

*Canadian Agency for
Drugs and Technologies
in Health*

*Agence canadienne
des médicaments et des
technologies de la santé*



HTA

**Guidelines for the Economic
Evaluation of Health
Technologies: Canada**

3RD EDITION, 2006

Supporting Informed Decisions

Until April 2006, the Canadian Agency for Drugs and Technologies in Health (CADTH) was known as the Canadian Coordinating Office for Health Technology Assessment (CCOHTA).

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

This report and the French version entitled *Les lignes directrices de l'évaluation économique des produits pharmaceutiques au Canada* are available on CADTH's web site.

Production of this report is made possible by financial contributions from Health Canada and the governments of Alberta, British Columbia, Manitoba, New Brunswick, Newfoundland and Labrador, Northwest Territories, Nova Scotia, Nunavut, Ontario, Prince Edward Island, Québec, Saskatchewan, and Yukon. The Canadian Agency for Drugs and Technologies in Health takes sole responsibility for the final form and content of this report. The views expressed herein do not necessarily represent the views of Health Canada or any provincial or territorial government.

Reproduction of this document for non-commercial purposes is permitted provided appropriate credit is given to CADTH.

CADTH is funded by Canadian federal, provincial and territorial governments.

Legal Deposit - 2006
National Library of Canada
ISBN: 1-897257-08-2 (print)
ISBN: 1-897257-09-0 (online)

PUBLICATIONS MAIL AGREEMENT NO. 40026386
RETURN UNDELIVERABLE CANADIAN ADDRESSES TO
CANADIAN AGENCY FOR DRUGS AND TECHNOLOGIES IN HEALTH
600-865 CARLING AVENUE
OTTAWA ON K1S 5S8

**In memory of
Bernie O'Brien, BA, MSc, PhD
(1959-2004)**

The Canadian Agency for Drugs and Technologies in Health dedicates this work to the memory of Bernie O'Brien, an outstanding health economist, researcher, and educator, but above all, our highly valued and respected colleague. Internationally renowned as a pioneer in economic evaluation as it relates to health care, Dr. O'Brien was the Director of PATH (Program for Assessment of Technology in Health), Professor in the Department of Clinical Epidemiology and Biostatistics, and an Associate of the Centre for Health Economics and Policy Analysis at McMaster University; Associate Director of the Centre for Evaluation of Medicines at St. Joseph's Healthcare, and Director of the Clinical Effectiveness Research of the Father Sean O'Sullivan Research Centre, Hamilton, Ontario. His wisdom, foresight, and enthusiasm are sorely missed.

This report is a review of existing public literature, studies, materials and other information and documentation (collectively the “source documentation”) which are available to CADTH. The accuracy of the contents of the source documentation on which this report is based is not warranted, assured or represented in any way by CADTH and CADTH does not assume responsibility for the quality, propriety, inaccuracies or reasonableness of any statements, information or conclusions contained in the source documentation.

CADTH takes sole responsibility for the final form and content of this report. The statements and conclusions in this report are those of CADTH and not of its Panel members or reviewers.

CADTH Scientific Advisory Panel Reviewers

André-Pierre Contandriopolous, PhD

Full Professor
Health Administration
Université de Montréal
Montréal QC

Douglas Coyle, MA MSc PhD

Associate Professor, Department of Epidemiology
and Community Health
University of Ottawa
Ottawa ON

David Hailey, MSc PhD GradRIC

Professor, Public Health Sciences
University of Alberta
Edmonton AB

Philip Jacobs, DPhil

Professor
University of Alberta
Edmonton AB

Muhammad Mamdani, MA MPH PharmD

(At the time of document review) Senior Scientist
Institute for Clinical Evaluative Sciences
Toronto ON

CADTH Advisory Committee on Pharmaceuticals Reviewers

The following individuals reviewed early drafts of the third edition.

Roy Dobson, BSc Pharm MBA PhD

Assistant Professor
University of Saskatchewan
Saskatoon SK

RS Nakagawa, BSc FCSHP

(At the time of document review) Director Pharmacy
Fraser Health Authority
Surrey BC

John A. Hoar, BA MPA MDE

Pharmaceutical Economist
Nova Scotia Department of Health
Halifax NS

Angie Wong, BScPhm MSc

Acting Associate Director, Pharmaceutical Services
Drug Programs Branch
Ontario Ministry of Health and Long-Term Care
Toronto ON

CADTH Devices and Systems Advisory Committee

Paul Childs, BA (Hon)
Manager, Business Planning
Business Intelligence Program
Alberta Health and Wellness
Edmonton AB

External Reviewers

Ron Goeree, MA
Assistant Professor
Department CE&B
McMaster University
Hamilton ON

Robert Lee, BSc BScEd MSc
Assistant Professor
University of Calgary
Calgary AB

Braden J. Manns, MD MSc FRCPC
Assistant Professor
University of Calgary
Calgary AB

Orlando Manti, BA (Hon) MA
Senior Economist
Patented Medicine Prices Review Board
Ottawa ON

Christopher McCabe, BA MSc PhD
Senior Lecturer in Health Economics
University of Sheffield
Sheffield UK

Nicole Mittmann, MSc PhD
Scientist
Sunnybrook and Women's College Health
Sciences Centre
Toronto ON

Angela Rocchi, MSc
Principal
Axia Research
Hamilton ON

Alan Shiell, PhD
Professor
University of Calgary
Calgary AB

George W. Torrance, PhD
Professor Emeritus
McMaster University
Hamilton ON

Contributors

Mo Amin, PhD made substantial contributions to draft 2 sections 2, 6, 7, 9, 12, and 14 of the Economic Guidelines in terms of writing; and reviewed drafts of various sections of the document.

Bruce Brady, MA MSc contributed extensively to all parts of the Economic Guidelines; led the project; provided coordination and planning; and contributed to the design, writing and revision of all parts of the document.

Allan Brown, MBA MA wrote drafts of Appendix 2, and helped with organizing the project.

Philip Jacobs, DPhil wrote early drafts of section 10.

Karen Lee, MA reviewed and revised drafts of sections of the Economic Guidelines, and provided comments on drafts of the Economic Guidelines.

Lynda McGahan, MSc drafted information relating to the evaluation of non-drug technologies.

Shaila Mensinkai, MA MLIS conducted literature searches for existing guidelines and for information on the evaluation of non-drug technologies, wrote the literature search part of Appendix 2, reviewed drafts of the document, and verified and formatted bibliographic references.

The authors are grateful to the following individuals for their contributions to the third edition.

- Michael Cheng contributed information on the economic evaluation of non-drug technologies.
- Doug Coyle provided an early draft of sections 3 and 8.
- Jeffrey Hoch provided comments on section 12.
- Debra Marshall contributed information on the economic evaluation of non-drug technologies.
- Nicole Mittmann contributed to early drafts of sections 4 and 6.
- Bernie O'Brien provided detailed comments on, and suggestions for revising, the second edition of the Economic Guidelines.

Other individuals are acknowledged for their contributions.

- Michel Boucher helped with organizing the project; and reviewed and provided comments on the second edition of the Economic Guidelines.
- Eugenia Palylyk-Colwell undertook structural editing of the third draft of the document.
- Vijay Shukla helped with organizing the project; and reviewed and provided comments on various drafts of the Economic Guidelines.

Conflicts of Interest

Nicole Mittmann has been a consultant to several pharmaceutical companies.

Angela Rocchi has had health economic research contracts with multiple Canadian pharmaceutical companies since 1990. From December 2003 to January 2004, she facilitated a meeting and prepared a response to draft 1 of the CCOHTA Guidelines for RX&D.

George Torrance is a developer of the Health Utilities Index (HUI) instrument mentioned in the Economic Guidelines, and a principal of Health Utilities Inc., a company that provides instruments and consulting services to users of the HUI.

FOREWORD TO THIRD EDITION

The objective of the Guidelines for the Economic Evaluation of Health Technologies: Canada is to assist the “doers” of economic evaluations (i.e., analysts) to produce credible and standardized economic information that is relevant and useful to decision makers in Canada’s publicly funded health care system. The guidance provided sets standards for the conduct and reporting of high quality economic evaluations that can be reviewed and compared by decision makers.

The principles in the third edition apply to a variety of health technologies, including those that promote health, prevent and treat conditions, or improve rehabilitation and long-term care. In the past, the Economic Guidelines were primarily directed toward the evaluation of drugs. The audience for the economic evaluations has been mainly publicly funded drug programs and pharmaceutical manufacturers that submit economic information to support the formulary listing of drug products. Increasingly, however, economic evaluations are being used to inform decisions about other health care technologies, such as vaccines, devices, medical and surgical procedures, disease prevention and screening activities, health promotion activities, and health care delivery initiatives such as telemedicine. Such technologies refer not only to individual products but also to strategies for the management or treatment of a condition. The third edition of the Economic Guidelines has been written to address the information needs of this broader audience.

