity analysis confirmed the robustness of the results.

Conclusions: For osteoporotic women with a previous fracture and/or aged over 70 and/or with a BMD at least 3.5 standard deviations below the young adult mean, alendronate is the optimal pharmacotherapy based on threshold value of $50,000 per QALY. For all other women, no pharmacotherapy is cost effective.

INTRODUCTION

Osteoporosis is a progressive bone disease characterized by low bone mass which increases bone fragility leading to an increase in fracture risk (1). Osteoporosis can be diagnosed by the presence of fragility fractures (2). For individuals without fracture, osteoporosis can be diagnosed based on the World Health Organization criteria: having a bone mineral d*e*nsity (BMD) which is at least 2.5 standard deviations (SD) below the mean peak bone mass of an average young female (1). The prevalence of osteoporosis in post menopausal women increases from approximately 6% in those aged 50-59 years to over 42% in those aged over 80 (3).

The most common fractures associated with osteoporosis are fractures of the hip, vertebrae or wrist. The major source of morbidity from osteoporosis arises from hip fractures which is also associated with higher costs and greater mortality (4,5). The incidence of osteoporosis increases with age as does the risk of fracture (3,6). Fracture risk is also related to the previous incidence of fracture and the degree of low bone mass.

A recent study estimated that the annual health care costs in 2007-08 associated with osteoporosis in Canada were $2.3 billion with an additional $1.6 billion for associated use of long term care facilities (5). The cost of treating hip fractures comprised more than half of the acute care costs associated with osteoporosis.

Pharmacological treatments for osteoporosis include bisphosphonates (alendronate, risedronate and etidronate, zoledronic acid) and the RANKL (receptor activator of nuclear factor kappa-B ligand) inhibitor; denosumab. Both classes of drugs attempt to maintain the balance between osteoblasts which create bone and osteoclasts which destroy bone (7). Postmenopausal osteoporosis occurs when there is imbalance between osteoblasts and osteoclasts leading to reduced bone mass (8). Bisphosphonates slow bone loss by causing the cell death of osteoclasts. Denosumab inhibits the protein which activates osteoclasts.

The annual costs of osteoporosis drug treatments vary considerably although such costs may by partly offset by reducing the economic burden of fracture. The objective of this study was to assess the cost effectiveness of the various pharmacotherapies and identify, which treatments are optimal, depending on a woman's age, history of fracture and BMD.

METHODS

Study Question

The principal form of analysis was a cost utility analysis, where outcomes are expressed in terms of Quality Adjusted Life Years (QALYs), which are derived from summing annual utility scores (43). In addition, cost-effectiveness analysis was conducted with effectiveness measured by both the number of fractures avoided and by change in life expectancy. Analysis is presented in terms of the incremental cost per life year (ICER) or per QALY (ICUR) gained.

The model adopts the perspective of the health and social care system in that the costs of health, social services and long term care (LTC) are included (9). For the base case analysis, costs and benefits were discounted at 5% per annum (9). Sensitivity analysis was conducted with discounting at 0 and 3 % (9).

The cost effectiveness of treatments will vary by women's age, severity of osteoporosis and fracture history. Base analysis is presented for a cohort of 75 year old osteoporotic women with a BMD 2.5 SD less than a young female and with no previous fracture. Further analysis is conducted varying the cohort's age groups (65, 70, 75, 80, 85 ,90), baseline BMD (2.5, 3, 3.5 and 4 SD less than a young female) and fracture history (no previous fracture and previous fracture).

Model Design

Analysis was conducted using a decision analytic model for osteoporosis developed based on the most recently available data relevant to the Canadian population. The model used is an update of a model employed in several previous analyses and follows guidelines both for the design of economic models in general and for the design of economic models specific to osteoporosis (9-13). The model is a Markov model with one year cycle length and a lifetime horizon. The model incorporates the sequelae associated with osteoporosis (e.g. fracture) and the transition of women through different states related to the development of osteoporosis, history of fracture and residential status (Figure 1).

The model is populated with relevant transition probabilities and estimates of the costs and utilities associated with each health state (Table 1) (3,5, 12, 14-33). Data sources were obtained through review of the available literature and focused on identifying the most recent relevant and appropriate data from a Canadian context (34). The primary form of analysis is deterministic with further stochastic analysis using Monte Carlo simulation (35). The model is developed within a Microsoft Excel workbook.

Transition Probabilities

Given the structure of the model outlined above, the following transition probabilities are required to allow a simulation of progression through the model:

- Probabilities of hip, wrist and spine fracture
- Probability of developing osteoporosis
- Probabilities of being admitted to LTC
- Probabilities of death

Each of these probabilities will vary by a number of factors. The probability of developing osteoporosis will increase with age (3). The probability of fracture increases with age, declining BMD, residence in LTC and previous history (14). Admission to LTC increases with age (15). The probability of death increases with age, incidence of fracture and residence in LTC (16-20). Specific data on probabilities by risk factors were often unavailable. However, data for alternative parameters were available which allowed computation of the necessary parameters through calibration of the model. For example, although non fracture related mortality by age for those living in the community (i.e. excluding mortality post fracture) is unavailable, age specific all-cause mortality and the relative risks of mortality associated with fracture and living in LTC is available. Thus, the model is calibrated such that the combination of mortality rates post fracture with and without fracture and with and without LTC residence replicates age specific all-cause mortality.

The prevalence of osteoporosis and the incidence of osteoporotic fractures by age group for the Canadian population was obtained from the Canadian Multicentre Osteoporosis Study (CaMos) a prospective cohort study involving a random sample of 10061 Canadians aged over 25 (3,14). Age-specific transition rates for admission to LTC were derived from the proportion of individuals living in LTC by age obtained from Canadian census data (15). To populate the model it was necessary to derive age specific probabilities for fracture based on BMD, fracture history and residential status. First, data was obtained on the relative risk of fracture by type for women with osteoporosis compared to no osteoporosis and the relative risk of fracture for women with osteoporosis and a BMD -2.5 SD below the age adjusted mean (21). The former is used to calibrate the model to provide a probability of fracture for women with osteoporosis, the latter is then used to allow estimation of fracture risks by BMD. To allow estimation of fracture probabilities by different degrees of bone loss, an estimate of the relative risk of fracture for each one standard deviation decrease (1.5) from a meta analysis of prospective cohort studies was

used (22). The relative risks of hip, wrist and spine fractures given a previous osteoporotic fracture was obtained from a systematic synthesis of the literature relating to the risk of future fracture, given a history of prior fracture (23). The relative risk of fracture given residence in LTC was obtained from an analysis of hospital discharge data from Ontario (24).

Age specific all-cause mortality rates for women were obtained from Statistics Canada (16). To calibrate the probability of death dependent on age, fracture history and residence, relative risks of death based on incident fractures and LTC residence were obtained from the literature (17-20).

Costs

For each particular state within the decision model, there is an associated estimate of costs (2013 Canadian dollars). The cost of the health care resources associated with the treatment of a hip fracture in Ontario including immediate acute care, rehabilitation and institutionalization were obtained from an analysis of healthcare administrative claims and billing data available at the Institute for Clinical Evaluative Sciences (ICES)(25). The cost of treating a women with vertebral and wrist fractures which do not require hospitalization were obtained from a previous Canadian economic study (12). The costs of treating a women with vertebral and wrist fractures which required hospitalization were obtained from an analysis of administrative data held by the Canadian Institute of Health Information (26). The proportions of women with these fractures requiring hospitalization was derived from the prevalence data of fracture from CaMos and a recent burden of illness study (5, 14). Sensitivity analysis was conducted assuming all fracture patients were hospitalized.

The costs of alendronate, etidronate, risedronate, denosumab and zoledronic acid were obtained from the Ontario Drug Benefit Formulary (27). The yearly cost of each product was obtained by summing the acquisition costs of the medication, an additional 8% markup, and the appropriate number of dispensing fees at a cost of $8.83 each (4 for alendronate, etidronate and risedronate, two for zoledronic acid and 1 for denosumab).

Utilities

Utility values for women with normal health status were obtained by age from the Canadian Community Health Survey (CCHS) (28). Disutilities associated with hip and wrist fractures and vertebral fractures not requiring hospitalization were derived from a recent systematic review of utility values for fractures (29). The disutility of a vertebral fracture not requiring hospitalization was obtained from the most recently available study (30).

Treatment Specific Parameters

The effect of alternative treatments on the risks of fractures were derived from data from a network meta-analysis of randomized controlled clinical trials (31). This analysis included ibandronate which is not used in Canada in the treatment of osteoporosis so was excluded from analysis. A previous economic analysis comparing denosumab to risedronate and alendronate had effectiveness data more favorable towards denosumab. Thus, a sensitivity analysis was conducted utilizing this data,

Treatment duration was assumed to be five years with sensitivity analysis adopting various lengths of treatment duration (2, 3 and 4 years).

There is evidence that patients experience continued reductions in the risk of fracture after stopping therapy. Thus, the model allows for a fracture set time whereby there is a linear reduction of benefit after the curtailment of therapy for all bisphosphonates up to a set period of time. Base analysis assumed a set time of 2 years (12,14, 36, 37).

Adherence was incorporated into the model by assuming that a proportion of patients would stop treatment within 12 months of commencement – those who continue would continue for the maximum duration of treatment specified in the analysis. It is necessary to incorporate differential adherence with pharmacotherapies based primarily on their frequency. Data on 12-month adherence to daily alendronate, etidronate and risedronate was based on data for Ontario from a previous economic study (12). Improved adherence for weekly bisphosphonates versus daily was modeled based on relative adherence rates from an analysis of administrative data (32). Improved adherence with denosumab versus weekly bisphosphonates was modeled based on data from an RCT (33). No data were available to measure long term adherence with zoledronic acid. In theory, adherence should be higher given the reduced frequency of dosing. Thus, for the base case the rate of discontinuations with zoledronic acid was assumed to be half that of denosumab with sensitivity analysis assuming the same adherence as denosumab and full adherence.

Sensitivity Analysis

Univariate sensitivity analysis was conducted to assess the robustness of the study's results to changing the main assumptions within the model, namely:

- Changing set time in the model (0 and 5 years)

- Changing duration of treatment (2, 3 and 4 years)
- Assuming efficacy data for risedronate, alendronate and denosumab as from a previous publication (14)
- Assuming all patients with vertebral and wrist fractures require hospitalization
- Assuming full adherence for all therapies
- Assuming equal adherence with denosumab and zoledronic acid
- Assuming full adherence for zoledronic acid
- Assuming full adherence for denosumab and zoledronic acid
- Assuming no increase in mortality due to vertebral fracture
- Assuming no increase in mortality due to hip or vertebral fracture
- Assuming no disutility from hip fracture after 1 year
- Adopting discount rates of 0 and 3%
- Assuming full adherence for denosumab and zoledronic acid, assuming all vertebral and wrist fractures required hospitalization and assuming efficacy data for risedronate, alendronate and denosumab as from a previous publication (14)

Threshold analysis was conducted to ascertain the necessary price reduction for zoledronic acid and denosumab that would be required for these products to be considered optimal given an assumed maximum willingness to pay for a QALY of $50,000 (38).

In addition, probabilistic sensitivity analysis was conducted through Monte Carlo simulation. Estimates of input parameters from a sample rather than population data were assumed to be random variables rather than fixed values. Analysis involves re-running the model employing different values for each data input randomly selected from a probability density function which is characterized by the mean value, a measure of dispersion (standard error) and type of distribution. This was repeated to obtain 5000 estimates of the costs and quality adjusted life years for each treatment.

Standard distributions were used for each data element: beta distributions for probabilities and utility values, log normal distributions for relative effects, normal distributions for disutilities and

gamma distributions for cost data. Uncertainty concerning the relative cost effectiveness of the alternative therapies is depicted by a cost effectiveness acceptability curve (CEAC) which depicts the likelihood a specific treatment option is given alternate values placed on a QALY (39).

RESULTS

Deterministic Analysis

For the base case population, zoledronic acid was the most effective treatment in terms of QALYs and life years gained and reduction in fracture costs. The next most effective treatment was alendronate (Table 2). Risedronate and etidronate were all dominated by alendronate as they were less effective (in terms of life years and QALYS) and were more costly. Denosumab was dominated by both alendronate and zoledronic acid. The incremental cost per QALY gained for alendronate versus no therapy was $10,895 ($16,488 per life year gained). The incremental cost per QALY gained for zoledronic acid versus alendronate was $1,609,172 ($2,713,512 per life year gained).

Analysis of Variability

Results varied by the population's age, bone mineral density and fracture history (Table 3). Typically, the incremental effect of alendronate improved the older the patient population (except for the comparison of 85 and 90 year olds), the lower the populations BMD and given the previous incidence of fracture. For several patient populations, alendronate was dominant over no therapy. However, for patients aged 65 with no previous fracture and a BMD no less than 3 SD below the mean of a young female, the incremental cost per QALY gained for alendronate versus no therapy exceeded $50,000.

Zoledronic acid remained more effective than alendronate for all patient populations with the associated incremental cost per QALY gained ranging from $224,826 to $4,591,999. For all patient populations; risedronate, etidronate and denosumab were subject to dominance or extended dominance.

Sensitivity Analysis

The conclusions to be drawn from the results of the sensitivity analysis did not differ from the base analysis (Table 4). The ICUR for alendronate versus no therapy remained less than $27,000 in all scenarios. In one scenario, full adherence for all treatments, zoledronic acid was dominated by alendronate. In the other scenarios, the ICUR for zoledronic acid versus alendronate exceeded $1.1 million. In all scenarios, risedronate and etidronate were dominated by alendronate. In the analysis using different relative effectiveness denosumab was found to be more effective than alendronate but with an ICUR of $1.6 million. Similarly, within the sensitivity analysis whereby all patients with wrist or vertebral fractures were hospitalized, denosumab was subject to extended dominance through alendronate and zoledronic acid. In this scenario, the ICUR for denosumab versus alendronate was $2.3 million.

Within the base case analysis, the price for denosumab would need to be reduced by 82.6% for denosumab to be considered the optimal treatment strategy. For zoledronic acid the price reduction would have to be 77.3% for a similar finding.

Probabilistic Analysis

The probabilistic analysis found similar results to the deterministic analysis. For the base case population, the incremental costs per QALY gained for alendronate versus no therapy was $14,588 and for zoledronic acid versus alendronate was $1,481,003. For threshold values of a QALY greater than $17,000, alendronate was optimal in more than 50% of replications. For threshold values up to $119,000, the probability for both denosumab and zoledronic acid of being optimal was 0%. At threshold value of a QALY greater than $1.58 million, the probability that zoledronic acid was optimal exceeded that of alendronate (Figure 2).

DISCUSSION

Based on the available literature this is the first economic evaluation comparing this wide range of osteoporotic treatments which is independent of manufacturer sponsorship. There are a number of publications which have suggested that manufacturer sponsored economic evaluations may be susceptible to bias (40,41).

The results of this analysis are in contrast to two previous studies of denosumab funded by the manufacturer. Chau and colleagues reported an ICUR of $60,266 (2010 Can$) for denosumab versus alendronate for a base case of 72 year old women with previous fracture and Jonsson and colleagues reported an ICUR of €27,090 (2009 euros) for 71 year old women with 34% with previous fracture (14,42). These findings contrast with our result that alendronate dominated denosumab.

The difference in results between the two studies and the current analysis may be due to four issues. First, it is unclear that the two manufacturer sponsored studies adequately accounted for calibration; failure to calibrate has been shown to greatly over estimate the cost effectiveness of osteoporotic treatments (43). Secondly, the two other studies used different NMA results which slightly favoured the denosumab in terms of reductions in hip fractures. Thirdly, the studies made the assumption that there was greater disutility in first year post vertebral fracture than hip fracture. Finally the studies used different estimates of treatment adherence than in the current study.

Sensitivity analysis addressed the last three of these differences. Analysis assuming different efficacy as used in the two manufacturer sponsored studies found denosumab no longer dominated by alendronate but with an ICUR of $1.6 million. Analysis assuming that all vertebral and wrist fractures required hospitalization; an extreme assumption but in line with assuming the high disutility with vertebral fracture in the manufacturer sponsored studies; found denosumab no longer dominated by alendronate but subjected to extended dominance through alendronate and zoledronic acid. The ICUR for denosumab versus alendronate was $2.2 million. Finally, analysis assuming full adherence with denosumab and zoledronic acid, found denosumab no longer dominated by alendronate but subjected to extended dominance through alendronate and zoledronic acid. The ICUR for denosumab versus alendronate was $10.8 million. An analysis incorporating all of the above found the ICUR for denosumab versus alendronate was $1.1 million still higher than the manufacturer's estimate.

There are a number of limitations to our analysis. Efficacy data was based on a network meta analysis of randomized controlled trials the majority of which were placebo controlled rather than active comparators. The lack of head to head clinical trials is worrisome. It would be beneficial that future novel treatments for osteoporosis are compared with current existing therapies especially oral bisphosphonates, which are available in low-cost generic format.

The base analysis did not address a potential subgroup of patients who are intolerant to oral bisphosphonates as this was not the focus of the analysis. For certain subgroups relating to more elderly patients with previous fracture, the ICUR for zoledronic acid versus no therapy was

less than $50,000. This suggests that for this narrow subgroup, annual therapy with zoledronic acid may be cost effective. In all populations, denosumab was dominated by zoledronic acid.

Costs were based on the most recent available Canadian data, and drug costs were obtained from the Ontario drug formulary. Although drug costs may vary by provincial formulary, it is unlikely that list prices would vary substantially enough to alter the conclusions of this analysis. However, there is greater potential that provincial drug plans may have negotiated lower prices for the newer osteoporotic treatments though those prices remain confidential. Threshold analysis was conducted and demonstrated that the costs of denosumab and zoledronic acid needed to be reduced substantially to allow these therapies to become optimal.

In conclusion this study demonstrated that treatment with alendronate was the most cost effective strategy for women who either have a previous fracture are aged over 70 and/or have a bone mineral density at least 3.5 standard deviations below the young adult mean. The newer treatment strategies were not cost effective and in the case of denosumab were found to be less effective than alendronate. The required price reductions to allow these newer therapies to be considered cost effective were substantial. Thus, the conclusion from this analysis is that denosumab and zoledronic acid should not be covered under provincial drug formularies for the general osteoporotic population unless substantial price reductions can be negotiated. Zoledronic acid may be optimal for patients intolerant to weekly bisphosphonate therapy at high risk of fracture, though the size of this potential sub group is unclear and clinical evidence relating specifically to this subgroup is unavailable.