The third edition of the Economic Guidelines follows publications in November 1994 (first edition) and October 1997 (second edition). The third edition reflects the experience gained through using the second edition, and takes into account the methodological developments that have occurred in the economic evaluation of health technologies since 1997. The preparation of the

third edition began with the development of a protocol, which set the following principles:

- provide clear, concise, and practical guidance of a high standard for “doers”
- meet the needs of decision makers for reliable, consistent, and relevant economic information
- identify preferred methods where “best practice” was identified or where there was general agreement among decision makers
- provide succinct information and advice in areas where methodological issues remain unresolved
- allow for flexibility, innovation, and alternative approaches, particularly where methodological issues are unresolved
- assume that the reader is technically literate about the methods of economic evaluation, so that lengthy explanations can be avoided.

Throughout the process, the inherent tensions among these principles required that compromises be made. Practical considerations included the relevance of methods to the needs of decision makers, and the use of more simplified and comprehensible methods where additional complexity was judged to be unnecessary. Notwithstanding such considerations, the inherent time, effort, and cost required to produce economic evaluations consistent with the Economic Guidelines still had to be weighed against the (often greater) cost of wrong funding decisions being made as a result of implementing the findings of a poor quality evaluation.

In preparing the third version of the Economic Guidelines, consideration was given to all the comments received from reviewers. Decisions relating to methodological issues were achieved through consensus.

CADTH takes sole responsibility for the content of the Economic Guidelines.

HIGHLIGHTS OF THIRD EDITION

Format: Each section of the Economic Guidelines addresses a specific topic on the conduct or reporting of economic evaluations. Guideline Statements summarizing the key points of guidance for the analyst to follow are provided at the front of the Economic Guidelines. The strength of the recommendation is implied by the wording. The analyst should follow the recommended guidance when it has been phrased in the “active” voice, whereas more flexibility on the part of the analyst is implied by the use of wording such as “encouraged,” “preferred,” or “consider.”

Reference Case: The Reference Case is the set of preferred methods that an analyst should follow when conducting the base case analysis in an economic evaluation. The purpose of the Reference Case is to aid decision making by enhancing the consistency by which economic evaluations are conducted and reported, thereby improving the comparability among evaluations.

Relevance: Decision makers must have information that is relevant to the circumstances of the decision that they must make. The starting point for meeting a decision maker’s needs is to frame the study question of an economic evaluation in a way that directly addresses the decision, problem, or policy question. Doing so will clarify the scope, design, and reporting of the evaluation. The Economic Guidelines also emphasize the use of “real world” data, and the simulation of “real world” scenarios. When an evaluation is intended to inform a specific decision in more than one jurisdiction or setting, alternative data and assumptions should be included in the analysis (e.g., using sensitivity analyses) to take into consideration meaningful differences between the jurisdictions or settings. In some cases, it may be useful to analyze situations where inappropriate, suboptimal, or unintended use of the technology is anticipated. It is recognized, however, that meeting the “real world” information needs of decision makers is not without challenges.

Flexibility: Although a prime objective of the Economic Guidelines is to encourage the use of consistent approaches for analyzing and reporting evaluations, it is recognized that the Guideline Statements or Reference Case may not apply, or they may be impractical in a particular situation. As a result, the analyst has the flexibility to choose alternative approaches to address the circumstances surrounding the evaluation. Some sections in the Economic Guidelines provide advice for the analyst to consider when no direction on methodological issues has been established. For example, in the Economic Guidelines, a deterministic sensitivity analysis is regarded as a practical and acceptable approach to analyzing uncertainty, even though a probabilistic sensitivity analysis provides a more complete assessment of uncertainty and is more likely to produce an unbiased estimate of costs and effects. A key concern is whether using alternative approaches reduces the quality of the information provided by the evaluation. Analysts should state if the methods used in their evaluation are consistent with the Guideline Statements, and justify any deviations.

Transparency: A key concept in the Economic Guidelines is the need for transparency in the reporting of an evaluation. Analysts should provide complete information on the methods, inputs, and results of an evaluation. Transparency allows users to critically appraise the methodological quality of the evaluation, and to satisfy themselves that potential biases have been appropriately handled. It is also crucial to present information in a way that is useful to the decision maker. All steps in the analysis should be presented in a disaggregated manner before aggregation into cost-effectiveness results. A standard reporting format has been included in Appendix 3 for analysts to use to ensure thorough and consistent reporting.

The third edition of the Guidelines can be downloaded from CADTH’s web site (<http://www.cadth.ca>) or is available in hard copy by contacting CADTH.

TABLE OF CONTENTS

FOREWORD TO THIRD EDITION	iv
ABBREVIATIONS AND CONVENTIONS	vii
1 INTRODUCTION	1
1.1 Economic Evaluations.....	1
2 GUIDELINE STATEMENTS	4
2.1 Study Question	4
2.2 Types of Evaluations.....	4
2.3 Target Population.....	5
2.4 Comparators	5
2.5 Perspective	5
2.6 Effectiveness.....	6
2.7 Time Horizon.....	6
2.8 Modelling.....	7
2.9 Valuing Outcomes.....	7
2.10 Resource Use and Costs	8
2.11 Discounting	9
2.12 Variability and Uncertainty	9
2.13 Equity	10
2.14 Generalizability	10
2.15 Reporting	11
3 GUIDELINES IN DETAIL	12
3.1 Study Question	12
3.2 Types of Evaluations.....	13
3.3 Target Population.....	15
3.4 Comparators	16
3.5 Perspective	18
3.6 Effectiveness.....	19
3.7 Time Horizon.....	23
3.8 Modelling.....	23
3.9 Valuing Outcomes.....	26
3.10 Resource Use and Costs	28
3.11 Discounting	32
3.12 Variability and Uncertainty	32
3.13 Equity	34
3.14 Generalizability	36
3.15 Reporting	38
4 REFERENCES	40
APPENDIX 1: Presenting Results of the Analysis	A-1
APPENDIX 2: Review of Existing Economic Evidence.....	A-4
APPENDIX 3: Standard Reporting Format	A-6
APPENDIX 4: Glossary.....	A-12

ABBREVIATIONS AND CONVENTIONS

CBA	cost-benefit analysis
CCA	cost-consequence analysis
CEA	cost-effectiveness analysis
CEAC	cost-effectiveness acceptability curve
CMA	cost-minimization analysis
CUA	cost-utility analysis
DSA	deterministic sensitivity analysis
EVPI	expected value of perfect information
HRQL	health-related quality of life
HUI	Health Utilities Index
ICER	incremental cost-effectiveness ratio
PICOS	population (or participants), intervention, comparator (or control), outcomes, and study design
PSA	probabilistic sensitivity analysis
QALY	quality-adjusted life-years
RCT	randomized controlled trial
SA	sensitivity analysis
WTP	willingness to pay

The following conventions are used in the Economic Guidelines:

- “analysts” or “doers” are those individuals who conduct economic evaluations
- “comparator” or “alternative” is the technology to which the intervention is compared
- “condition” is a “medical condition” that includes “disease”
- “consequences” of technologies most often refer to “health outcomes” (also referred to as “outcomes,” “effects,” or sometimes “benefits”), although at times they may also refer to other types of consequences such as process factors (e.g., cases found)
- “economic evaluation” is referred to as an “evaluation,” “analysis,” or “study”
- “intervention” is the health technology of interest for assessment
- “users” of economic evaluations most often refer to “decision makers” in Canada’s publicly funded health care system.

1 INTRODUCTION

The primary audience for the Economic Guidelines is composed of economists and health service researchers, in the public and private sectors, who conduct economic evaluations. In turn, the primary audience for the economic evaluations includes Canadian decision and policy makers who are responsible for the funding decisions regarding health technologies. This group includes health policy advisors in the Federal/Provincial/Territorial Ministries of Health, and those working in jurisdictional drug plans, regional health authorities, hospitals, and other health care facilities. In addition, national initiatives such as the Common Drug Review, rely on such information to inform its recommendations. A secondary audience for evaluations includes academics, medical specialist groups, health care providers, patients, patient advocacy groups, manufacturers, media, and the general public.

1.1 Economic Evaluations

The main purpose of an economic evaluation is to “identify, measure, value and compare the costs and consequences of alternatives being considered”¹ to inform “value for money” judgments about an intervention or program.² In this context, “consequences” are most often the health outcomes of the alternatives being compared, although there may be other types of consequences, such as those relating to process (e.g., cases found).

Central to this area of economics are the concepts of “opportunity cost” and “incremental change.” Economics deals with the exchange between people and the trade-offs that they make.³ In publicly funded health care systems, limited resources mean that every available intervention cannot be provided in every situation for all who need or want it. Choices must be made among effective health care interventions, and the decision to fund one means that others cannot be funded. The

opportunity cost of funding the chosen intervention can be seen as the health benefits that could have been derived from funding the next best alternative. Furthermore, the choice of the best course of action depends on weighing only the “incremental changes” in costs and consequences between the alternatives being compared.⁴ Consequently, it is unnecessary to weigh the full range of possible costs and consequences of each alternative.