REFERENCES

1. World Health Organization. Assessment of osteoporosis at the primary health care level. Summary Report of a WHO Scientific Group. WHO, Geneva 2007.
2. Qaseem A, Snow V, Shekelle P, et al. Pharmacologic treatment of low bone density or osteoporosis to prevent fractures: a clinical practice guideline from the American College of Physicians. Ann Intern Med. 2008 Sep 16;149(6):404-15.
3. Tenenhouse A, Joseph L, Kreiger N et al. Estimation of the prevalence of low bone density in Canadian women and men using a population-specific DXA reference standard: the Canadian Multicentre Osteoporosis Study (CaMos). Osteoporos Int. 2000;11(10):897-904
4. Haentjens P, Magaziner J, Colón-Emeric CS et al. Meta-analysis: excess mortality after hip fracture among older women and men. Ann Intern Med. 2010 Mar 16;152(6):380-90.
5. Tarride JE, Hopkins RB, Leslie WD, et al. The burden of illness of osteoporosis in Canada. Osteoporos Int. Nov 2012;23(11):2591-2600.
6. Lix LM, Azimaee M, Osman BA et al. Osteoporosis-related fracture case definitions for population-based administrative data. BMC Public Health. 2012 May 18;12:301
7. Fontenot HB, Harris AL. Pharmacologic management of osteoporosis. J Obstet Gynecol Neonatal Nurs. 2014 Mar;43(2):236-45
8. Raisz LG. Pathogenesis of osteoporosis: concepts, conflicts, and prospects. J Clin Invest. 2005 Dec;115(12):3318-25.
9. Guidelines for the economic evaluation of health technologies: Canada [3rd Edition]. Ottawa: Canadian Agency for Drugs and Technologies in Health; 2006.
10. Coyle D, Cranney A, Lee KM, Welch V, Tugwell P. Cost effectiveness of nasal calcitonin in postmenopausal women: use of Cochrane Collaboration methods for meta-analysis within economic evaluation. Pharmacoeconomics 2001;19:565-75.
11. Waldegger L, Cranney A, Hing M, Coyle D. Cost-effectiveness of hip protectors in institutional dwelling elderly. Osteoporosis International 2003;14: 243-250.
12. Coyle D, Hadj Tahar A, Murphy G et al. Teriparatide and bisphosphonates for treatment of osteoporosis in women: a clinical and economic analysis [Technology report no 70]. Ottawa: Canadian Agency for Drugs and Technologies in Health; 2006.
13. Coyle D, Tosteson ANA. Towards a reference case for economic evaluations of osteoporosis treatments. J Rheumatol. 2003;68s:31-6.
14. Chau D, Becker DL, Coombes ME, Ioannidis G, Adachi JD, Goeree R Cost-effectiveness of denosumab in the treatment of postmenopausal osteoporosis in Canada. J Med Econ. 2012;15 Suppl 1:3-14.

15. Statistics Canada. Living arrangements of seniors. Families, households and marital status. Structural type of dwelling and collectives, 2011 Census of Population. Ottawa: Statistics Canada; 2013.
16. Statistics Canada. Health Statistics Division. Deaths 2009. Ottawa: Statistics Canada; 2012.
17. Holvik K, Ranhoff AH, Martinsen MI, Solheim LF. Predictors of mortality in older hip fracture inpatients admitted to an orthogeriatric unit in Oslo, Norway. J Aging Health. 2010 Dec;22(8):1114-31.
18. Kado DM, Browner WS, Palermo L, Nevitt MC, Genant HK, Cummings SR. Vertebral fractures and mortality in older women: a prospective study. Study of Osteoporotic Fractures Research Group. Archives of Internal Medicine 1999;159:215-20
19. Haentjens P, Magaziner J, Colón-Emeric CS et al. Meta-analysis: excess mortality after hip fracture among older women and men. Ann Intern Med. 2010 Mar 16;152(6):380-90
20. Gandjour A, Weyler EJ. Cost-effectiveness of preventing hip fractures by hip protectors in elderly institutionalized residents in Germany. Value Health. 2008 Dec;11(7):1088-95.
21. Kanis JA, Borgstrom F, Johnell O, Jonsson B. Cost-effectiveness of risedronate for the treatment of osteoporosis and prevention of fractures in postmenopausal women. Osteoporos Int. 2004 Nov;15(11):862-71.
22. Marshall D, Johnell O, Wedel H. Meta-analysis of how well measures of bone mineral density predict occurrence of osteoporotic fractures. BMJ 1996;312:1254-1259.
23. Klotzbuecher CM, Ross PD, Landsman PB, Abbott TA, Berger M. Patients with prior fractures have an increased risk of future fractures: a summary of the literature and statistical synthesis. J Bone Miner Res 2000;15(4):721-39.
24. Crilly RG, Tanner DA, Kloseck M, Chesworth BM. Hip fractures in long-term care: is the excess explained by the age and gender distribution of the residents? J Aging Res. 2010 Aug 24;2010:291258.
25. Nikitovic M, Wodchis WP, Krahn MD, Cadarette SM. Direct health-care costs attributed to hip fractures among seniors: a matched cohort study. Osteoporos Int. 2013 Feb;24(2):659-69
26. Canadian Institute for Health Information. Patient Cost Estimator. Ottawa: Canadian Institute for Health Information 2013.
27. Ontario Ministry of Health and Long-Term Care. Ontario drug benefit formulary/comparative drug index. Edition 42 Effective March 27, 2014 Toronto: Queen's Printer for Ontario 2014.
28. Statistics Canada. Canadian Community Health Survey - Annual Component, 2010. Ottawa: Statistics Canada; 2011.

29. Peasgood T, Herrmann K, Kanis JA, Brazier JE. An updated systematic review of health state utility values for osteoporosis related conditions. Osteoporos Int. 2009 Jun;20(6):853-68.
30. Oleksik A, Lips P, Dawson A et al. Health-related quality of life in postmenopausal women with low BMD with or without prevalent vertebral fractures. J Bone Miner Res. 2000 Jul;15(7):1384-92.
31. Maher M, Sullivan SM, Chen L, Hsieh S-C, Peterson J, Wells GA. The effectiveness and safety of denosumab compared with bisphosphonates in the prevention of osteoporosis in post-menopausal women: a systematic review and network meta-analysis.
32. Emkey RD, Ettinger M. Improving compliance and persistence with bisphosphonate therapy for osteoporosis. Am J Med. 2006 Apr;119(4 Suppl 1):S18-24.
33. Freemantle N, Satram-Hoang S, Tang ET et al. Final results of the DAPS (Denosumab Adherence Preference Satisfaction) study: a 24-month, randomized, crossover comparison with alendronate in postmenopausal women. Osteoporos Int. 2012 Jan;23(1):317-26.
34. Coyle D, Lee KM, Cooper N J. Use of evidence in decision models. In Shemilt I, Mugford M, Vale L, Marsh K, Donaldson C. Evidence-Based Decisions and Economics: Health Care, Social Welfare, Education and Criminal Justice, Wiley-Blackwell, Oxford, UK, 2010.
35. Doubilet P, Begg CB, Weinstein MC, Braun P, McNeil BJ. Probabilistic sensitivity analysis using Monte Carlo simulation: a practical approach. Medical Decision Making 1985;5:157-177
36. Tosteson AN, Jonsson B, Grima DT, O'Brien BJ, Black DM, Adachi JD. Challenges for model-based economic evaluations of postmenopausal osteoporosis interventions. Osteoporosis International. 2001;12:849-57.
37. Tonino RP, Meunier PJ, Emkey R, et al. Skeletal benefits of alendronate: 7-year treatment of postmenopausal osteoporotic women. J Clin Endocrinol Metab 2000;85:3109-3115.
38. Winquist E, Bell CM, Clarke JTR et al. An evaluation framework for funding drugs for rare disease. Value in Health 2012:15; 982–986
39. Briggs AH. Handling uncertainty in cost effectiveness models. Pharmacoeconomics 2000; 17:479–500.
40. Bell CM, Urbach DR, Ray JG, Bayoumi A, Rosen AB, Greenberg D, Neumann PJ. Bias in published cost effectiveness studies: systematic review. BMJ. 2006 Mar 25;332(7543):699-703.
41. Garattini L, Koleva D, Casadei G. Modeling in pharmacoeconomic studies: funding sources and outcomes. Int J Technol Assess Health Care2010;26:330-3.

42. Jönsson B, Ström O, Eisman JA et al. Cost-effectiveness of Denosumab for the treatment of postmenopausal osteoporosis. Osteoporos Int. 2011 Mar;22(3):967-82
43. Coyle D. The impact of failure to calibrate on results in economic evaluation. Med Decis Making. 2005; 25:E32.

TABLES AND FIGURES

Table 1 Parameter Estimates

Parameter	Base Value	Probability Distribution	Source
Natural History Data			
Relative increase in risk of fracture for 1 SD decrease in bone density below age adjusted mean	1.5	Lognormal (1.4, 1.6)	22
Relative risk of hip fracture given previous fracture	2.4	Lognormal (1.9, 3.2)	23
Relative risk of wrist fracture given previous fracture	2	Lognormal (1.7, 2.4)	23
Relative risk of spine fracture given previous fracture	1.9	Lognormal (1.7, 2.3)	23
Relative increase in risk of hip fracture in LTC	1.5	Lognormal (1.3, 1.7)	24
RR risk of mortality post hip fracture with LTC	3.24	Lognormal (2.37, 4.43)	17
RR of mortality post vertebral fractures	1.16	Lognormal (1.03, 1.3)	18
RR of mortality post hip fractures	2.87	Lognormal (2.52, 3.27)	19
RR of mortality LTC	1.16	Lognormal (1.1, 1.2)	20
Annual probability of vertebral fracture			14
50-54	0.0007	Fixed	
55-59	0.0011	Fixed	
60-64	0.0016	Fixed	
65-69	0.0025	Fixed	
70-74	0.0038	Fixed	
75-79	0.0059	Fixed	
80-84	0.0089	Fixed	
85-89	0.0137	Fixed	
90-94	0.0210	Fixed	
Annual probability of wrist fracture			14
50-54	0.0018	Fixed	
55-59	0.0023	Fixed	
60-64	0.0033	Fixed	
65-69	0.0038	Fixed	
70-74	0.0044	Fixed	
75-79	0.0049	Fixed	
80-84	0.0058	Fixed	
85-89	0.0064	Fixed	
90-94	0.0074	Fixed	
Annual probability of hip fracture			14
50-54	0.0001	Fixed	
55-59	0.0003	Fixed	
60-64	0.0006	Fixed	
65-69	0.0014	Fixed	
70-74	0.0026	Fixed	
75-79	0.0062	Fixed	
80-84	0.0112	Fixed	
85-89	0.0197	Fixed	
90-94	0.0259	Fixed	
Proportion of women residing in LTC			15
65-69	0.01	Fixed	
70-74	0.023	Fixed	

75-79	0.057	Fixed	
80-84	0.136	Fixed	
>85	0.334	Fixed	
Proportion of women who are osteoporotic			3
<50	0.0133	Fixed	
50-59	0.0605	Fixed	
60-69	0.1825	Fixed	
70-79	0.2699	Fixed	
>80	0.4209	Fixed	
Mortality in general population (females)			16
50-54	0.0026	Fixed	
55-59	0.0039	Fixed	
60-64	0.0062	Fixed	
65-69	0.0094	Fixed	
70-74	0.0158	Fixed	
75-79	0.0263	Fixed	
80-84	0.0474	Fixed	
85-89	0.0849	Fixed	
>90	0.1786	Fixed	
Increase in fracture with osteoporosis- wrist			21
50-54	2.54	Fixed	
55-59	2.33	Fixed	
60-64	2.15	Fixed	
65-69	2.05	Fixed	
70-74	1.96	Fixed	
75-79	1.81	Fixed	
80-84	1.77	Fixed	
85-89	1.56	Fixed	
>90	1.54	Fixed	
Increase in fracture with osteoporosis- vertebral			21
50-54	3.18	Fixed	
55-59	2.82	Fixed	
60-64	2.36	Fixed	
65-69	2.34	Fixed	
70-74	2.16	Fixed	
75-79	1.97	Fixed	
80-84	1.85	Fixed	
85-89	1.58	Fixed	
>90	1.51	Fixed	
Increase in fracture with osteoporosis- hip			21
50-54	6.3	Fixed	
55-59	5.16	Fixed	
60-64	4.21	Fixed	
65-69	3.51	Fixed	
70-74	2.95	Fixed	
75-79	2.43	Fixed	
80-84	2.09	Fixed	
85-89	1.6	Fixed	
>90	1.4	Fixed	
Increase in fracture for women with a T-score of 2.5 SD - wrist			21

50-54	1.54	Fixed	
55-59	1.36	Fixed	
60-64	1.2	Fixed	
65-69	1.06	Fixed	
70-74	0.94	Fixed	
75-79	0.82	Fixed	
80-84	0.73	Fixed	
85-89	0.64	Fixed	
>90	0.57	Fixed	
Increase in fracture for women with a T-score of 2.5 SD - vertebral			21
50-54	1.93	Fixed	
55-59	1.65	Fixed	
60-64	1.32	Fixed	
65-69	1.21	Fixed	
70-74	1.04	Fixed	
75-79	0.89	Fixed	
80-84	0.76	Fixed	
85-89	0.65	Fixed	
>90	0.56	Fixed	
Increase in fracture for women with a T-score of 2.5 SD - hip			21
50-54	3.87	Fixed	
55-59	3.02	Fixed	
60-64	2.35	Fixed	
65-69	1.82	Fixed	
70-74	1.42	Fixed	
75-79	1.11	Fixed	
80-84	0.86	Fixed	
85-89	0.66	Fixed	
>90	0.52	Fixed	
Proportion of vertebral fractures requiring hospitalization	0.061	Beta (1437, 21956)	5, 14
Proportion of wrist fractures requiring hospitalization	0.159	Beta (3697, 19565)	5, 14
<u>Relative Risk of Fracture versus No Treatment</u>			
Hip Fracture			31
Alendronate	0.5249	Lognormal (0.27, 0.98)	
Etidronate	1.03	Lognormal (0.16, 5.69)	
Risedronate	0.7399	Lognormal (0.51, 1.09)	
Denosumab	0.602	Lognormal (0.33, 1.08)	
Zoledronic acid	0.5826	Lognormal (0.36, 0.95)	
Vertebral Fracture			31
Alendronate	0.5499	Lognormal (0.42, 0.72)	
Etidronate	0.4891	Lognormal (0.24, 0.92)	
Risedronate	0.6125	Lognormal (0.47, 0.8)	
Denosumab	0.3282	Lognormal (0.23, 0.46)	
Zoledronic acid	0.2959	Lognormal (0.21, 0.41)	
Wrist Fracture			31
Alendronate	0.7391	Lognormal (0.39, 1.12)	
Etidronate	2.223	Lognormal (0.37, 16.2)	
Risedronate	0.6121	Lognormal (0.25, 1.47)	
Denosumab	1.163	Lognormal (0.55, 2.42)	

Zoledronic acid	1	Fixed	
Adherence with Treatment at 12 months			
Alendronate once a day	0.57	Beta (57,43)	12
Etidronate	0.65	Beta (65,35)	12
Risedronate once a day	0.62	Beta (62,38)	12
Relative reduction in non-adherence			
Weekly oral bisphosphonate versus daily	0.719	Lognormal (0.713, 0.726)	32
Denosumab versus bisphosphonate	0.54	Lognormal (0.31, 0.93)	33
Event Cost Data			
Cost of hip fracture			25
Women living in the community	47685.81	Gamma (47686, 378)	
Women living in LTC	18486.46	Gamma (18486, 380)	
Women who die following fracture	11524.40	Gamma (11524, 1349)	
2nd year post fracture	4846.89	Gamma (4847, 199)	
Cost of wrist fracture			
Ambulatory	322.41	Gamma (322, 81)	12
Hospitalized	4523.18	Gamma (4523, 1131)	26
Cost of vertebral fracture			
Ambulatory	146.55	Gamma (147, 37)	12
Hospitalized	8843.19	Gamma (8843, 2211)	26
Cost of no fracture living in the LTC	43709.45	Gamma (43709, 10927)	25
Cost of no fracture living in the community	8577.69	Gamma (8578, 2144)	25
Annual Drug Costs			
Alendronate	176.92	Fixed	27
Etidronate	122.39	Fixed	27
Risedronate	180.31	Fixed	27
Denosumab	774.61	Fixed	27
Zoledronic acid	733.29	Fixed	27
Utility Values			
No Fracture			
Women aged 65-69	0.8359	Beta (5920, 1162)	28
Women aged 70-74	0.8197	Beta (4878, 1073)	28
Women aged 75-79	0.7854	Beta (3792, 1036)	28
Women aged >80	0.7016	Beta (4623, 1966)	28
Disutilities from fracture			
Hip fracture - 1st year	0.228	Normal (0.228, 0.06)	29
Hip fracture - 2nd year	0.153	Normal (0.153, 0.04)	29
Vertebral fracture – hospitalized	0.3	Normal (0.3, 0.08)	29
Vertebral fracture - not hospitalized	0.075	Normal (0.075, 0.04)	30
Wrist	0.04	Normal (0.04, 0.01)	29

Beta and gamma distributions depicted by shape and scale parameters. Normal distributions depicted by mean and standard errors. Lognormal distribution depicted by 95% CI.

Table 2: Results for base population

	Total Costs	Drug Costs	Fracture Costs	QALYs	Life years	ICUR vs. No therapy^	Sequential Cost Utility Analysis^	Sequential Cost Effectiveness Analysis*
No Therapy	$3,365	$0	$3,365	7.530	10.537			
Alendronate	$3,465	$590	$2,874	7.539	10.543	$10,895	$10,895	$16,488
Zoledronic Acid	$5,815	$2,984	$2,830	7.541	10.544	$231,877	$1,609,172	$2,713,512
Etidronate	$3,737	$326	$3,412	7.530	10.537	Dominated	Dominated by no therapy and alendronate	Dominated by no therapy and alendronate
Risedronate	$3,670	$587	$3,083	7.536	10.541	$53,648	Dominated by alendronate	Dominated by alendronate
Denosumab	$5,826	$2,926	$2,900	7.539	10.543	$270,927	Dominated by zoledronic acid and alendronate	Dominated by zoledronic acid and alendronate

* Incremental cost per life year gained

Table 3: Sequential Cost Utility Analysis Results for Alternative Patient Populations

Patients with no previous fracture

Age	No. of SDs, BMD is less than young adult mean			
	-2.5	**-3**	**-3.5**	**-4**
65	ICUR for A vs. NT= $98,347. ICUR for Z vs. A= $4,591,999.	ICUR for A vs. NT= $62,628. ICUR for Z vs. A= $3,644,184.	ICUR for A vs. NT= $34,544. ICUR for Z vs. A= $2,884,865.	ICUR for A vs. NT= $12,560. ICUR for Z vs. A= $2,279,080.
70	ICUR for A vs. NT= $49,580. ICUR for Z vs. A= $2,680,222.	ICUR for A vs. NT= $26,656. ICUR for Z vs. A= $2,101,633.	ICUR for A vs. NT= $8,583. ICUR for Z vs. A= $1,642,386.	A dominates NT. ICUR for Z vs. A= $1,279,867.
75	ICUR for A vs. NT= $10,895. ICUR for Z vs. A= $1,609,172.	A dominates NT. ICUR for Z vs. A= $1,251,712.	A dominates NT. ICUR for Z vs. A= $969,380.	A dominates NT. ICUR for Z vs. A= $747,769.
80	A dominates NT. ICUR for Z vs. A= $1,059,966.	A dominates NT. ICUR for Z vs. A= $823,135.	A dominates NT. ICUR for Z vs. A= $635,878.	A dominates NT. ICUR for Z vs. A= $488,762.
85	A dominates NT. ICUR for Z vs. A= $754,076.	A dominates NT. ICUR for Z vs. A= $586,198.	A dominates NT. ICUR for Z vs. A= $452,878.	A dominates NT. ICUR for Z vs. A= $347,652.
90	A dominates NT. ICUR for Z vs. A= $826,190.	A dominates NT. ICUR for Z vs. A= $644,656.	A dominates NT. ICUR for Z vs. A= $499,710.	A dominates NT. ICUR for Z vs. A= $384,625.

Patients with previous fracture

Age	No. of SDs, BMD is less than young adult mean			
	-2.5	**-3**	**-3.5**	**-4**
65	ICUR for A vs. NT= $15,204. ICUR for Z vs. A= $2,335,023.	A dominates NT. ICUR for Z vs. A= $1,903,776.	A dominates NT. ICUR for Z vs. A= $1,551,666.	A dominates NT. ICUR for Z vs. A= $1,264,170.
70	A dominates NT. ICUR for Z vs. A= $1,473,833.	A dominates NT. ICUR for Z vs. A= $1,201,162.	A dominates NT. ICUR for Z vs. A= $978,532.	A dominates NT. ICUR for Z vs. A= $796,762.
75	A dominates NT. ICUR for Z vs. A= $921,077.	A dominates NT. ICUR for Z vs. A= $749,702.	A dominates NT. ICUR for Z vs. A= $609,785.	A dominates NT. ICUR for Z vs. A= $495,556.
80	A dominates NT. ICUR for Z vs. A= $605,251.	A dominates NT. ICUR for Z vs. A= $492,559.	A dominates NT. ICUR for Z vs. A= $400,570.	A dominates NT. ICUR for Z vs. A= $325,492.
85	A dominates NT. ICUR for Z vs. A= $419,288.	A dominates NT. ICUR for Z vs. A= $340,928.	A dominates NT. ICUR for Z vs. A= $276,986.	A dominates NT. ICUR for Z vs. A= $224,826.
90	A dominates NT. ICUR for Z vs. A= $442,656.	A dominates NT. ICUR for Z vs. A= $359,708.	A dominates NT. ICUR for Z vs. A= $292,021.	A dominates NT. ICUR for Z vs. A= $236,804.

A = alendronate, NT = no pharmacotherapy, Z= zoledronic acid, ICUR = incremental cost per QALY gained

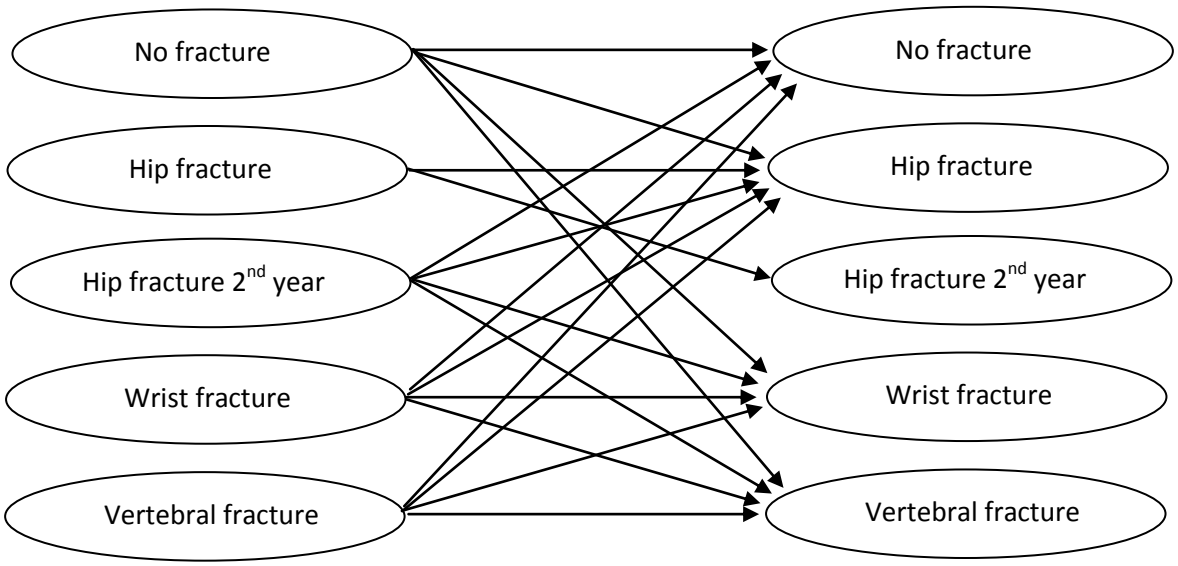
Table 4: Sensitivity Analysis

Scenario	Sequential Result
Base case	ICUR for alendronate versus no therapy = $10,895. ICUR for zoledronic acid versus alendronate = $1,609,172.
Set time = 0 years	ICUR for alendronate versus no therapy = $17,470. ICUR for zoledronic acid versus alendronate = $1,837,579.
Set time = 5 years	Alendronate dominates no therapy. ICUR for zoledronic acid versus alendronate = $1,203,740.
Duration of treatment = 2 years (set time =0)	ICUR for alendronate versus no therapy = $17,298. ICUR for zoledronic acid versus alendronate = $2,195,811.
Duration of treatment = 3 years (set time =0)	ICUR for alendronate versus no therapy = $17,353. ICUR for zoledronic acid versus alendronate = $1,985,781.
Duration of treatment = 4 years (set time =0)	ICUR for alendronate versus no therapy = $17,397. ICUR for zoledronic acid versus alendronate = $1,891,257.
Alternate efficacy data for risedronate, alendronate and denosumab	ICUR for alendronate versus no therapy = $26,154. ICUR for denosumab versus alendronate = $1,583,828. Risedronate subject to dominance by alendronate.
Full adherence to all therapies	ICUR for alendronate versus no therapy = $10,823.
Full adherence to zoledronic acid	ICUR for alendronate versus no therapy = $10,895. ICUR for zoledronic acid versus alendronate = $1,202,624.
Full adherence to denosumab and zoledronic acid	ICUR for alendronate versus no therapy = $10,895. ICUR for zoledronic acid versus alendronate = $1,202,624.
Equal adherence with denosumab and zoledronic acid	ICUR for alendronate versus no therapy = $10,895. ICUR for zoledronic acid versus alendronate = $2,593,864.
No increase in mortality with vertebral fracture	ICUR for alendronate versus no therapy = $11,155. ICUR for zoledronic acid versus alendronate = $1,837,180.
No increase in mortality with hip or vertebral fracture	ICUR for alendronate versus no therapy = $14,861. ICUR for zoledronic acid versus alendronate = $2,635,703.
No disutility from hip fracture after 1 year	ICUR for alendronate versus no therapy = $13,006. ICUR for zoledronic acid versus alendronate = $1,771,500.
All wrist and vertebral fractures require hospitalization	ICUR for alendronate versus no therapy = $1,605. ICUR for zoledronic acid versus alendronate = $794,715.
Discount rate =0%	ICUR for alendronate versus no therapy = $4,734. ICUR for zoledronic acid versus alendronate = $1,136,851.
Discount rate =3%	ICUR for alendronate versus no therapy = $8,442.

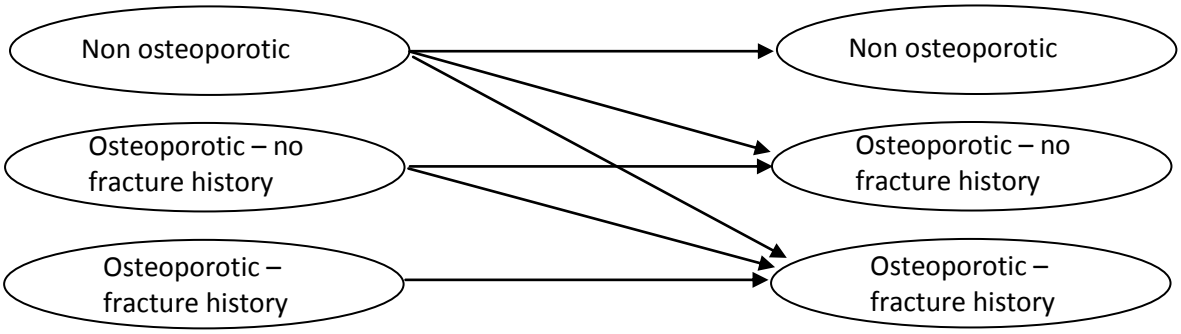
	ICUR for zoledronic acid versus alendronate = $1,427,320.
Alternate efficacy data for risedronate, alendronate and denosumab, full adherence to denosumab and vertebral fractures require hospitalization	ICUR for alendronate versus no therapy = $13,236. ICUR for denosumab versus alendronate = $1,093,104. Risedronate subject to dominance by alendronate.

Note: In all analyses, therapies that are not listed were subject to dominance or extended dominance.

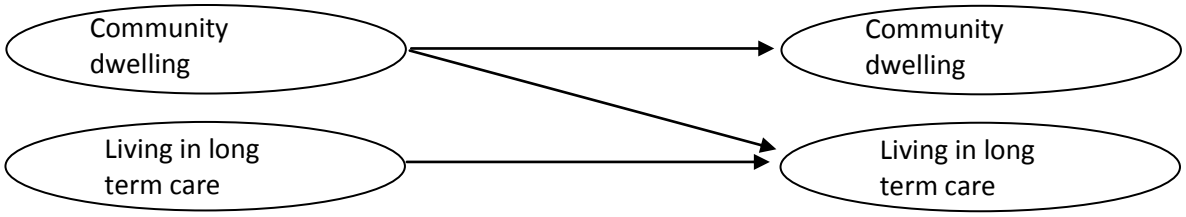
Figure 1: Design of Decision Model



a. Fractures

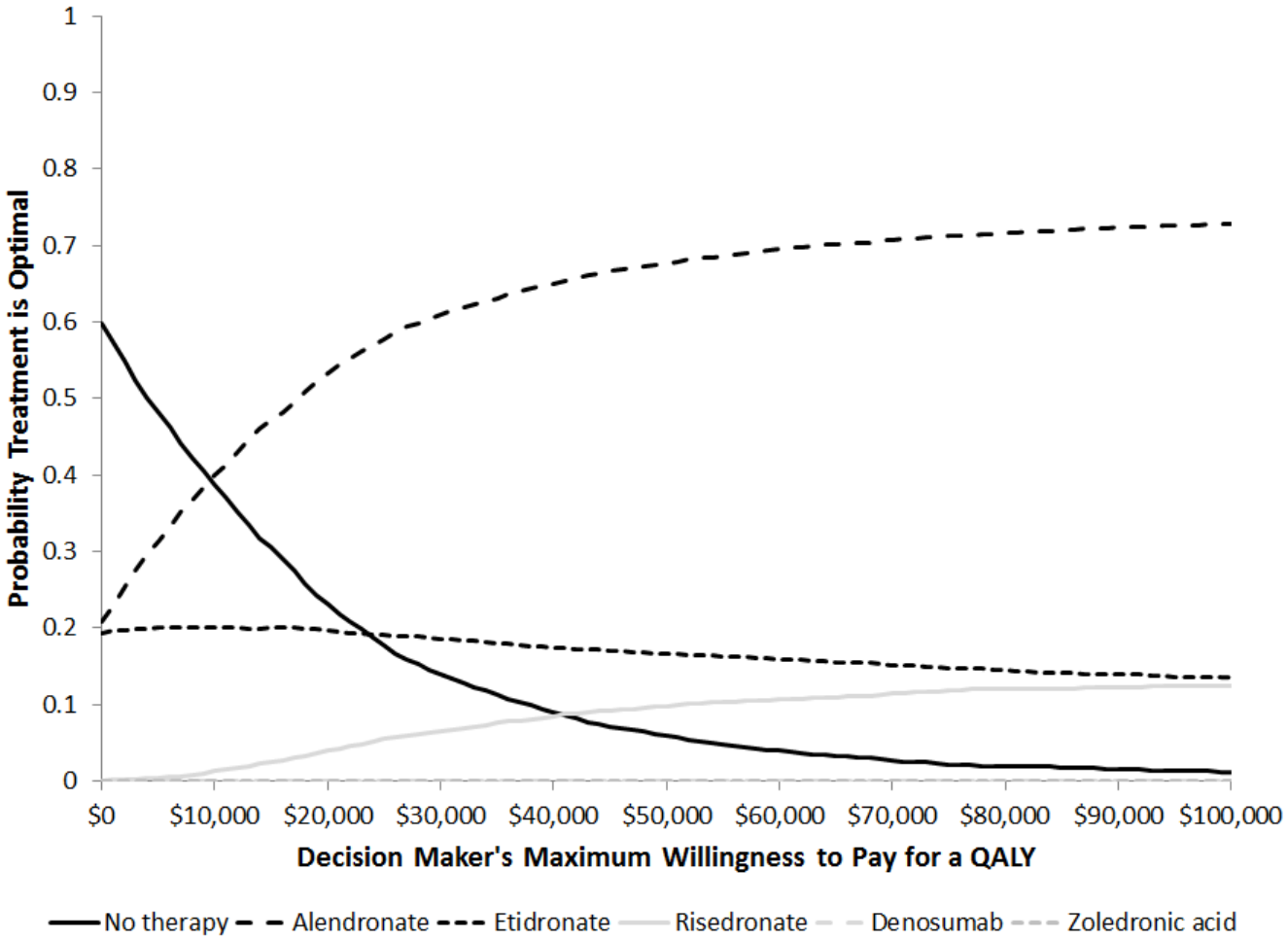


b. Osteoporotic status



c. Residential status

Figure 2: Cost Effectiveness Acceptability Curve



APPENDIX E.

OSTEOPOROSIS NETWORK META-ANALYSIS

TITLE:

The effectiveness and safety of denosumab compared with bisphosphonates in the prevention of osteoporosis in post-menopausal women: a systematic review and network meta-analysis

AUTHORS:

Maurica Maher,[1] Shannon M. Sullivan, [1] Li Chen, [1] Shu-Ching Hsieh, [1] Joan Peterson, [1] George A. Wells[1,2]

AFFILIATIONS:

1 University of Ottawa Heart Institute, Ottawa, ON

2 Department of Epidemiology & Community Medicine, University of Ottawa, Ottawa, ON

SHORT TITLE: Osteoporosis NMA

Keywords: network meta-analyses; osteoporosis; bisphosphonates, denosumab

Tables and Figures: 2 Figures, 3 Tables

ABSTRACT

Objectives and Methods: To compare the safety and effectiveness of denosumab and five bisphosphonates in preventing postmenopausal osteoporosis. Completed, updated and ongoing Cochrane systematic reviews of these therapies were included. Randomized controlled trials of dosing regimens authorized in North America were included. Outcomes were vertebral, non-vertebral, hip and wrist fractures and withdrawals due to adverse events (WDAEs). Bayesian network meta-analysis methods were applied.

Results: 34 trials, including two head-to-head trials and one multi-arm trial, were synthesized. Zoledronic acid was found to be the most effective at preventing vertebral fracture (OR=0.28,95%CrI:0.19, 0.39), closely followed by denosumab (OR=0.31,95%CrI:0.22,0.44), when compared with placebo. Both were more effective than alendronate 10mg/day (OR=0.54,95%CrI:0.33,0.82 and OR=0.6,95%CI:0.37,0.91, respectively). Only alendronate 10 mg/day (OR=0.78,95%CrI:0.60,0.96) and risedronate (OR=0.74,95%CrI: 0.57,0.91) were effective at preventing non-vertebral fractures compared to placebo. Only alendronate 10 mg/day (OR=0.55, 95%CrI:0.27,0.98) and zoledronic acid (OR=0.60,95%CrI:0.36,0.94) were effective at preventing hip fractures compared to placebo and no differences in effectiveness were observed between therapies. No differences in preventing wrist fractures or in WDAEs were observed when therapies were compared with placebo or with each other.

Conclusions: Denosumab and zoledronic acid were the most effective therapies in preventing vertebral fracture in postmenopausal osteoporotic women when compared with placebo. However, alendronate was the only therapy effective in preventing all of vertebral, non-vertebral and hip fractures, the latter of which is a more clinically relevant outcome.. Conclusions on comparative safety are limited by relying only on WDAEs and the challenges synthesizing data on rare but relevant serious adverse events in NMAs.

INTRODUCTION

Osteoporosis is a skeletal disorder characterized by compromised bone strength, which predisposes those affected to an increased risk of fractures (1). In 1994 the World Health Organization Study Group on osteoporosis defined the disease as “a hip BMD level of more than 2.5 standard deviations (SDs) below the mean bone mineral density (BMD) for young, white, adult women” (2). Osteoporosis can also be diagnosed on a history of a previous osteoporotic or fragility fracture. Osteoporosis and its resultant fragility fractures primarily impact the elderly. It has been estimated that 1 in 4 women in Canada and 1 in 8 men have osteoporosis (3). Osteoporotic fractures can be classified as vertebral and non-vertebral (non-spine) fractures. Vertebral fractures are the most common type of osteoporotic fractures (4) and occur in 20% of postmenopausal women (5). These types of fractures are often asymptomatic and may go undetected (6). The morbidity, mortality, and financial consequences of osteoporotic fractures in the elderly are considerable. There is a significant increased risk of death after hip and vertebral fractures (7,8), which persists for 1 to 2 years after the event. Hip fractures are the most serious type of osteoporotic fracture and are associated with permanent disability, institutionalization, and increased mortality. (9)

Bisphosphonates have become the most commonly used treatments in postmenopausal osteoporosis (10-11). All of the commonly used bisphosphonates for osteoporosis have been shown to decrease vertebral fracture risk in randomized controlled trials (5), however, not all have been demonstrated to decrease a reduction in hip and other non-vertebral fractures. Non-vertebral fractures occur less frequently than vertebral fractures (12), and not all studies were powered to detect a difference with respect to non-vertebral fractures. Denosumab, a RANK ligand (RANKL) inhibitor, is also authorized in Canada for the treatment of postmenopausal women with osteoporosis at high risk for fracture (defined as a history of osteoporotic fracture, or multiple risk factors for fracture) or for patients who have failed or are intolerant to other osteoporosis therapies (13). The efficacies of alendronate, risedronate, and etidronate for the primary and secondary prevention of osteoporotic fractures in postmenopausal women have all been the subject of individual Cochrane reviews (14-16) and reviews on ibandronate, zoledronic acid and denosumab are in progress. However, few head-to-head trials have been conducted comparing osteoporosis therapies on fracture outcomes. In these cases, indirect treatment comparisons and network meta-analyses (NMAs), can be used to compare the results of treatments through the use of common comparators when there is a lack of available direct, or head-to-head, comparisons (17-18). Although some NMAs of osteoporosis therapies have been

published, none have specifically focused on how the safety and effectiveness of denosumab compares with bisphosphonates, the most commonly prescribed class of osteoporosis therapies. There are key differences between denosumab and bisphosphonates including the route of administration and dosing schedule, which may make it an appealing treatment option and understanding its comparative effectiveness with bisphosphonates in this context is important.

METHODS

Searches Strategy: Studies that compared alendronate, risedronate and etidronate to each other or placebo were identified through previously conducted Cochrane systematic reviews and meta-analyses. (14-16). These searches were updated to 2011 and also identified trials of ibandronate, zoledronic acid and denosumab as part of the process for conducting Cochrane reviews for these therapies. Databases searched included Medline (1948 to December 2011), EMBASE (1980-December 2011), CENTRAL, DARE, HTA Database, NHS EED, Scopus and the WHO International Clinical Trials Registry Platform, using the earlier data limit of 2000 but no language restriction. Search terms and strategies were based upon the original Cochrane systematic reviews, which have been previously reported. The United States Food and Drug Administration and the European Medicines Agency websites and bibliographies of relevant systematic and non-systematic reviews were also searched.