1.1.1 Use in decision making

A high quality economic evaluation should provide decision makers with information that is useful, relevant, and timely. In addition, evaluations should be based on rigorous analytical methods, be balanced and impartial (credible), and be transparent and accessible to the reader.

There are many situations where economic evaluations can assist decision makers:

- decisions by various levels of government or administrative bodies (e.g., regional health authorities, hospitals, drug plans) to fund a program, service or technology
- pricing decisions by government regulators and technology manufacturers
- clinical practice guidelines⁵
- priorities for research funding by governments and research-based firms
- post-marketing surveillance and updates of economic information based on the use of the technology in the “real world” (which can then be used to inform one of the other types of decisions).

Economic evaluations can provide “value-for-money” information to those making decisions about the allocation of limited health care resources. In particular, economic evaluations can be used to identify interventions that are worth providing and those that are not. Furthermore, evaluations can be used with other approaches to help set priorities, such as program-budgeting marginal-analysis.^{6,7}

There are concerns about the adequacy of economic evaluations for decision-making purposes. Evaluations often lack transparency, which can lead to improper interpretation of the results, and cast doubt on the credibility of the evaluation.⁸ There is also criticism about the dissemination and timeliness of the information, although this is not unique to economic evaluations.^{9,10} Problems of reliability include the inappropriate choice of assumptions and methods in analyses (e.g., data extrapolation techniques), and limitations of the methods (e.g., valuing lost productivity).⁸⁻¹¹ Evaluations have also been criticized for not taking into account the dynamic nature of conditions, outcomes, and costs, and for not taking a comprehensive view of all the factors that can have an impact on the cost-effectiveness of an intervention, such as interactions with existing programs.¹⁰ Problems of relevance include the use of inappropriate comparators, the lack of “real world” data in the analysis, the lack of appropriate subgroup analysis, and poor generalizability of results.⁸⁻¹⁰

Evaluations do not assess all the economic implications of a technology, in particular, the financial consequences of decisions.^{8,12} Budget impact analysis provides complementary information on budgetary expenditure and affordability issues. A comparison of some features of economic evaluation and budget impact analysis is presented in Table 1. Although some data requirements and analytical methods are common to both types of analyses, there are key differences between the two, including the decision maker question that they address. Economic evaluations generally do not distinguish between financial costs and economic (opportunity) costs, which can differ in some situations. Consequently, a reference to “cost

savings” in evaluations generally indicates the value of resources freed up (e.g., release of hospital beds), which may not translate into actual financial savings.

There are systemic barriers to using economic evaluations for decision making, including problems of “silo budgeting,” and a lack of economic expertise by some decision-making bodies, which can lead to the improper interpretation of evaluations.¹²

These factors help explain why economic evaluations have not been used more often for decision making in the health sector. It is difficult to argue, however, that disregarding economic evaluations will lead to better management of limited health care resources. Beyond the usefulness of the actual results of an analysis, economic evaluations synthesize evidence and assumptions in a way that provides users of the information with a structured way of thinking and useful insights about the implications of decisions. This requires that decision makers take a broad view of the impact of a technology, and decisions that are more explicit and transparent. The ultimate test of an evaluation is whether it leads to better decisions in the presence of uncertainty, and results in the more efficient and effective use of resources.

The need for better and more complete economic information by decision makers is reflected in the growing number of guidelines that have been produced worldwide.¹³ By providing standards for the conduct and reporting of economic evaluations, guidelines can address current limitations of evaluations and lead to better studies. Following these guidelines will not eliminate the possibility of bias in evaluations, given the inherent art and judgments that are pervasive in their conduct.

Table 1: Comparison of economic evaluation and budget impact analysis

	Economic Evaluation	Budget Impact Analysis
Question addressed	Is it good value for money?	Is it affordable?
Goal	Efficiency of alternatives	Plan for financial impact
Health outcomes	Included	Excluded
Measure	Added cost per unit of benefit or outcome	Total expenditure (\$)
Time horizon	Usually longer term (may be lifetime)	Usually short (1 to 5 years)

1.1.2 *Timing of evaluations*

Economic evaluations can be undertaken at any point in the life cycle of a technology. The timing of a study ultimately depends on the needs of the decision makers. If an evaluation is conducted late in the life cycle, there is a risk that the findings will not be of use to the decision maker, because the funding decision has been made, or the intervention has diffused into clinical practice, though the findings could inform decisions about changes to reimbursement status or the intended target population. If a technology is evaluated early in its life cycle, before evidence on its effectiveness is clear, there is a risk that the uncertainty about the costs and effects would be larger than if it is evaluated later. Often, the effectiveness of technologies depends on the setting, and sometimes on the operator's experience if there is a learning curve associated with it.

Performing evaluations is an iterative process. Study findings can be updated as more information on the intervention's impacts and "real-world" experience becomes available. A well conducted

evaluation will identify the most important sources of uncertainty, and thereby will direct the gathering of evidence to those areas. This produces more accurate estimates of an intervention in the long term. Bayesian approaches are particularly well suited for this purpose. These approaches can be used to help determine whether to fund a technology or whether additional information should be collected before making such a decision. These approaches can also be useful for re-evaluating technologies that are in use, or where utilization problems have been identified. They also aid in updating decisions about products that have been given probationary funding based on preliminary evidence, with a view to collecting further information on its "real world" use and cost-effectiveness before determination as a full benefit. This is an important aspect of using economic evaluations for decision making.

Suggested readings for those wishing to obtain more information on conducting economic evaluations include Gold *et al.*,⁴ Drummond *et al.*,^{1,14} and Muenning *et al.*¹⁵

2 GUIDELINE STATEMENTS

2.1 Study Question

- 2.1.1 *State the study question to be addressed by the evaluation. The question should be well defined, stated in an answerable form, and relevant to the decision facing the target audience. Relevant and related secondary questions should be included (e.g., the impact of the intervention on subgroups).*
- 2.1.2 *Define the patients or population, intervention, and comparators relevant to the study question. The primary perspective of the study may also be stated in the question.*
- 2.1.3 *Identify the target audience for the study. Secondary audiences may also be listed.*

2.2.4 *Use a cost-effectiveness analysis (CEA) as the Reference Case when a CUA is an inappropriate choice. Use a final outcome (e.g., life-years gained), or if that is impossible, an important patient outcome. Only use a surrogate outcome if it has a well established link (i.e., validated) with one of those outcomes. Consider a CEA as a secondary analysis when the use of one important patient outcome measure [other than a quality-adjusted life-year (QALY) gained] in the denominator of the ICER can be justified, provided that there is a meaningful difference in such an outcome.*

2.2.5 *A cost-minimization analysis (CMA) is appropriate as the Reference Case when the evidence shows that the important patient outcomes of the intervention and comparators are essentially equivalent. Provide justification for conducting a CMA.*

2.2 Types of Evaluations

- 2.2.1 *State and justify the type(s) of economic evaluation chosen. Select the appropriate type of evaluation based on the nature of the research question, the condition of interest, and the availability of data on outcomes.*
- 2.2.2 *In the denominator of the incremental cost-effectiveness ratio (ICER), use a valid outcome measure that is most important to the health of the patient (i.e., important patient outcome).*
- 2.2.3 *Use a cost-utility analysis (CUA) as the Reference Case where meaningful differences in health-related quality of life (HRQL) between the intervention and comparators have been demonstrated.*

2.2.6 *A cost-benefit analysis (CBA) may be useful in some situations, but generally, it should be considered as a secondary analysis. Explain all the steps taken to convert outcomes into monetary values, and analyze key assumptions using a sensitivity analysis.*

2.2.7 *A cost-consequence analysis (CCA) is generally not expected to be used as the Reference Case, unless a CEA or a CUA are inappropriate to use. To enhance reporting transparency, use a CCA as an intermediate step in reporting the other types of economic evaluations.*

2.3 Target Population

- 2.3.1 *Specify the target population(s) for the intervention and its expected use.*
- 2.3.2 *Perform the analysis for the entire target population that is specified in the study question. This may include the population representing the majority or all of its expected use. The efficacy-effectiveness data used in the analysis should be relevant to the target population in the analysis.*
- 2.3.3 *Conduct stratified analysis of smaller, more homogeneous subgroups, where appropriate, if there is variability (heterogeneity) in the target population.*
- 2.3.4 *Analysts are encouraged to analyze situations where it is anticipated that there will be inappropriate, suboptimal, or unintended use of the intervention.*

reflect appropriate (high quality) care. It can be regarded as the first choice in practice or care, as recommended in clinical practice guidelines.

- b) *Where the alternatives are different treatment strategies, distinguish between situations where the intervention is an additional element in the strategy, a different treatment sequence, or a distinct alternative that could replace another element in the treatment strategy. Comparators may be alternative packages of care that consist of many elements. Analyze each strategy separately and explain the alternatives.*
- c) *At times, it may be prudent to analyze the entry of future comparators, including the anticipated entry of lower cost technologies (e.g., generic drugs).*

- d) *For drugs, the alternative agents listed in a formulary may be the most relevant, although those that are not listed should not be excluded. The comparators should include the lowest cost available alternative that is often used for the same indication. Include the cost of the drug and any drug administration costs. Dosing regimens used in the analysis should reflect the dose and duration supporting the effectiveness data for the agent.*

2.4 Comparators

- 2.4.1 *Relate the choice of comparators to the study population, and the local context or practice in which the decision is being made. In principle, consider all technically feasible, acceptable, and relevant alternatives as potential comparators. Then, select the appropriate comparators. Describe and justify the comparators that are chosen for evaluation, and justify those that are not chosen.*
- 2.4.2 *In the Reference Case, use “usual care” (i.e., the most common or frequently used care) which the intervention is intended to replace. In some cases, “usual care” may include more than one relevant, widely used alternative for the same indication.*
- 2.4.3 *Consideration should be given to the following when choosing comparators.*
 - a) *Add “recommended care” as a comparator when usual care does not*

2.5 Perspective

- 2.5.1 *State the perspective(s) of the study in terms of the costs included in the evaluation.*
- 2.5.2 *In the Reference Case, use the perspective of the publicly funded health care system.*
- 2.5.3 *Consider reporting separately the costs associated with adopting a wider perspective, where it is likely that they have a substantial impact on the results of the analysis. Quantify such costs separately, where possible, or at least discuss their likely magnitude and impact on the results of the analysis.*

2.6 Effectiveness

- 2.6.1 *Use a systematic review of the available literature to form the basis for evidence about the efficacy-effectiveness of the intervention. Justify failure to conduct a systematic review. Report the included studies and methods used to conduct the review and analyze or combine data.*
- 2.6.2 *Where feasible and scientifically credible, translate efficacy data into the best quantitative estimate of effectiveness in the Reference Case, using the best available evidence and appropriate modelling techniques. This may involve linking surrogate outcomes to important patient outcomes or extrapolating data beyond the duration of the trial.*
- 2.6.3 *Where feasible in the Reference Case, incorporate “real world” factors that modify the effect of the intervention, where there are established links to important patient outcomes based on the best available evidence. These factors include patients’ adherence to treatment, screening and diagnostic accuracy, and health care providers’ compliance and skill. State the nature of the factor, measures used to quantify the effect, and the methods and assumptions used for modelling.*
- 2.6.4 *The evaluation of medical devices should focus more broadly on the entire episode of care rather than on only the technical performance of the device. The outcomes of medical and surgical procedures, and diagnostic technologies may depend on the operator’s skill and experience. The extensive use of sensitivity analysis may be required to properly evaluate situations where the evidence of efficacy-effectiveness is weak.*
- 2.6.5 *Where feasible, include the impact of adverse events associated with the intervention if they are clinically or economically important, and analyze them appropriately. Depending on the nature, frequency, duration, and severity, adverse events may have an impact on patients’*

adherence, mortality, morbidity, health-related quality of life (HRQL) (utilities), or resource use. Value these in a manner that is consistent with the principles outlined in the Economic Guidelines.

- 2.6.6 *In the Reference Case, extrapolate data based on the best quantitative estimate of the relevant parameters, using the best available evidence and appropriate modelling techniques. Describe the strength of the evidence for extrapolating data and assess uncertainty through a sensitivity analysis. Unless such an analysis is based on high quality evidence, identify it as speculative, and give appropriate caveats in the report.*

2.7 Time Horizon

- 2.7.1 *Base the time horizon on the natural course of the condition and the likely impact that the intervention will have on it. State and justify the time horizon(s) of the evaluation.*
- 2.7.2 *In the Reference Case, ensure that the time horizon is long enough to capture all relevant differences in future costs and outcomes of the alternatives being analyzed. Apply the same time horizon to costs and outcomes. Consider using a lifetime time horizon, and justify where a shorter time horizon is used.*
- 2.7.3 *If the long-term costs and outcomes are modelled, it may be appropriate to present the shorter-term analysis based on primary data, and the longer-term analysis using the extrapolated or modelled data. Multiple time horizons might be appropriate for exploring alternative scenarios in some cases. Explain the causal relationships and techniques that are used to extrapolate or model the data.*

2.8 Modelling

2.8.1 Modelling considerations

- a) *Follow good modelling practices when constructing the model used to conduct the evaluation. Analysts are encouraged to consult good modelling practice guidelines as required.*
- b) *Describe the model, including its scope, structure, and assumptions. Provide justification for assumptions and choices.*
- c) *Use a model structure that is appropriate for addressing the study question. Build the model in such a way to permit updating of results as more data become available.*
- d) *Explain and justify any causal relationships and extrapolation techniques used in the model. Base the extrapolation of data on valid techniques that reflect reasonable scientific evidence, and test through sensitivity analysis.*
- e) *Formally validate the model, and state how this was done.*

2.8.2 Data considerations

- a) *Systematically identify, collect, and assess the data used in the model.*
- b) *Report and identify all data sources. Explain and justify all parameter choices and assumptions.*
- c) *Describe the quality (e.g., strength of evidence) of the data used in the model. Be explicit about data limitations and how they were dealt with. Try to quantify the impact of the limitations on the uncertainty of the evaluation results.*
- d) *Gather the best available evidence on key model parameters for which the model results are most sensitive. Justify any failure to gather the best available evidence of such parameters.*

- e) *Use caution when expert opinion is used to establish parameter values. Justify its use; and describe the source of the opinion, the method of elicitation, and the results of the exercise. Assess such estimates through a sensitivity analysis.*
- f) *Use appropriate methods to analyze or combine data from different sources. Explain and justify the methods used, and report the results of the analysis. Report limitations in the methods or data used, and where feasible, test through a sensitivity analysis.*
- g) *Incorporate data into the model using appropriate techniques, and explain the methods used. If data are incorporated as point estimates, use mean estimates of parameters in the base case. If estimates are incorporated as probability distributions, state and justify the form of the distributions.*

2.9 Valuing Outcomes

- 2.9.1 *Use appropriate preference-based measures to value meaningful differences between the intervention and alternatives in terms of HRQL.*
- 2.9.2 *Measure the outcome for a CUA in terms of the QALYs gained. Report changes in the length of life and quality-weight separately, and report the procedure for combining them. State the assumptions and methods used to estimate QALYs. Justify using alternative outcome measures in a CUA.*
- 2.9.3 *Preferences (utilities) can be measured directly or indirectly. Study the alternative methods a priori and select in advance the one that is most appropriate for the condition and study question. Justify the selection and method, report on the validity and reliability of the method selected, and explain the steps undertaken to measure preferences.*

2.9.4 *Where preferences are measured directly, use the standard gamble or time trade-off approaches. To avoid double-counting, subjects in exercises measuring preferences should be asked to value lost leisure time in terms of changes in preferences, and to assume that health care costs and income losses are fully reimbursed.*

2.9.5 *A representative sample of the general public, suitably informed, is the preferred source for preferences. Patients who have direct experience of the relevant health states may be an acceptable source. Describe the population from which the preferences were derived, and their relevance to the Canadian population.*

2.9.6 *Willingness-to-pay methods for valuing outcomes in a CBA are regarded as a secondary type of analysis. Explain the steps to convert outcomes into monetary terms. Validate key assumptions, and test through a sensitivity analysis.*

2.10 Resource Use and Costs

2.10.1 General

a) *Systematically identify, measure, and value resources that are relevant to the study perspective(s). Classify resources in categories that are appropriate to the relevant decision maker (e.g., primary care, drug plan, hospitals).*

2.10.2 Resource identification

a) *Exclude protocol-driven costs taken from clinical trials. Transfer payments should be excluded from the public payer and societal perspectives.*

b) *Unrelated costs that are incurred during normal life-years should be excluded from the evaluation. Unrelated costs that are incurred during life-years gained from the intervention may be included at the analyst's discretion in a sensitivity analysis.*

2.10.3 Resource measurement

a) *Report quantities of resources in physical units.*

b) *Report the costing method used and justify the approach taken. Measure and value with greater precision those resources that contribute most to total and incremental costs. Where lower quality cost estimates are used, use a sensitivity analysis to determine the impact of cost assumptions.*

c) *Where feasible, base resource use estimates on data for Canadian routine practice. Where resource use data are from international sources, clinical trials, or non-observational sources (clinical practice guidelines), validate or adjust them for Canadian routine practice, using appropriate methods.*

2.10.4 Resource valuation

a) *Conceptually, use economic (opportunity) costs as the basis for valuing resources. In principle, use total average cost (including capital and allocated overhead costs) as the unit cost measure.*

b) *Report the valuation methods used, and justify the approach where appropriate. Use market prices where available. Standard costs can be used where available and appropriate. Where costs are directly calculated or imputed, they should reflect the full economic cost of all relevant resources at normal operating levels.*

c) *When evaluating the public payer perspective, use the full cost (i.e., contributions paid by the public payer, private insurers, and patients) of the intervention and comparators in the Reference Case. For interventions involving cost-sharing arrangements with patients that are likely to have a noticeable impact on the results, use a sensitivity analysis to assess the implications of variations in the proportion of the cost of the intervention and comparator paid by*