Inclusion Criteria: Included studies were active and placebo-controlled randomized trials in post-menopausal women diagnosed with osteoporosis on the basis of BMD and history of vertebral fracture. Only studies that included doses equivalent to authorized doses in North America were included in this analysis to make the study most relevant to clinical practice. Outcomes of interest were vertebral, non-vertebral, hip and wrist fractures and withdrawals due to adverse events. Fracture outcomes could be defined either clinically or radiographically and collected as primary outcomes, secondary outcomes are adverse events.

Data Extraction and Quality Assessment: Data was extracted using standardized data abstraction forms based on templates from the Cochrane collaboration (19). Abstraction included: 1) study design; 2) inclusion criteria and major exclusion criteria; 3) characteristics of participants,

including age, country, age at menopause and presence of vertebral fractures; 4) type of intervention including study drug, duration of use, dose, frequency, and route of administration; 5) and outcomes assessed, including number of patients with vertebral, non-vertebral, hip and wrist fractures, total withdrawals, withdrawals due to adverse events, and serious adverse events.

Methodological quality was assessed using the Cochrane Collaboration's risk of bias tool, which addressing adequacy of randomization and allocation concealment, blinding (participants/personnel and outcome assessment), completeness of outcome data, selective reporting and the presence of any other bias (19)

Bayesian mixed treatment comparisons: Mixed treatment comparison (MTC) models were performed using Bayesian methods (18). The software package WinBUGS (version 1.4.3) was used, which is the Windows© version of BUGS (Bayesian Inference Using Gibbs Sampling). This software was created at the MRC Biostatistical Unit in Cambridge, U.K., and can be downloaded free of charge (www.mrc-bsu.cam.ac.uk/bugs/welcome.shtml). The software uses a Monte Carlo chain simulation to construct the results. In this study 40 000 iterations were performed for each analysis. Random effects and fixed effects models were used for this analysis and the results compared to each other, with the intention that the random effects analysis would be the primary analysis and the fixed effects model would serve as a sensitivity analysis.

For each comparison informative priors were used based on a review of meta-analyses that investigated between-study heterogeneity (20). In some instances priors had to be modified to allow the model to converge, especially for those outcomes with fewer studies (such as wrist fractures). Priors were then combined with the extracted trial data to derive posterior distributions for all unknown model parameters. Results are reported as odds ratios (OR) with 95% credible intervals (95% CrI); credible intervals are considered to be the Bayesian equivalent of a classical confidence interval. Relative measures are presented in matrix tables showing the comparison of each treatment to each other and placebo (odds ratios, 95% CrI) for each outcome. Bolded values indicate significant results. The resultant odds ratios and the placebo group event rates were used to calculate the absolute measure, number needed to benefit or harm using the online tool available from: http://www.nntonline.net/visualrx/.

Clinical heterogeneity was assessed by examining the clinical and methodological diversity of the studies. Statistical heterogeneity was formally assessed formally based on the I^2 test statistic. I^2 Values of 25%, 50%, and 75% were interpreted as low, moderate, and high degrees of

heterogeneity. Inconsistency was evaluated through the use of a random effects inconsistency model applied in WinBUGS and plotting the posterior mean deviance contributions for the consistency and inconsistency models for each outcome.(21) Assessment of model fit was based on the deviance information criterion (DIC) and comparison of residual deviance to number of unconstrained data points.

Network diagrams outlining the available evidence for each outcome were created using NodeXL (vertebral, non-vertebral, hip and wrist fractures; withdrawals due to adverse events). NodeXL is a free, open-source template for Microsoft® Excel® that can be used to create network graphs (available at: http://nodexl.codeplex.com/). In the network diagram, the size of the treatment spheres represents the number of patients exposed to that treatment in all of the studies; the lines show direct comparisons between treatments (most often versus placebo); and the width of the lines indicates the number of randomized controlled trials (RCTs).

Subgroup and Sensitivity Analyses: Planned subgroup analyses were primary versus secondary prevention trials and treatment duration (studies greater than or equal to 3 years and those less than 3 years). Planned sensitivity analyses were random effects versus fixed-effect models.

RESULTS

Thirty-three studies were included in the analysis.(22-56) Twenty-five studies were included based on the original Cochrane meta-analyses for alendronate, risedronate and etidronate. An additional eight studies were included after updating these searches and including ibandronate, zolderonic acid and denosumab trials. The number of included studies for each outcome varied as not every outcome was reported for every study. Figure 1 outlines the search for included studies.

INSERT FIGURE 1. PRISMA Flow Diagram for Network Meta-Analysis of Osteoporosis Therapies

Of the 33 studies, there were only two “head-to-head” studies, which compared ibandronate and alendronate, and the remainder were placebo-controlled trials. Key trial characteristics are documented in Supplementary Table 1. Studies were conducted over a wide range of geographic regions including North America, Asia, Europe and Oceania. The majority of trials evaluated alendronate, one of the more established bisphosphonates, with trials conducted between 1996 and 2011. The earliest trials evaluated etidronate, beginning in 1990 and the most recent trials evaluated zoledronic acid in 2012. Fracture outcomes could be reported as primary or secondary outcomes or as adverse events and this varied across trials. Fracture was defined in multiple ways across studies and did not always distinguish between clinical and radiographic fractures. Although all women were postmenopausal, the mean age of trial populations ranged from 53 years (Hoskings 1998, Hooper 2005, Pouilles 1997, McClung 2009) to 78 years (Greenspan 2002, McClung 2001). The mean years since menopause was not always reported in individual trials, however, when it was, years since menopause ranged from 3 years (Pouilles 1997) to 31 years (Greenspan 2002). The proportion of patients with a prevalent fractures was variable, when reported, ranging from zero, in many studies, to 100% in some studies. Risk of bias of individual trials is reported in Supplementary Table 2. Risk of bias was high or unclear in the majority of trials for most categories other than selective outcome reporting.

Vertebral Fractures

Eighteen studies had vertebral fracture outcome data that could be included in the NMA (total participants 30 667). There were two trials that included comparisons between bisphosphonates: one trial compared alendronate 10 mg/day to ibandronate 150 mg/month orally and placebo, and the other trial compared ibandronate 2.5 mg/day orally to ibandronate 20 mg intermittently orally and placebo. All other studies compared various therapies to placebo. Figure 2 shows the network diagram for vertebral fractures.

INSERT FIGURE 2. Evidence network for vertebral fractures.

All treatments, except alendronate 5 mg/day and ibandronate 150 mg/month orally, were effective at preventing vertebral fractures when compared with placebo (Table 3). The ORs ranged from a high of OR=0.60 (95%CrI: 0.45,0.78) for risedronate to a low of OR=0.28 (95%CrI: 0.19, 0.39) for

zoledronic acid, indicating that zoledronic acid is the most effective therapy for vertebral fracture prevention.

The comparisons of each therapy to each other identified that both denosumab and zoledronic acid were significantly better at preventing vertebral fractures than alendronate 10 mg/day (OR=0.6: 95% CrI 0.37, 0.91 and OR=0.54; 95% CrI 0.33, 0.82, respectively) and risedronate (OR=0.54; 95% CrI 0.33, 0.82 and OR=0.48; 95% CrI 0.30, 0.73, respectively). In addition, zoledronic acid was better than ibandronate 20 mg intermittently orally (OR=0.57; 95% CrI 0.30, 0.98).

INSERT TABLE 1. Network meta-analysis results: vertebral fractures. [Odds Ratio (95% Credible Interval)].

Non-Vertebral Fractures

Twenty-eight studies had non-vertebral fracture outcome data that was included in the analysis (total participants 44 216). There were three trials that included comparisons between bisphosphonates, namely: alendronate 10 mg/day to ibandronate 2 mg every 3 months iv; alendronate 10 mg/day to ibandronate 150 mg/month orally and placebo; and ibandronate 2.5 mg/day orally to ibandronate 20 mg intermittently orally and placebo. All other studies compared various therapies to placebo (See Supplementary Figure 1 for network diagram). Only alendronate 10 mg/day orally (OR=0.78; 95% CrI 0.60, 0.96) and risedronate (OR 0.74; 95% CrI 0.57, 0.91) were effective at preventing non-vertebral fractures compared to placebo (see Table 2). None of the comparisons between therapies were statistically significant.

INSERT TABLE 2. Network meta-analysis results: non-vertebral fractures. Odds Ratio (95% Credible Interval)

Hip Fractures

Twelve studies had hip fracture outcome data that was included in the analysis (total participants 31 482). All studies compared the various therapies to placebo, and there were no studies that compared drug treatments to each other (See Supplementary Figure 2 for network diagram). The analysis showed that only alendronate 10mg/day (OR 0.55; 95% CrI 0.27,0.98) and zoledronic acid (OR 0.60; 95% CrI 0.36, 0.94) were effective at preventing hip fractures compared to placebo (See Table 3). None of the comparisons between therapies were statistically significant.

INSERT TABLE 3. Network meta-analysis results: hip fractures. Odds Ratio (95% Credible Interval)

Wrist Fractures

Nine studies reported wrist fracture outcome data that were included in the analysis (total participants 18 916). All studies compared the various therapies to placebo, and there were no studies that compared drug treatments to each other (See Supplementary Figure 3 for network diagram). None of the treatments were effective at preventing wrist fractures compared with placebo or each other (See Supplementary Table 3).

Withdrawals Due to Adverse Events

Twenty-six studies reported withdrawals due to adverse event outcome data that were included in the analysis (total participants 43 224). There were two trials that included comparisons between bisphosphonates: one trial compared alendronate 10 mg/day to ibandronate 150 mg/month orally and placebo, and the other trial compared ibandronate 2.5 mg/day orally to ibandronate 20 mg intermittently orally and placebo. All other studies compared various therapies to placebo (See Supplementary Figure 4 for network diagram). None of the therapies resulted in statistically significantly more withdrawals from treatment due to adverse events compared to placebo or each other (See Table 4).

INSERT TABLE 4. Network meta-analysis results: hip fractures. Odds Ratio (95% Credible Interval)

Number Needed to Treat (NNT) Compared with Placebo

For vertebral fractures, the number need to benefit (NNB) for statistically significant effect estimates ranged from 18 for zoledronic acid to 33 for risedronate. For non-vertebral fractures, the NNB was 80 for alendronate 10 mg/day and 70 for risedronate, which were the only therapies that had statistically significant effect estimates. For hip fractures, NNB were 188 for alendronate 10 mg/day and 212 for zoledronic acid. Supplementary Table 6 displays the NNT for all therapies versus placebo.

Subgroup Analyses: Trial Duration and Osteoporosis Prevention

Subgroup analyses examined differences in outcomes: for trials greater than or equal to three years versus trials less than three years in duration and for primary prevention versus secondary prevention trials. Primary prevention studies were defined as those that included women whose bone density was within 2 SD of the mean or those in which the prevalence of vertebral fracture at baseline was less than 20. Trials were classified as secondary prevention trials if they included women with bone density that was less than 2 SD values below mean bone mass, if they included only women who had a previous vertebral fracture, or if the average age of participants was greater than 62 years.

However, when denosumab was added to the subgroup analyses, estimates were unstable. No clear trends were observed when analyses excluding denosumab were considered. The efficacy of osteoporosis treatment for fracture prevention was enhanced, in some cases, by longer treatment duration, however this was not the case for every fracture outcome. For all outcomes and therapies the majority of trials were secondary prevention studies; often there were only one or no primary prevention trials for each therapy, therefore, the conclusions that can be drawn from these subgroup analyses are limited (Data not provided).

Sensitivity analyses: Fixed and Random Effects Models.

Sensitivity analyses were conducted that examined the differences in results between random and fixed effects models. Credible intervals were narrower for the fixed effect model, however,

the choice of model did not impact the statistically significance of any effect estimates (see Supplementary Table 6).

DISCUSSION

Summary of Findings

All treatments, with the exception of alendronate 5 mg/day and ibandronate 150 mg/month orally, were effective at preventing vertebral fractures. Zoledronic acid was found to be the most effective at preventing vertebral fracture, closely followed by denosumab, based on the magnitude of the mixed treatment effect size. Both denosumab and zoledronic acid were significantly better at preventing vertebral fractures than alendronate 10 mg/day and risedronate. Zoledronic acid was also significantly more effective than ibandronate 20 mg intermittently orally. However, variations in the definition of vertebral fractures and the inclusion of asymptomatic radiographic fractures in most trials, limit the clinical value of these data.

Only alendronate 10 mg/day and risedronate were effective at preventing non-vertebral fractures compared to placebo and no differences in effectiveness were observed between therapies. Only alendronate 10 mg/day and zoledronic acid were effective at preventing hip fractures compared to placebo and no differences in effectiveness were observed between therapies. However, only 12 studies provided data on hip fracture, despite the fact this is the most clinically relevant and patient-important outcome in osteoporosis. No therapies were effective at preventing wrist fractures compared with placebo or each other and no differences in WDAEs were observed when therapies were compared with placebo or each other.

Comparison to Other Network Meta-Analyses

There are five previously published NMAs that have compared the efficacy of osteoporosis therapies in fracture prevention, including both bisphosphonates and denosumab (58-62). All of the NMAs relied primarily on a Bayesian analysis framework and random effects models, although additional analyses (e.g. frequentist, fixed effects) were also conducted as sensitivity or

comparative analyses. However, differences exist in the scope of all these analyses, as described below and none of them attempted to compare safety outcomes.

The first of these studies conducted by Jansen et al in 2011(58) examined alendronate, ibandronate, risedronate, etidronate, and zoledronic acid and included only eight studies. The outcomes examined were vertebral, non-vertebral and hip fractures and the outcome measure was relative risk. They concluded that zoledronic acid was the most effective therapy for fracture prevention compared to placebo and the other therapies in the study, however they only presented the results of a fixed effect model. Although this would have led to narrower credible interval, similar results to our analysis, using a random effects model were obtained. All of the treatments were effective for the prevention of vertebral fractures compared to placebo with zoledronic acid being the most effective, risedronate being the most effective for preventing non-vertebral fractures and zoledronic acid being the best therapy for preventing hip fractures. The results were similar to those obtained from the current study, with the exception of alendronate at 10 mg/day which was found to be effective for non-vertebral fracture prevention in the current study.

The study performed by Hopkins et al (59) included 30 studies and 9 drugs; denosumab, strontium, and teriparatide were also included and the outcomes examined were vertebral, non-vertebral, hip and wrist fractures. Studies were included if they included fractures as either primary or secondary outcomes. However, only placebo-controlled trials were included in the analysis, therefore, this was not a true network meta-analysis. For the bisphosphonates the authors concluded that all the drugs were effective at preventing vertebral fractures except etidronate, and that zoledronic acid was the most efficacious for this outcome. For non-vertebral fractures alendronate and risedronate were the only effective therapies and this is consistent with the results of the current study. Zoledronic acid was not found to be effective at preventing hip fractures in this study, which differs from the results of the first NMA (69) and the current study.

The final NMA study available for comparison (60) included 34 studies in total with the following treatments: alendronate, etidronate, ibandronate, risedronate, zoledronic acid, strontium, teriparatide, denusomab, and raloxifene. This study conducted a primary analysis with only studies that had fracture outcomes as endpoints, however the authors conducted a sensitivity analysis with studies that had fractures as adverse event outcomes. This study also used classical and Bayesian analytic techniques to conduct random effect analyses. The authors found

that all the bisphosphonates with the exception of etidronate were effective at preventing vertebral fractures, that risedronate and zoledronic acid were the most effective for reducing non-vertebral and hip fractures, and that alendronate was effective at reducing non-vertebral fractures.

Murad et al. 2012 (61) conducted a NMA of the comparative effectiveness of different drug therapies in reducing fragility fractures that included 116 trials of bisphosphonates, teriparatide, selective estrogen receptor modulators, denosumab or calcium and vitamin D. However, broad inclusion criteria were applied in this NMA and included individuals at risk of developing fragility fractures and not just post-menopausal women with osteoporosis, therefore, results are not directly relevant to the current study.

Migliore et al. 2013 (62) conducted a NMA of five osteoporosis therapies to prevent vertebral fractures in post-menopausal osteoporosis that included nine double-blind RCTs with a duration of at least three years. This NMA focused on treatment rankings generated by the Bayesian analysis and found that both zoledronic acid and denosumab had a high probability of being the best treatment. No analysis of non-vertebral fractures of safety outcomes was conducted and relevant bisphosphonates, etidronate and ibandronate, were not included in the analysis.

Generally the findings of the current study are consistent with previously published NMAs with respect to fracture prevention. Small differences in credible intervals that change whether a result is significant or not may be the result of slightly different inclusion criteria (not all the previous studies included RCTs with fracture as a secondary outcome), or the selection of the informative priors used in the construction of the models used in the Bayesian analysis. None of the NMAs analyzed safety outcomes of the different osteoporosis therapies. Based on the analysis of WDAEs in this study, no differences in treatments were observed. However, additional rare and serious adverse events associated with osteoporosis therapies may be important to consider, although these are challenging to analyze in a NMA due to their low event rates and sometimes heterogenous data collection in RCTs.

The value of results obtained from NMAs, are highly dependent on the quality of the individual trials that are included in the analysis. Risk of bias was assessed for all of the included trials in this NMA and found to be high or unclear in the majority of trials. Comparing the efficacy of the

different bisphosphonates in reduction of fracture risk has been hampered by a lack of head-to-head trials using fracture risk as a primary endpoint. (11,63) Some trials have been conducted as open-label studies as there are blinding challenges with the various dosage forms and dosing schedules (some products are administered orally daily, weekly, or monthly, while others are given intravenously yearly or quarterly). Some trials reported fracture incidence as a secondary or safety endpoint or as part of overall adverse event reporting, creating heterogeneity in the data synthesized. The incidence of “non-vertebral” fracture determination from published studies can also be challenging as the definition varies from trial to trial, may or may not include hip fracture, and may not be fully explained in publications.

Strengths and Limitations of This Analysis

The results of this network meta-analysis are robust as the review was performed according to the accepted methodology for systematic reviews and recommendations for the conduct of network meta-analysis. The literature search was comprehensive and included multiple databases and also searched reference lists to locate additional publications. In addition, inclusion and exclusion criteria were explicitly specified, a conservative data analysis using a random effects model was performed, and the results of the sensitivity analyses were consistent with the primary analyses. However, there are several limitations to this type of study that must be considered when interpreting the results. These involve the different doses and routes of administration of the various bisphosphonates, the quality of fracture assessment and classification, the assessment of WDAEs, and the large numbers of withdrawals in the larger and longer-term studies.

Bisphosphonates used for fracture prevention in osteoporosis therapies differ in terms of their dosage, dosing frequency, efficacy across fracture locations and route of administration. Ibandronate was not used in the same doses and route of administration in all the studies that included it as an active comparator. In the RCTs included in this NMA, alendronate was used in two different oral doses, and zoledronic acid was administered intravenously. For ibandronate there was often only one trial per dose and they were small trials. Although there is a substantial population regarding zoledronic acid and denosumab, in both cases, the vast majority of that population is from a single study. The inclusion criteria of this NMA stipulated that trials had to be conducted in women with postmenopausal osteoporosis, but there was a range of ages of study participants and those in primary prevention studies tended to be younger than those in the secondary prevention studies. These differences in the products and populations could have introduced some heterogeneity.