- the public payer. Use the same proportions for the intervention and comparators, unless there is a reason to do otherwise.*
- d) *Adjust any cost obtained from earlier times to the current period. Use appropriate methods, and provide justification when converting costs (i.e., resource quantities and unit costs) from another country to Canadian currency.*
- e) *Consider a separate analysis of the impact of the intervention on lost time by patients and informal caregivers, where it is likely to have a substantial impact on the results.*
- f) *Use the friction cost approach to value lost time from paid work. Report the friction period and unit cost used to value lost productivity. Gross wage rates plus the costs associated with recruiting and training replacement workers can be used to value long-term absences from work. Exclude the lost time from paid work due to premature death that occurs beyond the friction period.*
- g) *There are several acceptable methods for valuing lost time by patients and informal care-givers, but there is no preferred alternative.*
- h) *Describe the methods, data, and assumptions used to measure and value lost time by patients and informal caregivers. Present quantities and unit costs of lost time separately before combining them. Conduct a sensitivity analysis using alternative methods and assumptions.*

2.11 Discounting

- 2.11.1 *In the Reference Case, discount the costs and health outcomes that occur beyond one year to present values at the (real) rate of 5% per year.*
- 2.11.2 *Conduct sensitivity analyses using (real) discount rates of 0% and 3%.*

- 2.11.3 *When different discount rates are used from those recommended, present results in a sensitivity analysis, and justify the relevance.*

2.12 Variability and Uncertainty

2.12.1 Handling variability

- a) *Variability can be attributed to diverse clinical practice patterns in different geographical areas or settings, or to inherent variability in the patient population (i.e., patient heterogeneity). Handle variability in practice patterns through further analysis.*
- b) *Deal with variability in the population by stratifying the target population into smaller, more homogeneous groups. Identify the basis for the stratification. Define subgroups preferably at the planning stage, because post-hoc analysis may be unacceptable, unless a strong justification is given.*

2.12.2 Handling uncertainty

- a) *Uncertainty can be attributed to two types of model inputs: parameter and model (structure, methods, and assumptions). Deal with both types of uncertainty systematically and thoroughly, and fully assess the impact on the results and conclusions.*
- b) *In the Reference Case, at a minimum, conduct a deterministic sensitivity analysis (DSA).*
- *Perform the analysis for all model inputs to determine the impact on the results. Justify the omission of any model input from the sensitivity analysis.*
 - *Identify and fully assess the key model inputs contributing most to uncertainty. The choice of analysis should involve more than a one-way sensitivity analysis. Perform multi-way sensitivity analysis, threshold analysis, and analysis of extremes (e.g., best and worst case scenarios) for key model inputs.*

- *Assess the full range of plausible values for each parameter, and plausible alternatives for each assumption. State and justify the ranges of values selected, and the alternative assumptions used. Alternative assumptions should take into account the variability between the jurisdictions or settings of the target audience.*
- c) *A probabilistic sensitivity analysis (PSA) of parameter values that can be defined probabilistically is encouraged to more appropriately assess parameter uncertainty.*
- *The analysis should take the form of a Monte Carlo simulation. State and justify any assumptions regarding the range of values for key parameters, the form of probability distributions, and the number of Monte Carlo iterations.*
 - *Model uncertainty should be assessed through a DSA and model validation methods, with separate (probabilistic) results shown for each alternative analysis.*
 - *Parameter uncertainty can be assessed using a DSA and a PSA.*
- d) *Where a PSA has been used, quantify the contribution of each parameter to decision uncertainty. Value-of-information methods can be used to indicate where the collection of additional information may be helpful for making decisions.*

2.13 Equity

- 2.13.1 *State the implicit and explicit equity assumptions made in the evaluation. If possible, state the implications of the assumptions on the results of the analysis.*
- 2.13.2 *Identify the equity-relevant characteristics of the subgroups that may benefit from, or be adversely affected by, the intervention. Population characteristics such as age, sex, ethnicity, geographical area, socioeconomic group, or health status, may be relevant for equity purposes.*

2.13.3 *Analysts are encouraged to provide information on the distributional impact (e.g., benefits, harms, and costs) and cost-effectiveness of the intervention for those subgroups predetermined to be relevant for equity purposes.*

2.13.4 *Use equal equity weights for all outcomes in the Reference Case. Present the analysis in a disaggregated and transparent manner to allow decision makers to assess the distributional impacts and the trade-off between equity and the efficient allocation of resources.*

2.14 Generalizability

2.14.1 *Address generalizability in the design of the evaluation and in the interpretation of its findings. There are three aspects of generalizability to be addressed:*

- *distinction between efficacy and effectiveness of the intervention*
- *handling of data on costs and preferences (utilities) that are derived from another setting*
- *handling of data from trials involving several countries, including that of the decision maker.*

2.14.2 *Justify any data derived from outside Canada and verify for the Canadian setting. If data are adjusted for the Canadian setting, describe and justify the methods used. Report, analyze, and justify the use of cost data from multinational trials.*

2.14.3 *Where there is local variation in clinical practice or other model parameters, the Reference Case can be performed at a national (or aggregate) level using the most widespread or best available practice or data. A sensitivity analysis can be performed using regional or local practice and data. If a DSA is used, test the key model parameters throughout the range of values that apply in the jurisdictions representing the target audience.*

2.14.4 *Present the results in a disaggregated manner to facilitate the interpretation of results for different settings. Report the quantities of resources consumed and unit costs separately.*

2.14.5 *State the extent to which the findings of the evaluation can be generalized to the jurisdiction(s) or setting(s) of the target audience, including any study limitations that affect the generalizability of the evaluation findings.*

2.15 Reporting

2.15.1 *Report the evaluation in a transparent and detailed manner. Provide enough information to enable the audience to critically evaluate the validity of the analysis. Use a well structured report format (Appendix 3).*

2.15.2 *Include a summary and a conclusion of the evaluation that are written in non-technical language and that are accessible to the target audience.*

2.15.3 *Present the analysis in disaggregated detail first, showing total, undiscounted costs and outcomes separately for the intervention and each comparator. Introduce aggregations, incremental results, and value judgments as late as possible.*

2.15.4 *Report final results as incremental cost-effectiveness ratios (ICERs), based on incremental differences of expected costs and expected outcomes of the alternatives. Follow standard decision rules for estimating ICERs, including the exclusion of dominated alternatives. To aid understanding, analysts are encouraged to present the results of the analysis in graphical or visual form, in addition to tabular presentation.*

2.15.5 *Describe funding and reporting relationships of the evaluation, and disclose any conflicts of interest.*

2.15.6 *Make documents demonstrating quality assurance in the conduct of the evaluation available to decision makers. If requested, make a copy of the model available to decision makers for review.*

3 GUIDELINES IN DETAIL

3.1 Study Question

3.1.1 *Defined, decision-relevant question*

The first step in undertaking an evaluation is to develop the study question. This will help determine the scope, design, and reporting of the evaluation best suited for informing the decision, and prevent wasted effort in conducting the evaluation. The question should not be framed in terms of a broad issue.

The study question will be related to the decision problem prompting the analysis. In framing the study question, the analyst should define the patient population, intervention, and the appropriate comparators. The perspective of the study (e.g., public payer) may also be stated in the question. When the evaluation includes more than one perspective, the primary perspective should be used. A balance should be struck between defining these factors precisely enough to be relevant to the target audience (and thereby avoid wasted time and effort conducting unfocused research) while not being overly narrow, resulting in the lack of relevance and generalizability of the findings.

Patients or populations can be defined in terms of the condition (e.g., severity, stage, or risk level), demographic characteristics (e.g., age, sex, race, or health status), or setting (e.g., community, outpatient, or inpatient). The intervention and comparator(s) can be defined in terms of dose or treatment intensity, setting (e.g., primary care, health centre, or home), co-interventions, and method of delivery (e.g., intravenous or oral administration of drugs). The intervention and comparators may be treatment strategies rather than products.

For example, the question may be: “From the perspective of the public payer in Canada, what is the cost-effectiveness of the intervention for a particular population in a certain setting, compared to ‘usual care’?”

The evaluation can include secondary questions that are related to the main study question. This may include the impact of the intervention on subgroups (e.g., high risk versus average risk patients) or the impact of variations in treatment (e.g., monotherapy versus combination therapy).

3.1.2 *Target audience*

The study question must be relevant to the needs of the target decision makers. What is relevant will be determined by the question that the decision maker needs to answer, so that he or she can make a decision about the intervention. Although the target audience for an evaluation may be more than one decision maker, the evaluation should fit the purpose of informing a specific decision. Furthermore, the study question may lead to findings that are generalizable beyond the context of the target audience or jurisdiction. Where appropriate, the analyst should consult those with a good understanding of the problem requiring resolution (e.g., clinical experts or health service managers), to help frame a clear and relevant study question, and to better understand the broader context of the decision to be made.

The primary audience will have implications for the design of the evaluation and choice of data for the analysis. As a result, the primary target audience for the study should be identified. In Canada, the primary audience for an evaluation will often be more than one decision maker or jurisdiction. As a result, the evaluation should account for meaningful variation between these settings or jurisdictions. For example, the target audience could be a single entity, such as the Common Drug Review;¹⁶ in turn, this process influences the drug funding decisions of the federal, provincial, and territorial jurisdictions in Canada. Secondary audiences may include stakeholders who may use the information in an evaluation (e.g., academics, medical specialty groups, health care providers, patients, patient advocacy groups, manufacturers, media, and the general public).

3.2 Types of Evaluations

There are five types of economic evaluations (i.e., CUA, CEA, CMA, CBA, and CCA).¹⁷ The selection of the appropriate type of evaluation depends on the research question, the condition of interest, and the availability of data on outcomes. Analysts should justify the choice of outcome and type of evaluation chosen.

3.2.1 Types of outcome

The outcomes that are used to measure the health effects of interventions can be classified into three types. The outcomes are ranked in order of importance and relevance for the health of patients.

Final outcomes are related directly to the length and quality of life. Examples include deaths prevented, life-years gained, and QALYs gained.

Important clinical outcomes are valid outcomes of importance to the health of the patient. They include disease-specific events such as stroke and myocardial infarction. Final outcomes and condition-specific or generic measures of quality of life are excluded from this outcome.

A surrogate outcome is “a laboratory measurement or a physical sign used as a substitute for a clinically meaningful endpoint that measures directly how a patient feels, functions, or survives.”¹⁸ Examples for cardiovascular disease include blood pressure or cholesterol level.

- Validated surrogate outcomes are proven to be predictive of an important patient outcome. A surrogate outcome is valid only if there is a “strong, independent, consistent association” with an important patient outcome, and there is “evidence from randomized trials that... improvement in the surrogate end point has consistently lead to improvement in the target outcome.”¹⁸ (see a paper by Prentice).¹⁹
- Unvalidated (unproven) surrogate outcomes have not been proven to be predictive of an important patient outcome.

“Final outcomes” and “important clinical outcomes” are collectively referred to as “important patient outcomes” in the Economic Guidelines. These categories are not mutually exclusive, and these outcomes may reflect benefit

or harm, depending on whether the incidence is reduced or increased by the intervention.

Analysts are encouraged to select an outcome indicator that is most appropriate for the relevant condition, and most feasible, given the available data on outcomes for each alternative. The outcome selected should be accurately measured, and common to the alternatives being compared. Emphasis should be placed on using the relevant and valid outcomes of the highest importance for the health of patients. For cardiovascular disease, this could include all-cause mortality, and cardiovascular-related mortality such as fatal myocardial infarction or stroke; all-cause serious morbidity, and cardiovascular-related morbidity such as non-fatal myocardial infarction or stroke. In determining effectiveness, the evidence on final outcomes (e.g., life-years gained) is preferred to that of validated surrogate outcomes. Outcomes with less clear validation or relevance for patients should also be described, if relevant. If the protocol-defined primary efficacy outcomes from a clinical trial are not used, justification for the new choice of outcomes should be explained.

3.2.2 Types of economic evaluation

In a cost-utility analysis (CUA), outcomes are measured as health-related preferences, which are most often expressed as QALYs gained (i.e. a final outcome). This type of evaluation is useful when interventions have an impact on the HRQL, and on the length of life.

A CUA uses a generic outcome measure that permits decision makers to make broad comparisons across different conditions and interventions. This feature facilitates the allocation of resources based on maximizing health gains. Using a CUA is not without problems. For instance, the methods and instruments for measuring preferences often produce different scores for the same health state. In some cases, the public or decision makers may not consider QALYs to be wholly comparable across all conditions (e.g., chronic versus acute conditions, mild versus severe conditions).²⁰

A CUA should be used in the Reference Case where meaningful HRQL differences between the intervention and alternatives have been

demonstrated, and where appropriate preference (utility) data are available. Preferences should be derived using valid approaches.

In the literature, the term “cost-effectiveness analysis” is often used to refer to economic evaluations in general. In the Economic Guidelines, a CEA refers to a specific type of economic evaluation in which the outcomes are measured in natural (health) units, such as life-years gained, lives saved, or clinical event avoided or achieved.

A CEA should be used as the Reference Case when a CUA is an inappropriate choice. A CEA can be used as a secondary analysis when the use of an important patient outcome measure (other than a QALY gained) can be justified, provided that there is evidence of a meaningful difference in such an outcome compared with alternatives.

It is preferred that the outcome measure be a final outcome (e.g., life-years), or if that is impossible, an important clinical outcome. In general, a surrogate outcome should only be used as an outcome measure if it has a validated, well established link with an important patient outcome. Such an analysis should be appropriately tested through a sensitivity analysis. The linkage between different outcomes can be modelled.

A CEA is more straightforward to conduct than a CUA or CBA. A disadvantage of CEA is that the results can only be compared with the results of other technologies that are expressed using the same (or very similar) outcome measure. It does not facilitate the comparison of technologies and the allocation of resources across different conditions because of its reliance on one natural measure of health. Furthermore, a CEA may be inappropriate, when using one measure of outcome does not account for the full range of important patient outcomes due to an intervention.⁴

The results of a CEA or CUA should be expressed as an incremental cost-effectiveness ratio (ICER). The net benefit measure may be used as an additional (but not alternative) measure to the ICER, where a specific willingness-to-pay threshold has been assumed. The willingness-to-pay threshold and the associated ICER should be stated for each net benefit estimate.

In a cost-minimization analysis (CMA), alternatives are considered to be equivalent in terms of factors that are relevant to the decision (other than cost), and so, the lowest cost alternative is selected. A decision to conduct a CMA should not be taken at the inception of the evaluation but only after an assessment of the clinical evidence concerning the intervention and the appropriate alternatives. A CMA can be regarded as an extension of a CEA or a CUA where the outcomes are demonstrated to be equivalent, and so only the costs of the alternatives are compared.

The critical issue with the appropriate use of a CMA is whether there are meaningful differences between the intervention and alternatives in terms of important patient outcomes (including important adverse events). This decision should be justified, based on a high quality assessment of the intervention. If there is evidence of meaningful differences in any important patient outcomes, a CMA would be inappropriate.

Where the evidence demonstrates that the important patient outcomes of the intervention and alternatives are essentially equivalent, then a CMA is appropriate to use as the Reference Case. A clear justification for conducting a CMA should be provided. This may arise in two situations.

- The totality of the evidence shows that there are problems demonstrating the superiority of an intervention in terms of important patient outcomes, and it is appropriate to conclude that no meaningful difference with comparators exist. For instance, there may be a lack of good quality trials to conduct a meta-analysis, or a large number of participants may have dropped out of a key trial.
- The evidence of equivalence is demonstrated through a well designed and adequately powered trial (or meta-analysis of trials). A CMA of drugs should use the dosage of each comparator required to achieve the same therapeutic effect, and the dose equivalencies should be justified. Briggs *et al.*²¹ argue that a CMA has often been used inappropriately, because the evidence that there is no difference in the efficacy-effectiveness of treatments may have been based on clinical trials with inadequate statistical power. This has often been misinterpreted as evidence of equivalence (i.e., committing a type 2 error).

Beyond treatment effects, alternatives may differ in terms of other factors such as adherence or convenience of use (e.g., due to less frequent drug administration). These differences should only be considered as relevant for excluding a CMA where they have an established link to changes in important patient outcomes. Evidence of only a possible impact on such outcomes may be explored through a sensitivity analysis (e.g., threshold analysis) in a CEA, CUA, or CBA.

A cost-benefit analysis (CBA) values costs and outcomes in monetary terms. Values are usually obtained through using a willingness-to-pay approach, such as contingent valuation or conjoint analysis. The use of a CBA in health care decision making has been limited, despite a CBA being the only type of economic evaluation that directly addresses allocative efficiency (i.e., allocating resources between sectors). The difficulties with using a CBA in health technology assessment relate to methodological difficulties with measuring health outcomes in monetary terms, and ethical issues arising from assigning monetary values to health outcomes.²² In particular, willingness-to-pay approaches often depend on an individual's ability to earn income.

It may be appropriate to use a CBA in certain situations, such as when:

- a consequence of an intervention is difficult to value using QALYs (e.g., short-term symptom relief, patient reassurance or anxiety from screening)
- an attribute of an intervention is difficult to value using any health outcome (e.g., shorter or less frequent treatment, a more convenient dose form)
- a process outcome are major factors in analyzing an intervention (e.g., access to or satisfaction with care).