The quality of fracture assessment and classification is another potential source of heterogeneity, particularly in studies where the primary endpoint was not fracture, but BMD or another endpoint. In particular, in many trials, vertebral fractures may be overestimates as they were assessed using radiographs and would likely have remained undetected and asymptomatic in actual clinical practice.. Because of this, numbers needed to treat for vertebral fractures calculated based on this NMA are also likely overestimated. Fink et al (2005) conducted a study finding that 22.6% of radiographic fractures are clinically diagnosed. (64) Therefore, using these estimates, the NNT to prevent a vertebral fracture is more likely in the range of 80 to 147, rather than 18 to 33 as calculated based on the results of this NMA. Non-vertebral fractures may or may not include hip fractures and not all trials classified non-vertebral fractures in the same way. There were a smaller number of trials that reported hip or wrist fractures, limiting the amount of data available for analyses of these outcomes and the results of the hip fracture analysis are dominated by one large study. However, hip fractures have a high patient impact and are the most clinically relevant consequence of osteoporosis. (9)

The subgroup analyses that examined the differences in outcomes depending on trial duration and depending on primary and secondary prevention trials were limited by the number of studies for each therapy in each category ($\geq$ 3 years or < 3 years).

Finally, although the analyses of WDAEs did not reveal any statistically significant differences among the bisphosphonates, the lack of information concerning specific adverse events limits any conclusions about potential differences among the bisphosphonates from a safety perspective. This is frequently a challenge when evaluating results from published RCTs, which are not powered to evaluate specific adverse events. Since their introduction, concerns have arisen regarding the adverse event profile of bisphosphonates, including irritation of the gastrointestinal tract, osteonecrosis of the jaw, esophageal cancer, atrial fibrillation, and most recently atypical fractures such as low energy fractures, fractures of the femur, and pelvic insufficiency fractures. (65) The product monographs for the bisphosphonates marketed in Canada include warning statements regarding these adverse events. (66-69) These safety issues have been identified through clinical trials, post-market surveillance and reporting, as well as observational studies. The safety of bisphosphonates has been discussed in several review articles and specific adverse events have been the subject of observational studies for uncommon and rare adverse events, for example atypical femoral fractures. (70-71) Canadian product monograph warnings for

denosumab include hypocalcemia, hypersensitivity reactions, infections, dermatologic reactions, osteonecrosis of the jaw, atypical femoral fractures, malignancies and suppression of bone turnover.(13) Some of these adverse events are the similar to those noted with bisphosphonates, however, there has not been a comparison of the adverse event profile of the different bisphosphonates with each other or with denosumab, making the benefit-risk evaluations of these drugs challenging. Of note, approximately five years of long-term safety data exists for denosumab, which is less than the approximately ten years experience with bisphosphonates. (72)

Conclusions

Denosumab and zoledronic acid were the most effective therapies in preventing vertebral fracture in postmenopausal osteoporotic women when compared with placebo. Denosumab was more effective than alendronate and risedronate in preventing vertebral fractures. However, for preventing hip fractures, which have greater clinical relevance and patient impact, zoledronic acid and alendronate were the only effective therapies. Although more trials provided data on vertebral fracture than hip fracture, the relevance of these results are limited by the inclusion of asymptomatic, radiographic fractures,

No differences in WDAEs were detected among the various osteoporosis therapies, however, future analyses of specific rare and serious adverse events would help determine if there is a difference in the benefit-harm profile of different therapies, and could assist in making individual treatment recommendations by health care providers.

No advantages of denosumab compared with placebo or other therapies were identified related to preventing non-vertebral, hip or wrist fractures, although the sensitivity of this analysis to detect differences in these outcomes may be low due to the low event rates for these fragility fractures. Denosumab has a different mechanism of action and may have other advantages that physicians and patients wish to consider such as route of administration, dosing regimen, therefore, understanding its comparative effectiveness is an important component of being able to make informed treatment and/or policy decisions.

REFERENCES

1. Osteoporosis Prevention, Diagnosis, and Therapy. NIH Consensus Statement 2000 March 27-29; 17(1): 1–45.
2. World Health Organization. Assessment of fracture risk and its application to screening for postmenopausal osteoporosis: Report of a WHO study group. World Health Organ Tech Rep Ser 1994;843:1-129.
3. Hanley DA, Josse RG. Prevention and management of osteoporosis: consensus statements from the Scientific Advisory Board of the Osteoporosis Society of Canada. 1. Introduction. CMAJ 1996 Oct 1;155(7):921-923.
4. Compston J. Osteoporosis: social and economic impact. Radiol Clin North Am 2010 May;48(3):477-482.
5. Eastell R, Cedel SL, Wahner HW, Riggs BL, Melton LJ,3rd. Classification of vertebral fractures. J Bone Miner Res 1991 Mar;6(3):207-215.
6. Brown JP, Josse RG, Scientific Advisory Council of the Osteoporosis Society of Canada. 2002 clinical practice guidelines for the diagnosis and management of osteoporosis in Canada. CMAJ 2002 Nov 12;167(10 Suppl):S1-34.
7. Guggenbuhl P. Osteoporosis in males and females: Is there really a difference? Joint Bone Spine 2009 Dec;76(6):595-601.
8. Ioannidis G, Papaioannou A, Hopman WM, Akhtar-Danesh N, Anastassiades T, Pickard L, et al. Relation between fractures and mortality: results from the Canadian Multicentre Osteoporosis Study. CMAJ 2009 Sep 1;181(5):265-271.
9. Chrischilles EA, Butler CD, Davis CS, Wallace RB. A model of lifetime osteoporosis impact. Arch Intern Med 1991 Oct;151(10):2026-2032.
10. Warriner AH, Saag KG. Osteoporosis diagnosis and medical treatment. Orthop Clin North Am 2013 Apr;44(2):125-135.
11. Boonen S, Laan RF, Barton IP, Watts NB. Effect of osteoporosis treatments on risk of non-vertebral fractures: Review and meta-analysis of intention-to-treat studies. 2005;16 (10):1291-1298.
12. Bianchi G, Sambrook P. Oral nitrogen-containing bisphosphonates: a systematic review of randomized clinical trials and vertebral fractures. Curr Med Res Opin 2008 Sep;24(9):2669-2677.
13. Health Canada. Product monographs,denusomab. Available from: http://webprod5.hc-sc.gc.ca/dpd-bdpp/dispatch-repartition.do?lang=eng. Accessed March 19, 2014.
14. Wells GA, Cranney A, Peterson J, Boucher M, Shea B, Robinson V, et al. Alendronate for the primary and secondary prevention of osteoporotic fractures in postmenopausal women. Cochrane Database Syst Rev 2008 Jan 23;(1)(1):CD001155.
15. Wells GA, Cranney A, Peterson J, Boucher M, Shea B, Robinson V, et al. Etidronate for the primary and secondary prevention of osteoporotic fractures in postmenopausal women. Cochrane Database Syst Rev 2008 Jan 23;(1)(1):CD003376.
16. Wells G, Cranney A, Peterson J, Boucher M, Shea B, Robinson V, et al. Risedronate for the primary and secondary prevention of osteoporotic fractures in postmenopausal women. Cochrane Database Syst Rev 2008 Jan 23;(1)(1):CD004523.
17. Bucher HC, Guyatt GH, Griffith LE, Walter SD. The results of direct and indirect treatment comparisons in meta-analysis of randomized controlled trials. J Clin Epidemiol 1997 Jun;50(6):683-691.
18. Lu G, Ades AE. Combination of direct and indirect evidence in mixed treatment comparisons. Stat Med 2004 Oct 30;23(20):3105-3124.
19. Higgins JPT, Green S (editors). Cochrane handbook for systematic reviews of interventions Version 5.1.0 [updated March 2011]. The Cochrane Collaboration, 2011. Available from www.cochracne-handbook.org.
20. Turner RM, Davey J, Clarke MJ, Thompson SG, Higgins JP. Predicting the extent of heterogeneity in meta-analysis, using empirical data from the Cochrane Database of Systematic Reviews. Int J Epidemiol 2012 Jun;41(3):818-827.

21. Turner RM, Davey J, Clarke MJ, Thompson SG, Higgins JP. Predicting the extent of heterogeneity in meta-analysis, using empirical data from the Cochrane Database of Systematic Reviews. Int J Epidemiol 2012 Jun;41(3):818-827.
22. Cummings SR, San Martin J, McClung M, Siris E, Eastell R, Reid IR et al. for the FREEDOM Trial. Denosumab for prevention of fractures in postmenopausal women with osteoporosis. N Engl J Med, 2009, 361: 756-765.
23. Miller PD, Epstein S, Sedarati F, Reginster JY. Once-monthly oral ibandronate compared with weekly oral alendronate in postmenopausal osteoporosis: results from the head-to-head MOTION study. Curr Med Res Opin 2008 Jan;24(1):207-213.
24. Chesnut III CH, Skag A, Christiansen C, Recker R, Stakkestad JA, Hoiseth A, et al. Effects of oral ibandronate administered daily or intermittently on fracture risk in postmenopausal osteoporosis. 2004;19(8):1241-1249.
25. McClung MR, Bolognese MA, Sedarati F, Recker RR, Miller PD. Efficacy and safety of monthly oral ibandronate in the prevention of postmenopausal bone loss. Bone 2009 Mar;44(3):418-422.
26. Black DM, Delmas PD, Eastell R, Reid IR, Boonen S, Cauley JA, et al. Once-yearly zoledronic acid for treatment of postmenopausal osteoporosis. N Engl J Med 2007 May 3;356(18):1809-1822.
27. Bone HG, Downs RW,Jr, Tucci JR, Harris ST, Weinstein RS, Licata AA, et al. Dose-response relationships for alendronate treatment in osteoporotic elderly women. Alendronate Elderly Osteoporosis Study Centers. J Clin Endocrinol Metab 1997 Jan;82(1):265-274.
28. Fogelman I, Ribot C, Smith R, Ethgen D, Sod E, Reginster JY. Risedronate reverses bone loss in postmenopausal women with low bone mass: results from a multinational, double-blind, placebo-controlled trial. BMD-MN Study Group. J Clin Endocrinol Metab 2000 May;85(5):1895-1900.
29. Harris ST, Watts NB, Genant HK, McKeever CD, Hangartner T, Keller M, et al. Effects of risedronate treatment on vertebral and nonvertebral fractures in women with postmenopausal osteoporosis: a randomized controlled trial. Vertebral Efficacy With Risedronate Therapy (VERT) Study Group. JAMA 1999 Oct 13;282(14):1344-1352.
30. Hooper MJ, Ebeling PR, Roberts AP, Graham JJ, Nicholson GC, D'Emden M, et al. Risedronate prevents bone loss in early postmenopausal women: a prospective randomized, placebo-controlled trial. Climacteric 2005 Sep;8(3):251-262.
31. Meunier PJ, Confavreux E, Tupinon I, Hardouin C, Delmas PD, Balena R. Prevention of early postmenopausal bone loss with cyclical etidronate therapy (a double-blind, placebo-controlled study and 1-year follow-up). J Clin Endocrinol Metab 1997 Sep;82(9):2784-2791.
32. Montessori ML, Scheele WH, Netelenbos JC, Kerkhoff JF, Bakker K. The use of etidronate and calcium versus calcium alone in the treatment of postmenopausal osteopenia: results of three years of treatment. Osteoporos Int 1997;7(1):52-58.
33. Pouilles JM, Tremollieres F, Roux C, Sebert JL, Alexandre C, Goldberg D, et al. Effects of cyclical etidronate therapy on bone loss in early postmenopausal women who are not undergoing hormonal replacement therapy. Osteoporos Int 1997;7(3):213-218.
34. Reginster J, Minne HW, Sorensen OH, Hooper M, Roux C, Brandi ML, et al. Randomized trial of the effects of risedronate on vertebral fractures in women with established postmenopausal osteoporosis. Vertebral Efficacy with Risedronate Therapy (VERT) Study Group. Osteoporos Int 2000;11(1):83-91.
35. Watts NB, Harris ST, Genant HK, Wasnich RD, Miller PD, Jackson RD, et al. Intermittent cyclical etidronate treatment of postmenopausal osteoporosis. N Engl J Med 1990 Jul 12;323(2):73-79.
36. Wimalawansa SJ. A four-year randomized controlled trial of hormone replacement and bisphosphonate, alone or in combination, in women with postmenopausal osteoporosis. Am J Med 1998 Mar;104(3):219-226.
37. McClung MR, Geusens P, Miller PD, Zippel H, Bensen WG, Roux C, et al. Effect of risedronate on the risk of hip fracture in elderly women. Hip Intervention Program Study Group. N Engl J Med 2001 Feb 1;344(5):333-340.
38. Pols HA, Felsenberg D, Hanley DA, Stepan J, Munoz-Torres M, Wilkin TJ, et al. Multinational, placebo-controlled, randomized trial of the effects of alendronate on bone

density and fracture risk in postmenopausal women with low bone mass: results of the FOSIT study. Fosamax International Trial Study Group. Osteoporos Int 1999;9(5):461-468.
39. Storm T, Thamsborg G, Steiniche T, Genant HK, Sorensen OH. Effect of intermittent cyclical etidronate therapy on bone mass and fracture rate in women with postmenopausal osteoporosis. N Engl J Med 1990 May 3;322(18):1265-1271.
40. Valimaki MJ, Farrerons-Minguella J, Halse J, Kroger H, Maroni M, Mulder H, et al. Effects of risedronate 5 mg/d on bone mineral density and bone turnover markers in late-postmenopausal women with osteopenia: a multinational, 24-month, randomized, double-blind, placebo-controlled, parallel-group, phase III trial. Clin Ther 2007 Sep;29(9):1937-1949.
41. Black DM, Cummings SR, Karpf DB, Cauley JA, Thompson DE, Nevitt MC, et al. Randomised trial of effect of alendronate on risk of fracture in women with existing vertebral fractures. Fracture Intervention Trial Research Group. Lancet 1996 Dec 7;348(9041):1535-1541.
42. Hosking D, Chilvers CE, Christiansen C, Ravn P, Wasnich R, Ross P, et al. Prevention of bone loss with alendronate in postmenopausal women under 60 years of age. Early Postmenopausal Intervention Cohort Study Group. N Engl J Med 1998 Feb 19;338(8):485-492.
43. Yan Y, Wang W, Zhu H, Li M, Liu J, Luo B, et al. The efficacy and tolerability of once-weekly alendronate 70 mg on bone mineral density and bone turnover markers in postmenopausal Chinese women with osteoporosis. J Bone Miner Metab 2009;27(4):471-478.
44. Li Y, Zhang Z, Deng X, Chen, L. Efficacy and safety of risedronate sodium in treatment of postmenopausal osteoporosis. J Huazhong Univ Sci Technolog Med Sci 2005;25(5):527-529.
45. Fogelman I, Ribot C, Smith R, Ethgen D, Sod E, Reginster JY. Risedronate reverses bone loss in postmenopausal women with low bone mass: results from a multinational, double-blind, placebo-controlled trial. BMD-MN Study Group. J Clin Endocrinol Metab 2000 May;85(5):1895-1900.
46. McClung MR, Geusens P, Miller PD, Zippel H, Bensen WG, Roux C, et al. Effect of risedronate on the risk of hip fracture in elderly women. Hip Intervention Program Study Group. N Engl J Med 2001 Feb 1;344(5):333-340.
47. Ascott-Evans BH1, Guanabens N, Kivinen S, Stuckey BG, Magaril CH, Vandormael K, Stych B, Melton ME. Alendronate prevents loss of bone density associated with discontinuation of hormone replacement therapy: a randomized controlled trial. Arch Intern Med. 2003 Apr 14;163(7):789-94.
48. Cummings SR, Black DM, Thompson DE, Applegate WB, Barrett-Connor E, Musliner TA, et al. Effect of alendronate on risk of fracture in women with low bone density but without vertebral fractures: results from the Fracture Intervention Trial. JAMA 1998 Dec 23-30;280(24):2077-2082.
49. Herd RJ, Balena R, Blake GM, Ryan PJ, Fogelman I. The prevention of early postmenopausal bone loss by cyclical etidronate therapy: a 2-year, double-blind, placebo-controlled study. Am J Med 1997 Aug;103(2):92-99.
50. Greenspan SL, Parker RA, Ferguson L, Rosen HN, Maitland-Ramsey L, Karpf DB. Early changes in biochemical markers of bone turnover predict the long-term response to alendronate therapy in representative elderly women: a randomized clinical trial. J Bone Miner Res 1998 Sep;13(9):1431-1438.
51. Greenspan S, Field-Munves E, Tonino R, Smith M, Petruschke R, Wang L, et al. Tolerability of once-weekly alendronate in patients with osteoporosis: a randomized, double-blind, placebo-controlled study. Mayo Clin Proc 2002 Oct;77(10):1044-1052.
52. Fitzpatrick LA, Dabrowski CE, Cicconetti G, Gordon DN, Papapoulos S, Bone III HG, and Bilezikian JP. The effects of ronacaleret, a calcium-sensing receptor antagonist, on bone mineral density and biochemical markers of bone turnover in postmenopausal women with low bone mineral density. J Clin Endocrinol Metab 2011 Aug;96(8):2441-2449
53. Lyritis GP, Tsakalakos N, Paspati I, Skarantavos G, Galanos A, Androulakis C. The effect of a modified etidronate cyclical regimen on postmenopausal osteoporosis: a four-year study. Clin Rheumatol 1997 Jun;16(4):354-360.