See a paper by Ryan *et al.*²³ for an example of an application of a CBA in the latter situation. A paper by O'Brien and Viramontes²⁴ provides an example of using the willingness-to-pay approach to measure health state preferences. When using a CBA, the evaluation should explain the steps taken to convert the outcomes into monetary values. Key assumptions should be thoroughly tested through a sensitivity analysis.

In a cost-consequence analysis (CCA), the costs and outcomes of the alternatives are listed separately in a disaggregated format (e.g., intervention costs, hospital costs, clinical benefits, and adverse events). This type of evaluation can be useful for obtaining a picture of the impact of the intervention. It does, however, place the burden of aggregating, weighing, and valuing the components on the user of the study.

Generally, a CCA is not preferred for the Reference Case, although it can be useful in some cases. In particular, decision makers may value information presented in disaggregated form when using one measure of benefit (i.e., a CEA) that does not take account the full range of health effects of an intervention, or when it is difficult or misleading to combine multiple outcomes from an intervention in a QALY for a CUA. For example, focusing exclusively on a behaviour change in a population resulting from a health promotion activity may under-value other benefits from the activity, such as raising awareness. A CCA would also be acceptable to use when there is no unambiguous evidence to conclude that there is a "meaningful difference" in important patient outcomes. Such a situation may arise when a surrogate outcome is not validated, and so it cannot be effectively extrapolated to an important patient outcome. Furthermore, the transparency of other types of economic evaluations is improved when a CCA is used as an intermediate step in reporting the analysis, with the outcomes and costs presented in a disaggregate form before combining them in another type of evaluation.

3.3 Target Population

3.3.1 Commentary

The cost-effectiveness of a new intervention depends on the population being evaluated. The study question should specify the target population(s) for the intervention. This could include a description of the patient population for the indication approved by Health Canada. In cases where the evaluation is to be used for reimbursement purposes, reference to the reimbursement status and restricted use criteria of possible alternatives can guide the potential reimbursement status of the intervention.

The evaluation should analyze the entire target population as defined in the study question. Target populations may be defined using baseline demographic features that describe the type of patient (e.g., age, sex, socioeconomic status) with a specific condition, of a certain severity or stage, with or without co-morbidities or risk factors. In addition, populations can be defined by setting (e.g., community or hospital), geographic location, usual adherence rates, or typical patterns of treatment. The analyst should also describe the expected use or place in therapy of the intervention (e.g., replacement for current therapy, use with current therapy, use for non-responders only, use only for those with contraindications or with intolerance to current therapy). The efficacy-effectiveness data used in the analysis should be relevant for the target population.

It may be appropriate to conduct a stratified analysis of smaller, more homogeneous subgroups where there is variability (heterogeneity) in the target population. Variability may relate to differences in health outcomes, patients' preferences (utilities), and costs of the intervention among subgroups of patients.

Stratified analysis or sensitivity analysis can be used to evaluate the cost-effectiveness of situations where there is a potential for inappropriate, suboptimal, or unintended use of the intervention. These situations may occur for groups in or outside the target population.

- Health care payers often limit the reimbursement of interventions to more restricted subgroups of patients than those approved by Health Canada. Experience shows, however, that clinicians may not adhere to reimbursement criteria or conditions of use, and extend the intervention to patients who do not meet the clinical or demographic criteria or conditions (although patients may still fall within the approved indication). This can lead to over-prescribing and uncontrolled growth of the intervention. The evaluation can include an analysis of the anticipated use of the intervention in this larger population than intended by the reimbursement authority. Such situations may have implications for the selection of comparators, if for instance, those patients in the expanded group would not have otherwise been treated. Coyle *et al.*²⁵ provide

an example on how to undertake a stratified analysis for establishing efficient limited use criteria.

- Drugs that are listed as second-line therapy in a formulary may be used as first-line therapy.
- It may be difficult to precisely determine the patients who are appropriate candidates for the intervention. These difficulties may not be apparent in the clinical trials, but they do occur in clinical practice. If appropriate, no effect may be assumed for the unintended use of the intervention in misdiagnosed patients. For example, a study by Husereau *et al.*²⁶ uses a sensitivity analysis to show the impact of assumptions about misdiagnosed and (separately) late-presenting patients who are prescribed an antiviral treatment for influenza. As shown, a weighted-average ICER can be calculated using varying ratios of appropriate to inappropriate use.
- There is a potential for the non-approved off-label use of the intervention because of available clinical trial data for indications not (yet) approved by Health Canada or information from jurisdictions where the product has been marketed. If possible, the analyst should indicate how established such off-label uses are.
- With changes in reimbursement status or use, technologies that are funded are evaluated for potential changes (e.g., restrictions or delisting).

Conducting these types of analyses may require clinical data on the specific population or group of interest (e.g., average risk versus high risk patients). In some situations, it may be useful to obtain utilization data from jurisdictions where the intervention has been made available or to examine utilization data on similar interventions that have entered the market. Indirect comparisons may be necessary in some situations.

3.4 Comparators

3.4.1 General considerations

It is crucial to select the appropriate comparators for the analysis, as the choice will be important in determining the cost-effectiveness of the intervention and the relevance of the study to

decision makers. In principle, the comparator is the alternative that is most likely to be replaced in clinical practice should the intervention be adopted.

- Consider the study question, the indication or purpose of the intervention, and the target audience for the study. Also consider all approved, accepted, and technically feasible alternatives that are indicated for the condition as comparators. This does not necessarily mean that all such alternatives should be used as comparators in the analysis. Selection may be done through a process of elimination. It may be helpful to seek input from clinical or content experts during the process to identify all reasonable alternatives.
- Selecting comparators may be complicated when there is a range of approved alternatives for the same indication, or if there is variation in clinical practice across jurisdictions or patient subgroups (e.g., patients in nursing homes versus the general population). In practice, analysts may have to identify a small number of primary relevant comparators for analysis. In doing so, scientific rigor should be balanced against data availability, time constraints, and the feasibility of analyzing a large number of comparators.
- If there are no head-to-head clinical trials comparing the intervention and relevant comparators, the analyst may use indirect comparisons based on appropriate techniques. Methods used to synthesize indirect comparisons should be explained and justified. Any limitations of the methods, potential biases in the parameter estimates, and caveats about the interpretation of results, should be reported. A sensitivity analysis may also be used to assess the impact of assumptions about comparators.

3.4.2 “Usual care”

In the Reference Case, the comparator should be “usual care,” which is the most common or most widely used treatment in clinical practice for the condition. This is also referred to as “existing practice,” “current practice,” “typical care,” or “status quo.” The most commonly used treatment that the intervention is intended to replace can be the one used for the largest number of patients, based perhaps on utilization data and clinical

expert opinion. In addition, there may be the most prevalent type of care that dominates clinical practice or there may be two or three prevalent alternatives, in which case, all should be individually compared to the intervention being studied.

3.4.3 Additional considerations

Usual care may not always reflect the level of care that is recommended or that is clinically the most effective. In such situations, “recommended” or “appropriate” care should be included (in addition to usual care) when it is considered to be a feasible and relevant option. High quality, clinically appropriate care can be determined by referring to recommendations in evidence-based clinical practice guidelines or by clinical experts. In such cases, the strength of evidence for the alternatives should be provided. Several alternatives may be relevant where the preferred treatment is ambiguous.

For some interventions, comparisons may be made between treatment strategies rather than individual products. In such cases, distinguish between situations where the technology is an additional element in the strategy, a different treatment sequence, or an alternative that would replace another element in the strategy if the intervention were adopted. For example, in one study, the cost-effectiveness of a drug therapy for urinary incontinence was shown to depend on its place in therapy (e.g., used as a first-line or second-line therapy, after failure of an existing medication).²⁷ Alternative organizational models or packages of care (consisting of many different elements) may be compared. An example is a study that compared stroke rehabilitation in general medical wards to rehabilitation in specialized stroke units and to supported care in the home.²⁸ Strategies should be explained (e.g., when, under what circumstances, and for whom), and the elements of the alternative strategies defined.

It is good practice to anticipate future comparators, particularly lower cost technologies that may enter the market within the timeframe relevant to the analysis. Failure to do so can lead to an underestimation of the ICER of the new intervention. An example would be the anticipated entry of generic competitor drugs where it is known that the patent for the product will be

expiring.²⁹ The anticipated entry of lower cost technologies is also relevant to non-drug technologies (e.g., reusable versus low cost disposable surgical instruments). Such an analysis introduces additional sources of uncertainty in terms of the timing of entry and the price of the generic technology. Whenever feasible, such situations can be dealt with using a sensitivity analysis and a lower cost for the competitor technology.

Consider other reasonable alternatives. In some instance, “doing nothing” or “watchful waiting” may be appropriate comparators.