54. Li M, Xing X, Zhang Z, Liu J, Zhang Z, Liu D, et al. Infusion of ibandronate once every 3 months effectively decreases bone resorption markers and increases bone mineral density in Chinese postmenopausal osteoporotic women: a 1-year study. 2010;28(3):299-305.
55. Qin L, Choy W, Au S, Fan M, Leung P. Alendronate increases BMD at appendicular and axial skeletons in patients with established osteoporosis. J Orthop Surg Res 2007 May 21;2:9.
56. Ishida Y, Kawai S. Comparative efficacy of hormone replacement therapy, etidronate, calcitonin, alfacalcidol, and vitamin K in postmenopausal women with osteoporosis: The Yamaguchi Osteoporosis Prevention Study. Am J Med 2004 Oct 15;117(8):549-555.
57. Grey A, Bolland M, Wong S, Horne A, Gamble G, Reid IR. Low-dose zoledronate in osteopenic postmenopausal women: A randomized controlled trial. 2012;97(1):286-292.
58. 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 Feb;40(4):275-84.e1-2.
59. Hopkins RB, Goeree R, Pullenayegum E, Adachi JD, Papaioannou A, Xie F, et al. The relative efficacy of nine osteoporosis medications for reducing the rate of fractures in post-menopausal women. 2011;12(209).
60. Freemantle N, Cooper C, Diez-Perez A, Gitlin M, Radcliffe H, Shepherd S, et al. Results of indirect and mixed treatment comparison of fracture efficacy for osteoporosis treatments: a meta-analysis. Osteoporos Int 2013 Jan;24(1):209-217.
61. Murad MH, Drake MT, Mullan RJ, Mauck KF, Stuart LM, Lane MA et al. Comparative effectiveness of drug treatmetns to prevent fragility fractures: a systematic review and network meta-analysis. J Clin Endocrinol Metab, 2012, 97: 1871-1880.
62. Migliore A, Broccoli S, Massafra U, Cassol M, Frediani B. Ranking antireabsorptive agents to prevent vertebral fractures in postmenopausal osteoporosis by mixed treatment comparison meta-analysis. Eur Rev Med Pharmacol Sci, 2013, 17: 658-667.
63. Rizzoli R. Bisphosphonates for post-menopausal osteoporosis: are they all the same? 2011;104(4):281-300.
64. Fink HA, Milavetz DL, Palermo L, Nevitt MC, Cauley JA, Genant HK, Black DM, Ensrud KE; Fracture Intervention Trial Research Group. What proportion of incident radiographic vertebral deformities is clinically diagnosed and vice versa? J Bone Miner Res. 2005 Jul;20(7):1216-22.
65. Watts NB, Diab DL. Long-term use of bisphosphonates in osteoporosis. J Clin Endocrinol Metab 2010 Apr;95(4):1555-1565.
66. Health Canada. Product monographs, alendronate. Available from: http://webprod5.hc-sc.gc.ca/dpd-bdpp/dispatch-repartition.do?lang=eng. Accessed March 19, 2014.
67. Health Canada. Product monographs, etidronate. Available from: http://webprod5.hc-sc.gc.ca/dpd-bdpp/dispatch-repartition.do?lang=eng. Accessed March 19, 2014.
68. Health Canada. Product monographs, risedronate. Available from: http://webprod5.hc-sc.gc.ca/dpd-bdpp/dispatch-repartition.do?lang=eng. Accessed March 19, 2014.
69. Health Canada. Product monographs, zoledronic acid. Available from: http://webprod5.hc-sc.gc.ca/dpd-bdpp/dispatch-repartition.do?lang=eng. Accessed March 19, 2014.
70. Schilcher J, Michaelsson K, Aspenberg P. Bisphosphonate use and atypical fractures of the femoral shaft. N Eng J Med 2011 May 5;364(18):1728-1737.
71. Saleh A, Hegde VV, Potty AG, Lane JM. Bisphosphonate therapy and atypical fractures. Orthop Clin North Am 2013 Apr;44(2):137-151.
72. Reginster JY, Pelousse F, Bruyère O. Safety concerns with the long-term management of osteoporosis. Expert Opin Drug Saf. 2013 Jul;12(4):507-2

TABLES AND FIGURES

Figure 1. PRISMA Flow Diagram for Network Meta-Analysis of Osteoporosis Therapies

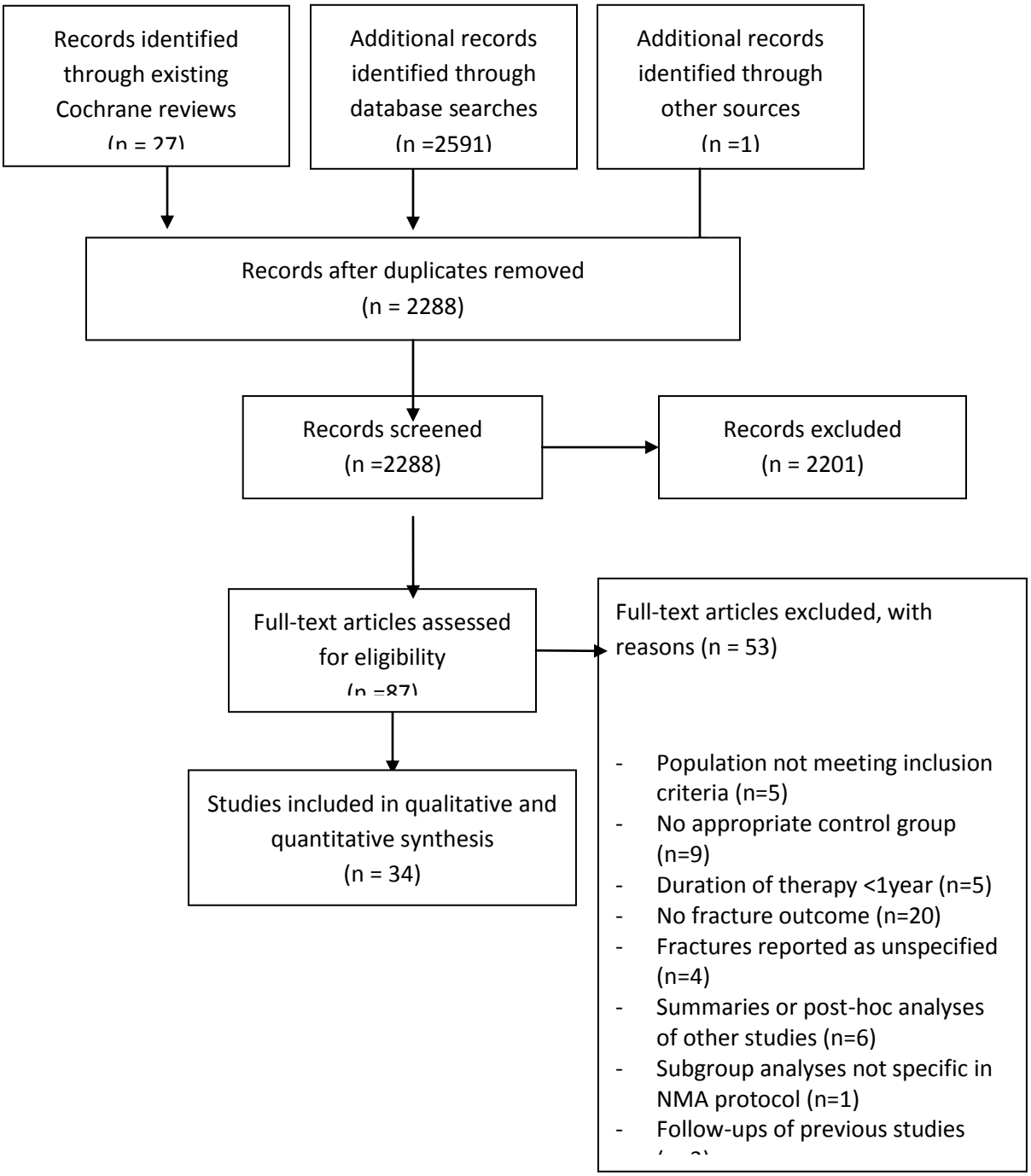


Figure 2. Evidence network for vertebral fractures.

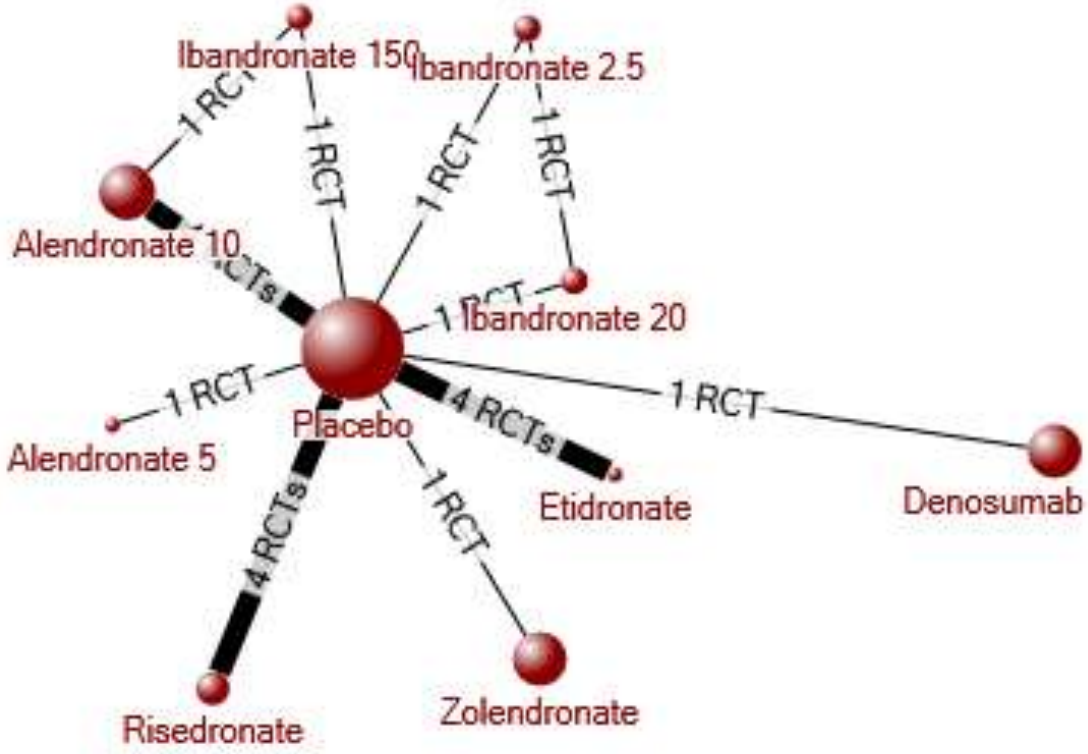


Table 1. Network Meta-Analysis Results: Vertebral fractures. Odds Ratio (95% Credible Interval)

Treatment	Placebo	ALN5	ALN10	RIS	ETI	IBA150	IBA2.5	IBA20	ZOL	DEN
Placebo	--	--	--	--	--	--	--	--	--	--
ALE5	0.77 (0.15,2.35)	--	--	--	--	--	--	--	--	--
ALE10	**0.53 (0.40,0.70)**	1.14 (0.22,3.68)	--	--	--	--	--	--	--	--
RIS	**0.49 (0.23,0.91)**	1.05 (0.16,3.62)	1.14 (0.75,1.66)	--	--	--	--	--	--	--
ETI	0.65 (0.13,1.98)	1.40 (0.13,6.20)	0.94 (0.51,1.57)	0.84 (0.36,1.63)	--	--	--	--	--	--
IBA150	**0.49 (0.29,0.78)**	1.05 (0.19,3.51)	1.21 (0.25,3.64)	1.1 (0.22,3.44)	1.48 (0.25,5.09)	--	--	--	--	--
IBA2.5	**0.52 (0.32,0.82)**	1.11 (0.20,3.75)	0.94 (0.51,1.57)	0.84 (0.46,1.40)	1.13 (0.45,2.46)	1.23 (0.22,3.89)	--	--	--	--
IBA20	**0.60 (0.45,0.78)**	1.27 (0.25,4.08)	1.00 (0.55,1.66)	0.89 (0.49,1.48)	1.20 (0.47,2.57)	1.30 (0.23,4.09)	1.09 (0.63,1.80)	--	--	--
ZOL	**0.28 (0.19,0.39)**	0.61 (0.11,1.97)	**0.54 (0.33,0.82)**	0.48 (0.30,0.73)	0.65 (0.28,1.33)	0.71 (0.13,2.17)	0.61 (0.32,1.05)	**0.57 (0.30,0.98)**	--	--
DEN	**0.31 (0.22,0.44)**	0.67 (0.13,2.19)	**0.6 (0.37,0.91)**	0.54 (0.33,0.82)	0.72 (0.31,1.48)	0.79 (0.15,2.42)	0.68 (0.35,1.18)	0.64 (0.34,1.10)	1.15 (0.67,1.83)	--

Note: ALE5, alendronate 5 mg/day po (35 mg/week); ALE10, alendronate 10 mg/day po (70 mg/week); ETI, 400 mg/month po; IBA150, ibandronate 150 mg/month po; IBA2.5, 2.5 mg/day po; IBA20, 20 mg/day po intermittently; RIS, risedronate 150 mg/month po; ZOL, zoledronic acid 5 mg/year iv. DEN, denosumab 60 mg/6months SC. Statistically significant results are in bold font.

Table 2. Network meta-analysis results: non-vertebral fractures. Odds Ratio (95% Credible Interval)

Treatment	Placebo	IBA2	IBA150	ALE10	ALE5	ETI	RIS	IBA2.5	IBA20	ZOL	DEN
Placebo	--	--	--	--	--	--	--	--	--	--	--
IBA2	4.40 (0.013,25.7)	--	--	--	--	--	--	--	--	--	--
IBA150	1.01 (0.41,2.09)	16.3 (0.34,72.8)	--	--	--	--	--	--	--	--	--
ALE10	**0.78 (0.60,0.96)**	12.9 (0.30,61.3)	0.91 (0.38,1.85)	--	--	--	--	--	--	--	--
ALE5	1.06 (0.57,1.79)	16.8 (0.036,80.9)	1.24 (0.40,2.92)	1.37 (0.71,2.44)	--	--	--	--	--	--	--
ETI	0.68 (0.39,1.11)	10.7 (0.024,55.3)	0.80 (0.27,1.88)	0.89 (0.48,1.50)	0.70 (0.30,1.41)	--	--	--	--	--	--
RIS	**0.74 (0.57,0.91)**	12.3 (0.029,60.0)	0.88 (0.34,1.85)	0.97 (0.70,1.31)	0.77 (0.39,1.34)	1.17 (0.64,1.97)	--	--	--	--	--
IBA2.5	1.11 (0.71,1.67)	18.1 (0.042,84.4)	1.31 (0.46,2.97)	1.44 (0.87,2.33)	1.14 (0.52,2.19)	1.74 (0.85,3.25)	1.51 (0.92,2.43)	--	--	--	--
IBA20	1.14 (0.73,1.70)	18.8 (0.043,88.2)	1.33 (0.48,3.05)	1.48 (0.90,2.38)	1.17 (0.54,2.25)	1.78 (0.87,3.32)	1.54 (0.95,2.48)	1.05 (0.67,1.56)	--	--	--
ZOL	0.74 (0.53,1.02)	12.2 (0.029,58.7)	0.87 (0.33,1.92)	0.97 (0.65,1.45)	0.76 (0.38,1.40)	1.16 (0.61,2.06)	1.01 (0.68,1.53)	0.70 (0.40,1.16)	0.68 (0.39,1.13)	--	--
DEN	0.80 (0.56,1.10)	13.3 (0.030,62.4)	0.94 (0.35,2.06)	1.04 (0.69,1.54)	0.83 (0.40,1.51)	1.26 (0.65,2.21)	1.09 (0.72,1.63)	0.76 (0.42,1.25)	0.74 (0.41,1.21)	1.11 (0.67,1.72)	--

Note: ALE5, alendronate 5 mg/day po (35 mg/week); ALE10, alendronate 10 mg/day po (70 mg/week); ETI, 400 mg/month po; IBA2, ibandronate 2 mg/every 3 months iv; IBA150, ibandronate 150 mg/month po; IBA2.5, 2.5 mg/day po; IBA20, 20 mg/day po intermittently; RIS, risedronate 150 mg/month po; ZOL, zoledronic acid 5 mg/year iv. DEN, denosumab 60 mg/6months SC. Statistically significant results are in bold font.

Table 3. Network meta-analysis results: Hip fractures. Odds Ratio (95% Credible Interval)

Treatment	Placebo	ALN10	ETI	RIS	ZOL	DEN
Placebo	--	--	--	--	--	--
ALE10	**0.55 (0.27,0.98)**	--	--	--	--	--
ETI	1.52 (0.16,5.95)	3.09 (0.28,12.7)	--	--	--	--
RIS	0.75 (0.51,1.09)	1.53 (0.67,3.05)	1.16 (0.12,4.75)	--	--	--
ZOL	**0.60 (0.36,0.94**)	1.22 (0.50,2.55)	0.92 (0.093,3.83)	0.82 (0.42,1.45)	--	--
DEN	0.62 (0.33,1.08)	1.27 (0.48,2.74)	0.96 (0.092,4.08)	0.86 (0.40,1.62)	1.11 (0.48,2.21)	--

Note: ALE10, alendronate 10 mg/day po (70 mg/week); ETI, 400 mg/month po;

RIS, risedronate 150 mg/month po; ZOL, zoledronic acid 5 mg/year iv. DEN, denosumab 60 mg/6months SC. Statistically significant results are in bold font

Table 4. Network meta-analysis results: Withdrawals due to Adverse Events. Odds Ratio (95% Credible Interval)

Treatment	Placebo	ALN10	ALN5	IBA150	ETI	RIS	IBA2.5	IBA20	ZOL	DEN
Placebo	--	--	--	--	--	--	--	--	--	--
ALE10	0.67 (0.36,1.14)	--	--	--	--	--	--	--	--	--
ALE5	1.17 (0.37,2.78)	1.91 (0.49,5.06)	--	--	--	--	--	--	--	--
IBA150	1.24 (0.34,3.40)	1.91 (0.57,5.03)	1.39 (0.23,4.86)	--	--	--	--	--	--	--
ETI	1.15 (0.40,2.58)	1.87 (0.53,4.73)	1.27 (0.25,3.94)	1.29 (0.22,4.10)	--	--	--	--	--	--
RIS	0.96 (0.54,1.61)	1.57 (0.65,3.20)	1.07 (0.29,2.87)	1.08 (0.24,3.01)	1.05 (0.31,2.66)	--	--	--	--	--
IBA2.5	1.13 (0.26,3.23)	1.84 (0.36,5.73)	1.27 (0.18,4.54)	1.27 (0.16,4.57)	1.24 (0.19,4.29)	1.27 (0.25,3.90)	--	--	--	--
IBA20	1.06 (0.25,3.03)	1.73 (0.34,5.31)	1.19 (0.17,4.27)	1.19 (0.14,4.30)	1.16 (0.18,4.00)	1.19 (0.23,3.68)	1.15 (0.27,3.28)	--	--	--
ZOL	1.42 (0.33,4.03)	2.31 (0.45,7.13)	1.59 (0.23,5.69)	1.60 (0.20,5.72)	1.54 (0.24,5.30)	1.59 (0.32,4.87)	1.90 (0.21,7.31)	2.03 (0.23,7.81)	--	--
DEN	1.41 (0.32,4.06)	2.30 (0.44,7.15)	1.58 (0.22,5.72)	1.59 (0.19,5.80)	1.54 (0.23,5.37)	1.59 (0.31,4.90)	1.91 (0.21,7.42)	2.02 (0.22,7.95)	1.52 (1.67,5.98)	--

Note: ALE5, alendronate 5 mg/day po (35 mg/week); ALE10, alendronate 10 mg/day po (70 mg/week); ETI, 400 mg/month po; IBA150, ibandronate 150 mg/month po; IBA2.5, 2.5 mg/day po; IBA20, 20 mg/day po intermittently; RIS, risedronate 150 mg/month po; ZOL, zoledronic acid 5 mg/year iv. DEN, denosumab 60 mg/6months SC

SUPPLEMENTARY MATERIAL 1

Supplementary Table 1. Patient and Trial Characteristics of Studies Included in NMA.

Study	PATIENT CHARACTERISTICS					TRIAL CHARACTERISTICS			
	Age (years), mean	Years since menopause	Prevalent Vertebral Fracture (%)	BMD (g/cm2)	T-score	Dose	Prevention (1° or 2°)	Location	Trial Duration (years)
Ascott Evans 2003	67.3 (6.6)	11.5 (7.3)	0	NR	Lumbar spine -2.3 (0.7)	AL 10mg/d	1	9 countries	NR
Black 1996	71.0 (5.6)	NR	100	Lumbar spine 0.790 (0.140) Femoral neck 0.565(0.070)	NR	AL 5mg/d for 2 yrs then 10mg/d for 1 yr	NR	USA	3
Black 2007	73.0 (5.4) P 73.1 (5.34) Z	NR	64.2 P 62.4 Z	Lumbar spine 0.79(0.140) P 0.79 (0.124) Z Total hip, femoral neck reported	% femoral neck <2.5 70.8 P , 72.6 Z -2.5 to 1.5 27.8 P, 25.9 Z >1.5 0.4P, 0.6Z	ZA 5mg/yr	2	Europe, NA, Oceania, Latin Am, Asia	3
Bone 1997	70.7 (5.9)	23.5 (9.6)	37.3	Lumbar spine 0.715(0.080)	NR	AL 1mg/d AL 2mg/d AL 5mg/d	2	NR	2
Chestnut 2004	69(6) P 69(6) I2.5 69(6)I20	20.8(7.8) P 20.9(8.0) I2.5 20.8(8.0) I20	100% (inclusion)	NR	Lumbar spine -2.8(0.9) -2.8(0.9) -2.7(0.9) Total hip -1.7(0.9) -1.7(0.8) -1.7(0.9) femoral neck also reported	IBD 2.5mg (n=982) IBD 20mg eod, 12 doses q3mos (n=982) PL (n=982)	2	Europe, NA	3

Study	PATIENT CHARACTERISTICS					TRIAL CHARACTERISTICS			
	Age (years), mean	Years since menopause	Prevalent Vertebral Fracture (%)	BMD (g/cm2)	T-score	Dose	Prevention (1° or 2°)	Location	Trial Duration (years)
Cummings 1998	67.6 (6.2)	NR	0	Posterior-anterior spine 0.842(0.130)	NR	AL 5mg/d, 2 yrs then 10mg for 2 yrs	1	USA	4
Cummings 2009	72.3 (5.2) D 72.3 (5.2) P	NR	23.8 D 23.4 P	NR	Lumbar spine -2.82(0.70) D -2.84(0.70) P total hip -1.89(0.81) D -1.91(0.81) P femoral neck -2.15(0.72) -2.17(0.71)	D 60mg q6mos	2	Europe, Latin Am, NA, Australia, NZ	3
Fitzpatrick 2011	64.2 (5.9)	Age at menopause 47.7(4.7)	42.2	NR	Lumbar spine -2.6(0.6) Hip -1.6(0.6)	AL 70mg/week	2	14 countries	6 mos (stopped early)
Fogelman 2000	64.7 (7.2)	17.7 (9.4)	29	Lumbar spine 0.740 (0.842) femoral neck 0.633 (0.937)	Lumbar spine -2.9 (0.7)	R 2.5mg/d (disc early) R 5mg/d	2	Europe	2
Greenspan 1998	70.0 (4.6)	NR	NR	Lumbar spine 0.991(0.105) femoral neck 0.638(0.092)	NR	AL 5mg/d for 1.5yrs then 10mg/d for 1 yr	2	USA	2.5
Greenspan 2002	78.5 (range 65-91)	NR	55 (history any fracture)	NR	Range at hip and spine: -3.5 to -2.4	AL 10mg/d	2	USA	2

Study	PATIENT CHARACTERISTICS					TRIAL CHARACTERISTICS			
	Age (years), mean	Years since menopause	Prevalent Vertebral Fracture (%)	BMD (g/cm2)	T-score	Dose	Prevention (1° or 2°)	Location	Trial Duration (years)
Grey 2012	67(8) P 62(8) Z	Late menopause	?	Lumbar spine 1.04(0.08) g/cm P 1.06(0.09) Z total hip, total body reported	Lumbar spine -1.2(0.7) P -1.0 (0.8) Z total hip reported	ZA 1mg (n=45) ZA 2.5 mg (n=45) ZA 5mg (n=45) given once	1	?	3+2yr extension
Harris 1999	68.7 (7.3)	24 (9.9)	80	Lumbar spine 0.833 (0.162) femoral neck 0.597 (0.103)	Lumbar spine -2.4 (1.4) femoral neck -2.6 (1.1)	R 2.5mg/d (disc early) R 5mg/d	2	NA	3
Herd 1997	54.8 (5.0)	5.5 (2.9)	0	Lumbar spine 0.835 (0.075) femoral neck 0.700 (0.090)	NR	E 400mg/d, 14d	1	UK	2
Hooper 2005	52.7 (3.2)	3.9 (5.3)	Prevalent vertebral deformities 18.3	Lumbar spine 1.077 (0.120) femoral neck 0.770 (0.100)	Lumbar spine -0.4 (0.1)	R 2.5 mg/d R 5 mg/d	1	Australia	2
Hosking 1998	53.3 (4.0)	5.9 (5.2)	NR	Lumbar spine 0.939 (0.129) Hip 0.846(0.166)	NR	AL 2.5mg/d 5mg/d	1	USA, UK, Denmark	2
Ishida 2004	69.3 (13.3)	19.8	100	Distal 1/3 radius 0.440 (0.073	NR	E 200mg/d 2wk	2	Japan	2
Li 2005	NR	NR	NR	NR	NR	R 5mg/d	2	China	1
Li 2010	65.1(7.3)	48.1 (3.4)	11.4	Lumbar spine 0.801(0.963) femoral neck 0.663(0.798)	NR	AL 70mg/w (n=79) IBD 2mg (n=79)	2	China	1
Lyritis 1997	72.0 (0.4)	25.8 (1.7)	100	Lumbar spine 0.569 (0.026) femoral neck 0.422 (0.032)	NR	E 400mg/d for 20d	2	Greece	4
McClung 2001	78.0 (9.7)	31.8 (19.3)	40.4	NR	Femoral neck -3.7 (0.6)	R 2.5 or 5 mg/d* (data NR separately)	2	NA, Europe, NZ, Australia	3

Study	PATIENT CHARACTERISTICS					TRIAL CHARACTERISTICS			
	Age (years), mean	Years since menopause	Prevalent Vertebral Fracture (%)	BMD (g/cm2)	T-score	Dose	Prevention (1° or 2°)	Location	Trial Duration (years)
McClung 2009	53.4 (3.8) P 53.7 (3.6) R	0.5-3 yrs, n(%) 49 (59.0) P 49 (63.6) R >3 yrs 34 (41.0) 28 (36.4)	0% (excluded)	Lumbar spine 0.906 (0.041) P 0.905 (0.042) R Total hip, trochanter, femoral hip also reported	Lumbar spine -1.6 (0.4) P -1.6 (0.4) R total hip, trochanter, femoral hip also reported	IBD 150mg/mo (n=77) PL (n=83)	1	USA	1
Menuier 1997	52.7 (0.5)	2.4 (1.8)	NR	Lumbar spine 0.900 (0.020) femoral neck 0.725 (0.016)	NR	E 400mg/d for 2wk	1	France	2yr tx, 1yr FU
Miller 2008	65.6	18.4	Prevalent NR Previous: 38.6	Lumbar spine 0.785(0.076) total hip 0.776(0.103)	Lumbar spine -3.2 total hip -1.7	AL 70mg/w (n=873) IBD 150mg (n=877)	2	NA, Latin Am, Europe, South Africa	1
Montessori 1997	62.5 (6.2)	14.9 (6.1)	35	Lumbar spine 0.674 (0.080) femoral neck 0.603 (0.083)	35	E 400mg/d for 14d	2	Netherlands	3
Pols 1999	62.8 (7.4)	15.9 (8.4)	NR	NR	NR	AL 10mg/d	2	Europe, Latin Am, Australia, Canada, South Africa, China	1
Pouilles 1997	53.8 (3.1)	2.6 (1.4)	NR	NR	NR	E 400mg/d for 14d	1	France	2
Qin 2007	59.8 (6.3)	10.1 (6.6)	NR	Lumbar spine 0.719 (0.099) femoral neck 0.641 (0.094)	Lumbar spine -2.2 (0.8)	AL 10mg/d	2	China	1

Study	PATIENT CHARACTERISTICS					TRIAL CHARACTERISTICS			
	Age (years), mean	Years since menopause	Prevalent Vertebral Fracture (%)	BMD (g/cm2)	T-score	Dose	Prevention (1° or 2°)	Location	Trial Duration (years)
Reginster 2000	71.0 (7.0)	24.7 (8.5)	100	Lumbar spine 0.785 (0.150) femoral neck 0.577 (0.099)	Lumbar spine -2.7 (1.4)	R 2.5mg/d* (disc early) R 5mg/d	2	Europe, Australia	3
Storm 1990	68.4 (0.9)	21.6 (1.3)	100	NR	NR	E 400mg/d, 2wk	2	Denmark	3
Valimaki 2007	65.9 (6.8)	18.3 (7.9)	NR	NR	Lumbar spine -1.8 (0.4) proximal femoral -1.2 (0.6)	R 5mg/d	1	Europe	2
Watts 1990	65.1 (0.6)	17.9 (0.9)	100	Lumbar spine 0.870 (0.020) femoral neck 0.675 (0.010)	Lumbar spine -2.7 (0.1)	E 400mg/d, 14d	2	USA	2
Wimalwansa 1998	64.9 (0.9)	15.1 (0.8)	100	Lumbar spine 0.825 (0.013) femoral neck 0.663 (0.046)	NR	E 400mg/d, 14d	2	USA or UK	4
Yan 2009	64.9 (6.2)	15.3 (6.8)	0	Lumbar spine 0.835 (0.100) femoral neck 0.705 (0.064)	NR	AL 70mg/w	2	China	1

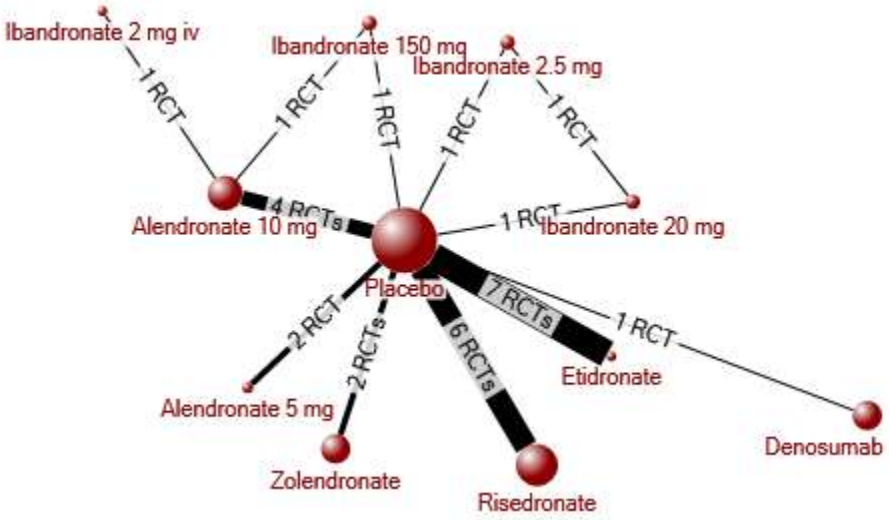
Abbreviations: AL=alendronate; BMD=bone mass density; D=denosumab; E=etidronate; IBD=ibandronate; NR=not reported; P=placebo; R=risedronate; ZA=zoledronic acid

* Head to head trials comparing different osteoporosis therapies; all other trials included a placebo-control group unless otherwise noted

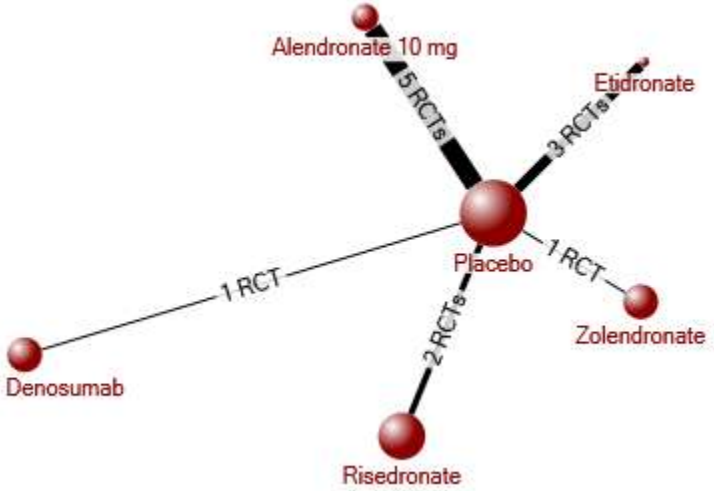
Supplementary Table 2. Risk of Bias Assessment for Studies Included in Network Meta-Analysis

Study	Random Sequence Generation	Allocation concealment	Blinding – objective outcomes	Blinding – subjective outcomes	Incomplete outcome data - efficacy	Incomplete outcome data – safety	Selective reporting	Other bias
AscotEvans 2003	Low	Unclear	Low	Low	Low	Low	Low	Low
Black 1996	Unclear	Unclear	Low	Low	Low	Low	Low	unclear
Black 2007	Low	Unclear	Low	Low	Low	Low	Low	Unclear
Bone 1997	Unclear	Unclear	Low	Unclear	Unclear	Unclear	Low	Low
Chestnut 2004	Unclear	Unclear	Unclear	Unclear	High	High	Low	Low
Cummings 1998	Low	Low	Low	Unclear	Low	Low	Low	High
Cummings 2009	Unclear	Unclear	Unclear	Unclear	Low	Low	Low	Unclear
Fitzpatrick 2011	Unclear	Unclear	Unclear	Unclear	High	High	Low	low
Fogelman 2000	Unclear	Unclear	Unclear	Unclear	High	High	Low	Low
Greenspan 1998	Unclear	Unclear	Unclear	Unclear	High	High	Low	unclear
Greenspan 2002	Low	Unclear	Low	Low	Unclear	Unclear	Low	Low
Grey 2012	Low	Low	Unclear	Unclear	Unclear	Unclear	Unclear	Low
Harris 1999	Low	Low	Low	Low	High	High	Low	low
Herd 1997	Unclear	Unclear	Unclear	Unclear	Low	Low	Low	Low
Hooper 2005	Low	Low	Unclear	Unclear	High	High	Low	Low
Hosking 1998	Unclear	Unclear	Unclear	Unclear	Unclear	Unclear	Unclear	Low
Ishida 2004	Low	Unclear	Unclear	Unclear	Low	Unclear	Low	Low
Li 2005	Unclear	Unclear	Low	Low	Low	Low	Low	Low
Li 2010*	Unclear	Unclear	Unclear	High	Low	Low	Low	Low
Lyritis 1997	Unclear	Unclear	Unclear	High	High	High	Low	Low
McClung 2001	Unclear	Unclear	Low	Low	High	High	Low	Unclear
McClung 2009	Unclear	Unclear	Unclear	Unclear	Low	Unclear	Unclear	Unclear
Menuier 1997	Unclear	Unclear	Low	Low	Low	Low	Low	Low
Miller 2008*	Unclear	Unclear	Low	Low	Low	Low	Low	Low
Montessori 1997	Low	Unclear	Unclear	High	Low	Low	Low	Low
Pols 1999	Unclear	Unclear	Low	Low	Low	Low	Low	Low
Pouilles 1997	Unclear	Unclear	Unclear	Unclear	Low	Low	Low	Low
Qin 2007	Unclear	Unclear	Unclear	Unclear	Low	Low	Low	Low
Reginster 2000	Unclear	Unclear	Low	Unclear	High	High	Low	Low
Storm 1990	Low	Unclear	Unclear	Unclear	High	High	Low	Low
Valimaki 2007	Unclear	Unclear	Low	Low	Unclear	Unclear	Low	Low
Watts 1990	Low	Unclear	Unclear	Unclear	Unclear	Unclear	Unclear	Low
Wimalwansa1998	Low	Unclear	Unclear	Unclear	Unclear	Unclear	Unclear	Low
Yan 2009	Low	Unclear	Low	Low	Low	Low	Low	Low

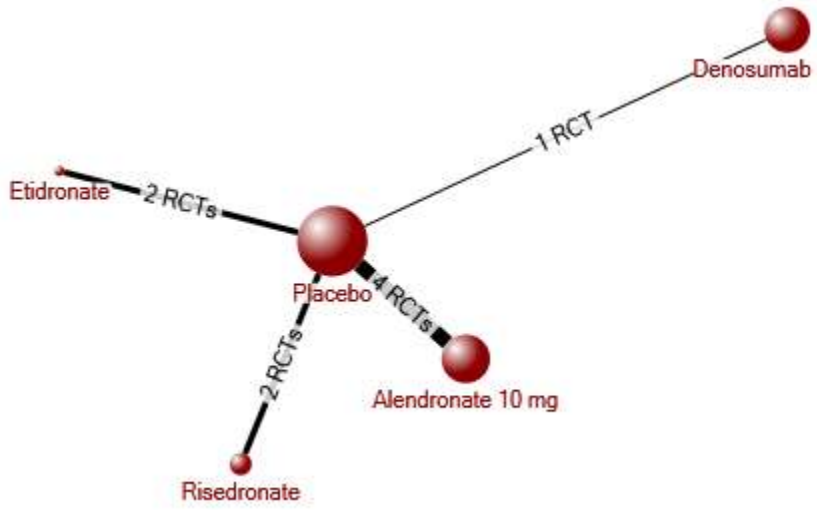
Supplementary Figure 1. Evidence network for non-vertebral fractures.



Supplementary Figure 2. Evidence network for hip fractures.



Supplementary Figure 3. Evidence network for wrist fractures.



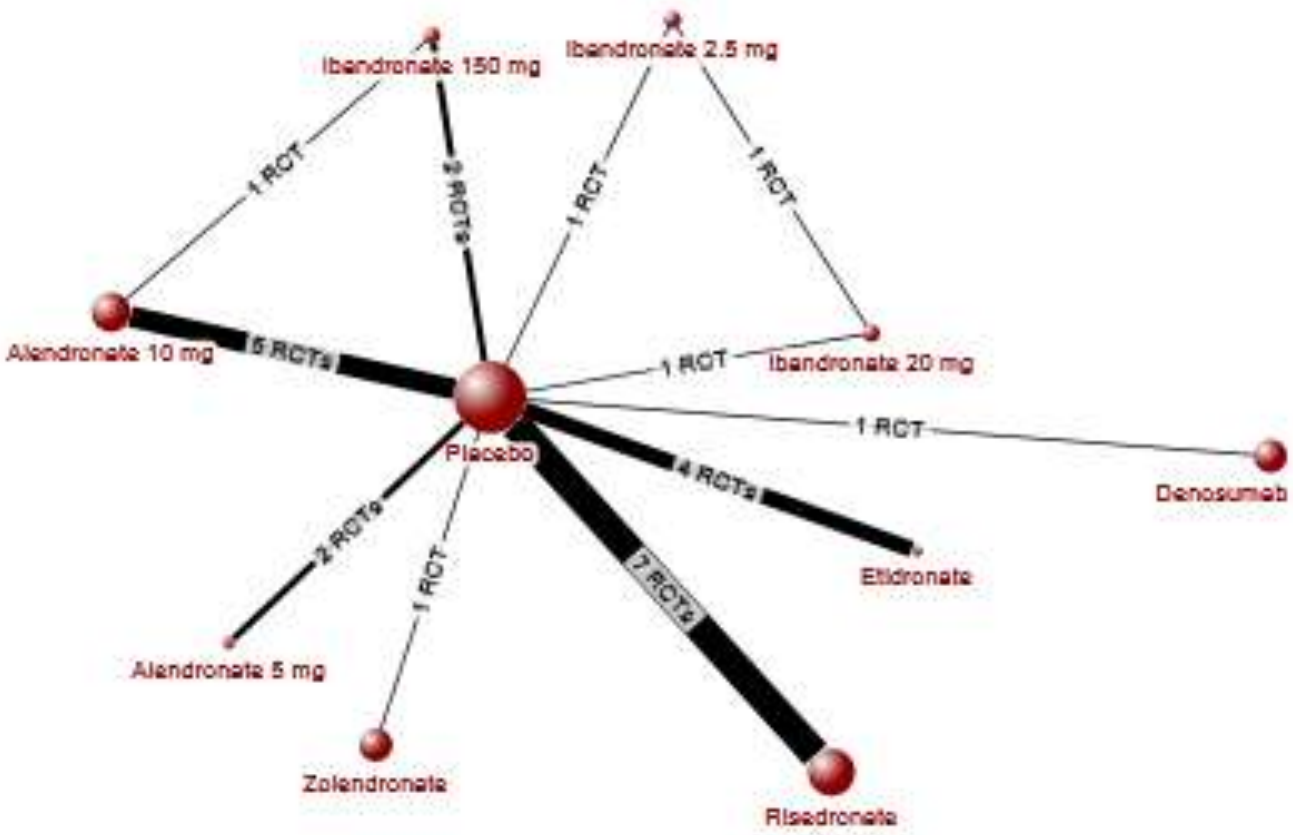
Supplementary Table 3. Network meta-analysis results: Wrist fractures. Odds Ratio (95% Credible Interval)

Treatment	Placebo	ALN10	ETI	RIS	DEN
Placebo	--	--	--	--	--
ALE10	0.74 (0.39,1.12)	--	--	--	--
ETI	4.13 (0.37,19.3)	6.10 (0.48,29.7)	--	--	--
RIS	0.67 (0.24,1.48)	0.98 (0.31,2.53)	0.45 (0.26,2.02)	--	--
DEN	1.25 (0.55,2.47)	1.84 (0.72,4.49)	0.85 (0.054,3.63)	2.32 (0.60,6.26)	--

Note: ALE10, alendronate 10 mg/day po (70 mg/week); ETI, 400 mg/month po;

ZOL, zoledronic acid 5 mg/year iv. DEN, denosumab 60 mg/6months SC.

Supplementary Figure 4. Evidence network for withdrawals due to adverse events



Supplementary Table 4. Number needed to treat for benefit or harm for therapies versus placebo for various outcomes

Treatment	Vertebral Fractures	Non-vertebral Fractures	Hip Fractures	Wrist Fractures	WDAEs
Alendronate 5 mg/day po	58 (B)	296 (H)	NA	NA	78 (H)
Alendronate 10 mg/day po	**28 (B)**	**80 (B)**	**188 (B)**	258 (B)	39 (B)
Etidronate 400 mg/month po	**26 (B)**	55 (B)	165 (H)	23 (H)	89 (H)
Ibandronate 2 mg/every 3 months iv	NA	7 (H)	NA	NA	
Ibandronate 150 mg/month po	38 (B)	1766 (H)	NA	NA	54 (H)
Ibandronate 2.5 mg/day po	**26 (B)**	162 (H)	NA	NA	102 (H)
Ibandronate 20 mg intermittently po	**27 (B)**	128 (H)	NA	NA	219 (H)
Risedronate 150 mg/month	**33 (B)**	**70 (B)**	340 (B)	203 (B)	326 (B)
Zoledronic acid 5 mg/year iv	**18 (B)**	67 (B)	**212 (B)**	NA	33 (H)
Denosumab 60 mg/6 months	**19 (B)**	88 (B)	229 (B)	270 (H)	33 (H)
Abbreviations: B=number needed to benefit; H=number needed to harm; iv=intravenously; po=orally; WDAE=withdrawals due to adverse events					

Note: Control event rates were based on pooled estimates obtained from placebo groups: 8.1% for vertebral fractures, 6.0% for non-vertebral factures, 1.2% for hip fractures, 1.5% for wrist fractures and 8.4% for withdrawals due to adverse events.

Note: Values generated from statistically significant results are bolded. Caution should be exercised when interpreting non-significant values as odds ratios could range from a NNB to a NNH.

Supplementary Table 5A. Random versus fixed effects comparisons for osteoporosis therapies versus placebo: Vertebral fractures.

Treatment	Random OR[a] (95% CrI[b])	Fixed OR (95% CrI)
Alendronate 5 mg/day po	0.77(0.15,2.35)	0.77 (0.15, 2.33)
Alendronate 10 mg/day po	0.53 (0.40,0.70)	0.52 (0.42, 0.65)
Etidronate 400 mg/month po	0.65 (0.13,1.98)	0.50 (0.23, 0.93)
Ibandronate 150 mg/month po	0.49 (0.29,0.78)	0.64 (0.13, 1.92)
Ibandronate 2.5 mg/day po	0.52 (0.32,0.82)	0.49 (0.32, 0.72)
Ibandronate 20 mg intermittently po	0.60 (0.45,0.78)	0.52 (0.34, 0.76)
Risedronate 150 mg/month po	0.49 (0.23,0.91)	0.59 (0.47, 0.74)
Zoledronic acid 5 mg/year iv	0.28 (0.19,0.39)	0.28 (0.22, 0.35)
Denosumab 60 mg/6 months SC	0.31 (0.22,0.44)	0.31 (0.24, 0.39)

[a]OR, Odds Ratio. [b]CrI, Credible Interval. Significant results are in bold font.

Supplementary Table 5B. Random versus fixed effects comparisons for osteoporosis therapies versus placebo: Non-vertebral fractures.

Treatment	Random OR[a] (95% CrI[b])	Fixed OR (95% CrI)
Alendronate 5 mg/day po	1.06 (0.57,1.79)	1.07 (0.61, 1.73)
Alendronate 10 mg/day po	**0.78 (0.60,0.96)**	0.81 (0.70, 0.93)
Etidronate 400 mg/month po	0.68 (0.39,1.11)	0.69 (0.40, 1.09)
Ibandronate 2 mg/every 3 months iv	4.40 (0.013,25.7)	14.2 (0.020, 32.1)
Ibandronate 150 mg/month po	1.01 (0.41,2.09)	1.03 (0.46, 2.03)
Ibandronate 2.5 mg/day po	1.11 (0.71,1.67)	1.09 (0.79, 1.49)
Ibandronate 20 mg intermittently po	1.14 (0.73,1.70)	1.12 (0.81, 1.52)
Risedronate 150 mg/month po	**0.74 (0.57,0.91)**	0.79 (0.69, 0.89)
Zoledronic acid 5 mg/year iv	0.74 (0.53,1.02)	0.73 (0.62, 0.86)
Denosumab 60 mg/6 months SC	0.80 (0.56,1.10)	0.80 (0.67, 0.95)

[a]OR, Odds Ratio. [b]CrI, Credible Interval. Significant results are in bold font.

Supplementary Table 5C. Random versus fixed effects comparisons for osteoporosis therapies versus placebo: Hip fractures.

Treatment	Random OR[a] (95% CrI[b])	Fixed OR (95% CrI)
Alendronate 10 mg/day po	**0.55 (0.27,0.98)**	0.54 (0.27, 0.94)
Etidronate 400 mg/month po	1.52 (0.16,5.95)	1.52 (0.17, 5.98)
Risedronate 150 mg/month po	0.75 (0.51,1.09)	0.74 (0.57, 0.94)
Zoledronic acid 5 mg/year iv	**0.60 (0.36,0.94**)	0.59 (0.41, 0.82)
Denosumab 60 mg/6 months SC	0.62 (0.33,1.08)	0.62 (0.36, 0.97)

[a]OR, Odds Ratio. [b]CrI, Credible Interval. Significant results are in bold font.

Supplementary Table 5D. Random versus fixed effects comparisons for osteoporosis therapies versus placebo: Wrist fractures.

Treatment	Random OR[a] (95% CrI[b])	Fixed OR (95% CrI)
Alendronate 10 mg/day po	0.74 (0.39,1.12)	0.87 (0.66, 1.12)
Etidronate 400 mg/month po	4.13 (0.37,19.3)	4.26 (0.39, 19.5)
Risedronate 150 mg/month po	0.67 (0.24,1.48)	0.65 (0.31, 1.19)
Denosumab 60 mg/6 months SC	1.25 (0.55,2.47)	1.18 (0.88, 1.55)

[a]OR, Odds Ratio. [b]CrI, Credible Interval.

Supplementary Table 5E. Random versus fixed effects comparisons for osteoporosis therapies versus placebo: Withdrawals due to Adverse events

Treatment	Random OR[a] (95% CrI[b])	Fixed OR (95% CrI)
Alendronate 5 mg/day po	1.17 (0.37,2.78)	1.28 (0.81, 1.94)
Alendronate 10 mg/day po	0.67 (0.36,1.14)	0.77 (0.66, 0.88)
Etidronate 400 mg/month po	1.15 (0.40,2.58)	1.13 (0.51, 2.18)
Ibandronate 150 mg/month po	1.24 (0.34,3.40)	0.92 (0.58, 1.38)
Ibandronate 2.5 mg/day po	1.13 (0.26,3.23)	0.93 (0.77, 1.11)
Ibandronate 20 mg intermittently po	1.06 (0.25,3.03)	0.87 (0.72, 1.04)
Risedronate 150 mg/month po	0.96 (0.54,1.61)	0.97 (0.88, 1.07)
Zoledronic acid 5 mg/year iv	1.42 (0.33,4.03)	1.16 (0.94, 1.42)
Denosumab 60 mg/6 months SC	1.41 (0.32,4.06)	1.16 (0.85, 1.56)

[a]OR, Odds Ratio. [b]CrI, Credible Interval.

SUPPLEMENTARY MATERIAL 2. SUPPORTING SUBGROUP ANALYSES DATA NOT SUBMITTED FOR PUBLICATION.

SUBGROUP ANALYSES: TRIAL DURATION

Table 1A. Bisphosphonates versus placebo, ≥ 3 years and < 3 years: Vertebral fractures.

Treatment	≥ 3 years		< 3 years	
	OR (95% CrI)[a]	**Number of studies**	**OR (95% CrI)**	**Number of studies**
Alendronate 5 mg/day po	N/A	N/A	0.62 (0.14, 2.42)	1
Alendronate 10 mg/day po	**0.51 (0.38, 0.69)**	2	0.79 (0.29, 2.11)	2
Etidronate 400 mg/month po	**0.35 (0.12, 0.86)**	3	0.69 (0.23, 2.05)	3
Ibandronate 150 mg/month po	N/A	N/A	0.75 (0.14, 4.06)	1
Ibandronate 2.5 mg/day	**0.49 (0.29, 0.79)**	1	N/A	N/A
Ibandronate 20 mg intermittently po	**0.51 (0.31, 0.84)**	1	N/A	N/A
Risedronate 150 mg/month po	**0.58 (0.42, 0.80)**	2	0.64 (0.32, 1.28)	2
Zoledronic acid 5 mg/year iv	**0.28 (0.19, 0.40)**	1	N/A	N/A

[a]OR (95% CrI), Odds Ratio (95% Credibility Interval). Significant results are in bold font.

Table 1B. Bisphosphonates versus placebo, ≥ 3 years and < 3 years: Non-vertebral fractures.

Treatment	≥ 3 years		< 3 years	
	OR (95% CrI)[a]	**Number of studies**	**OR (95% CrI)**	**Number of studies**
Alendronate 5 mg/day po	N/A	N/A	1.01 (0.55, 1.82)	2
Alendronate 10 mg/day po	0.84 (0.65, 1.07)	2	**0.48 (0.26, 0.86)**	6
Etidronate 400 mg/month po	0.72 (0.27, 1.85)	3	0.64 (0.33, 1.20)	4
Ibandronate 2 mg/every 3 months iv	N/A	N/A	0.49 (0.01, 24.68)	1
Ibandronate 150 mg/ month po	N/A	N/A	0.61 (0.24, 1.60)	2
Ibandronate 2.5 mg/day po	1.11 (0.74, 1.67)	1	N/A	N/A
Ibandronate 20 mg intermittently po	1.08 (0.72, 1.64)	1	N/A	N/A
Risedronate 150 mg/month po	**0.76 (0.50, 0.94)**	3	0.59 (0.28, 1.22)	3
Zoledronic acid 5 mg/year iv	0.73 (0.53, 1.00)	1	1.02 (0.03, 45.36)	1

[a]OR (95% CrI), Odds Ratio (95% Credibility Interval). Significant results are in bold font.

Table 1C. Bisphosphonates versus placebo, ≥ 3 years and < 3 years: Hip fractures.

Treatment	≥ 3 years		< 3 years	
	OR (95% CrI)[a]	**Number of studies**	**OR (95% CrI)**	**Number of studies**
Alendronate 10 mg/day po	0.47 (0.20, 1.07)	1	0.58 (0.18, 1.90)	4
Etidronate 400 mg/month po	0.64 (0.07, 4.60)	2	Data not usable	1
Risedronate 150 mg/month po	0.74 (0.50, 1.09)	2	N/A	N/A
Zoledronic acid 5 mg/year iv	**0.58 (0.35, 0.95)**	1	N/A	N/A

[a]OR (95% CrI), Odds Ratio (95% Credibility Interval). Significant results are in bold font.

Table 1D. Bisphosphonates versus placebo, ≥ 3 years and < 3 years: Wrist fractures.

Treatment	≥ 3 years		< 3 years	
	OR (95% CrI)[a]	**Number of studies**	**OR (95% CrI)**	**Number of studies**
Alendronate 10 mg/day po	0.87 (0.49, 1.41)	2	**0.33 (0.11, 0.86)**	2
Etidronate 400 mg/month po	1.03 (0.10, 10.49)	1	Data not useable	1
Risedronate 150 mg/month po	0.62 (0.24, 1.54)	1	0.52 (0.01, 18.82)	1

[a]OR (95% CrI), Odds Ratio (95% Credibility Interval). Significant results are in bold font.

Table 1E. Bisphosphonates versus placebo, ≥ 3 years and < 3 years: Adverse event withdrawals.

Treatment	≥ 3 years		< 3 years	
	OR (95% CrI)[a]	**Number of studies**	**OR (95% CrI)**	**Number of studies**
Alendronate 5 mg/day po	N/A	N/A	1.01 (0.27, 3.56)	2
Alendronate 10 mg/day po	0.90 (0.69, 1.15)	2	0.52 (0.20, 1.42)	4
Etidronate 400 mg/month po	0.77 (0.29, 2.04)	3	1.67 (0.37, 8.35)	4
Ibandronate 150 mg/month po	N/A	N/A	0.97 (0.24, 4.63)	1
Ibandronate 2.5 mg/day po	0.96 (0.68, 1.38)	1	N/A	N/A
Ibandronate 20 mg intermittently po	0.98 (0.69, 1.40)	1	N/A	N/A
Risedronate 150 mg/month po	0.95 (0.76, 1.14)	3	0.97 (0.35, 2.74)	4
Zoledronic acid 5 mg/year iv	1.16 (0.83, 1.63)	1	N/A	N/A

[a]OR (95% CrI), Odds Ratio (95% Credibility Interval). Significant results are in bold font.

SUBGROUP ANALYSES: SECONDARY AND PRIMARY PREVENTION TRIALS

Table 2A. Bisphosphonates versus placebo, secondary and primary prevention studies: Vertebral fractures.

Treatment	Secondary prevention		Primary prevention	
	OR (95% CrI)[a]	Number of studies	OR (95% CrI)	Number of studies
Alendronate 5 mg/day po	0.64 (0.14, 2.55)	1	N/A	N/A
Alendronate 10 mg/day po	**0.52 (0.36, 0.78)**	3	**0.54 (0.31, 0.95)**	1
Etidronate 400 mg/month po	**0.38 (0.17, 0.79)**	4	N/A[b]	2
Ibandronate 150 mg/month po	0.50 (0.12, 2.09)	1	N/A	N/A
Ibandronate 2.5 mg/ day po	**0.49 (0.29, 0.80)**	1	N/A	N/A
Ibandronate 20 mg intermittently po	**0.51 (0.31, 0.82)**	1	N/A	N/A
Risedronate 150 mg/month po	**0.57 (0.41, 0.76)**	3	0.97 (0.35, 2.75)	1
Zoledronic acid 5 mg/year iv	**0.28 (0.19, 0.40)**	1	N/A	N/A

[a]OR (95% CrI), Odds Ratio (95% Credibility Interval); [b]Zero events resulted in uninterpretable results. Significant results are in bold font.

Table 2B. Bisphosphonates versus placebo, secondary and primary prevention studies: Non-vertebral fractures.

Treatment	Secondary prevention		Primary prevention	
	OR (95% CrI)[a]	Number of studies	OR (95% CrI)	Number of studies
Alendronate 5 mg/day po	1.03 (0.61, 1.75)	2	N/A	N/A
Alendronate 10 mg/day po	**0.71 (0.56, 0.89)**	7	0.56 (0.87, 1.36)	1
Etidronate 400 mg/month po	0.71 (0.40, 1.22)	5	0.51 (0.14, 1.65)	2
Ibandronate 2 mg/every 3 months iv	0.70 (0.017, 28.87)	1	N/A	N/A
Ibandronate 150 mg/month po	0.82 (0.36, 1.88)	1	1.11(0.11, 11.52)	1
Ibandronate 2.5 mg daily po	1.11 (0.81, 1.52)	1	N/A	N/A
Ibandronate 20 mg intermittently po	1.08 (0.79, 1.49)	1	N/A	N/A
Risedronate 150 mg/month po	**0.79 (0.69, 0.89)**	4	0.69 (0.21, 2.15)	2
Zoledronic acid 5 mg/year iv	**0.73 (0.62, 0.85)**	1	0.94 (0.014, 43.74)	1

[a]OR (95% CrI), Odds Ratio (95% Credibility Interval). Significant results are in bold font.

Table 2C. Bisphosphonates versus placebo, secondary and primary prevention studies: Wrist fractures.

Treatment	Secondary prevention		Primary prevention	
	OR (95% CrI)[a]	**Number of studies**	**OR (95% CrI)**	**Number of studies**
Alendronate 10 mg/day po	**0.46 (0.26, 0.76)**	3	1.20 (0.70, 2.05)	1
Etidronate 400 mg/month po	2.20 (0.40, 17.03)	2	N/A	N/A
Risedronate 150 mg/month po	0.62 (0.27, 1.35)	1	0.46 (0.010, 17.88)	1

[a]OR (95% CrI), Odds Ratio (95% Credibility Interval). Significant results are in bold font.

Table 2D. Bisphosphonates versus placebo, secondary and primary prevention studies: Withdrawals Due to Adverse Events (WDAEs)

Treatment	Secondary prevention		Primary prevention	
	OR (95% CrI)[a]	**Number of studies**	**OR (95% CrI)**	**Number of studies**
Alendronate 5 mg/day po	1.25 (0.81, 1.94)	1	N/A	N/A
Alendronate 10 mg/day po	0.90 (0.68, 1.19)	8	0.89 (0.51, 1.28)	2
Etidronate 400 mg/month po	0.67 (0.20, 2.15)	3	3.81 (0.72, 35.53)	3
Ibandronate 150 mg/month po	0.93 (0.56, 1.56)	1	2.93 (0.69, 16.02)	1
Ibandronate 2.5 mg/day po	0.92 (0.77, 1.11)	1	N/A	N/A
Ibandronate 20 mg intermittently po	0.86 (0.72, 1.04)	1	N/A	N/A
Risedronate 150 mg/month po	0.97 (0.88, 1.07)	5	0.66 (0.29, 1.44)	2
Zoledronic acid 5 mg/year iv	1.16 (0.94, 1.42)	1	N/A	N/A

[a]OR (95% CrI), Odds Ratio (95% Credibility Interval)