3.4.4 *Comparator drugs*

- The comparator need not be an alternative drug listed in the formulary of the relevant jurisdiction. As a starting point, it may be useful to consider alternatives in the same therapeutic class (i.e., drugs with the same indication). This may include drugs with the same chemical composition or mechanism of action as the new drug. For example, the comparators for a new quinolone antibiotic may not be limited only to other quinolones (as all quinolones may not be approved for the same indications), but rather all other antimicrobial agents that are clinically used to treat the infections the new quinolone is intended for. If the new drug is in a new therapeutic class, then the comparator is usual care (and recommended care, if appropriate), which may involve treatment with a drug from another chemical class, if available, or a non-drug treatment.
- Select as a comparator the available alternative that is of the lowest cost and that is often used for the same indication. The selection should be based on the cost of the entire recommended dose of the drug for treating the condition and any drug administration costs (report such costs separately).
- The regimen used for costing should reflect the dose and duration supporting the efficacy or effectiveness data for the product used in the evaluation. State whether the dosing regimens that are used clinically differ from those used in the clinical efficacy trials. Actual

(versus recommended) dosing can be determined by reviewing the literature, by examining utilization data, or by conducting a survey of clinical experts. Where appropriate, use the dosage of each individual comparator required to achieve the same therapeutic effect, and justify the dose equivalencies used.

- The report should state a drug’s generic and brand names, therapeutic classification, dosage form, route of administration, recommended daily dosage, duration of treatment, daily cost, and cost per usual course (for the drug and any comparators). The differences in terms of indications, contraindications, cautions, warnings, and adverse effects should be reported.

3.5 Perspective

The perspective chosen for the evaluation should fit the needs of the target audience.³⁰ The perspective in the Reference Case should be that of the publicly funded health care system. In some jurisdictions, this perspective may include costs that are incurred by long-term care, social services, or community-based services.

The costs associated with adopting a wider perspective should be reported separately where it is likely that they have an impact on the results of the analysis. This may occur when an intervention permits patients to return to work sooner than otherwise, shifts costs to patients and their families (hospital-based care versus home care), or results in savings or additional costs to other public sector agencies (e.g., special education for children with learning disabilities). These costs should be quantified separately where possible, and be subjected to a sensitivity analysis. Where quantification is difficult, the likely magnitude of such costs and their impact on the results of the analysis should be discussed.

The types of costs associated with the individual perspectives are detailed in Table 2. Resources should be identified, measured, and valued.

Table 2: Perspectives of economic evaluations and their related costs				
Perspective	Types of Cost	Examples		
Societal perspective	Public payer	Direct costs to publicly funded services (other than health care)	Social services, such as home help, meals on wheels* Income transfer payments paid (e.g., disability benefits) Special education	
		Publicly funded health care system	Direct costs to publicly funded health care system	Drugs, medical devices Equipment, space, facilities, and associated overhead costs Aids and appliances paid by government Health care providers and other staff Medical services, including procedures Hospital services Emergency visits Ambulance services Diagnostic, investigational, and screening services Rehabilitation in a facility or at home* Community-based services, such as home care, social support* Long-term care in nursing homes*
			Direct costs to patients and their families	Out-of-pocket payments (including co-payments) for drugs, dental treatment, walking aids Cost of travel for treatment, paid caregivers Premiums paid to, and benefits received from, private insurers† Income transfer payments received (e.g., disability benefits)
			Time costs to patients and their families‡	Patient's time spent for travel and receiving treatment Lost time at unpaid work (e.g., housework) by patient and family caring for the patient
		Productivity costs	Lost productivity due to reduced working capacity, or short-term or long-term absence from work (during friction period) Costs to employer to hire and train replacement worker for patient	

*Some of these costs may be incurred by the publicly funded health care system, depending on the precise nature of these costs and the relevant jurisdiction.

†The costs to private insurers (i.e., insurance premiums received from, and benefits paid to, patients) have been included as a direct cost of the patient. These amounts can usually be assumed to cancel out, unless there is a good reason to do otherwise. Private insurance premiums paid by employers as a part of an employees' compensation package can be included as a part of lost productivity costs.

‡The classification system in the table excludes some (indirect) time costs to patients and families caring for them. The value of lost time at paid work (beyond the friction period) is considered to be a private cost, and is excluded in the societal perspective. Lost leisure time is not considered to be a cost, as it would be at least partly captured in the preference measure when the QALY is used as the health outcome measure.

3.6 Effectiveness

3.6.1 Efficacy versus effectiveness

There is a difference between efficacy and effectiveness measures. Efficacy refers to the performance of a health technology under controlled circumstances, often in the context of randomized controlled trials (RCTs). Administered according to a strict written protocol by research-oriented clinicians, trial participants are often selected according to restrictive inclusion and exclusion criteria, then encouraged to comply with treatment, and monitored with care. These trials

may be conducted in specialized centres, such as teaching hospitals. In contrast, effectiveness refers to the performance of a technology in the "real world" (i.e., routine use in clinical practice), with a variety of providers using the technology as they deem appropriate for a broad heterogeneous group of patients who are usually less well informed, less likely to be screened or diagnosed correctly, less compliant, and subject to co-morbid conditions and treatments that are excluded in the original efficacy trials. In addition, provider compliance and skill (e.g., the volume-outcomes relationship with some surgical procedures), and coverage of the population (e.g., vaccines), may be lower for

some technologies and may adversely affect outcomes. Overall, “real world” patients often are less likely to respond to treatment than are participants in RCTs.³¹

Decision-makers are primarily concerned with the “real life” impact that the intervention will have on patients who will be treated in routine practice. A key issue is whether the efficacy data obtained from a RCT would reflect the effectiveness that might be achieved in a “real world” setting (i.e., the external validity of the clinical trial). Where feasible, the outcomes and costs in an economic evaluation should be based on the effectiveness of the intervention, rather than its efficacy, for the evaluation to be relevant to the jurisdictions.²

3.6.2 *Gathering best evidence*

A sound clinical review of the intervention should form the basis of the evaluation. This should involve a systematic review of the available evidence on the efficacy and effectiveness of the intervention. Justification should be given for failing to undertake a systematic review of the efficacy and effectiveness of the intervention.

Analysts should describe the studies that are included in the systematic review and the methods used to conduct the review. The review may include studies with a variety of designs, reflecting different levels of internal and external validity. The conduct and reporting of individual studies are important: studies that are poorly conducted and reported (e.g., small studies with missing data and little detail about participant disposal) may not provide much evidence. Complementary information provided by each form of evidence can be analyzed further through a sensitivity analysis. The methods used to analyze or combine data (e.g. meta-analysis, indirect comparisons) should be explained and justified, and the results of the analysis should be reported (see Guideline Statement 8.10).

Non-drug technologies can often pose challenges for economic evaluations. Many of these technologies go through less rigorous clinical testing before approval, and are often subject to less stringent regulation at the marketing approval and post-marketing surveillance stages as compared to drugs. Clinical studies may have a weaker design (e.g., lack of randomized controlled

trials) or the long term follow-up of study participants may be lacking. Accordingly, the evidence base for efficacy may be of a lower standard as compared to drugs.³² The best available evidence may include more use of observational studies or registry databases.

The outcomes (and sometimes cost) of medical and surgical procedures, and diagnostic technologies may depend on the skill and experience of the operator. The impact on outcomes can be difficult to measure. The evaluation of medical devices should focus on the entire episode of care rather than on only the technical performance of the device. For example, the evaluation of a diagnostic device may involve assessing the impacts that the sensitivity and specificity of the device have on follow-up care and health outcomes. An extensive sensitivity analysis may be used to evaluate situations where the evidence of efficacy and effectiveness is weak. Appropriate caveats about the speculative nature of such an analysis should be reported.

Adverse events that are associated with the intervention should be included in the evaluation where they are clinically or economically important, based on meaningful differences between the intervention and alternatives. Analysts will need to consider their nature, frequency, duration, and severity before such a judgment can be made. Adverse events may have an impact on:

- treatment continuation or persistence (e.g., switching treatments, dosing reductions, or discontinuations)
- patients’ compliance with treatment
- mortality
- morbidity
- HRQL (utilities)
- resource use (e.g., physician visits, hospitalization, or prolongation of existing hospitalization).

These events should be valued in a manner consistent with the principles outlined in the Economic Guidelines. It is preferable that the evidence for an improved adverse event or side-effect profile be based on primary, rather than secondary, outcomes from trials. Unimportant differences in minor side-effects that were identified as secondary outcomes in trials may not be clinically relevant. If many minor side-effects

are experienced simultaneously, the analyst should comment on the possible implications for adherence with treatment, which could have an impact on effectiveness. Furthermore, the full magnitude of potential harm associated with an intervention may not emerge during pre-market clinical trials, so other study designs will usually have to be used, such as the spontaneous reporting of cases or post-marketing clinical trials.

For drugs, the adverse events should be those associated with the expected dosing range for patients using the product. Total withdrawals due to adverse events for all the studies, with a breakdown by reason, and the number of patients requiring dose reductions because of drug intolerance, should be considered as indices of patients' tolerance to the drug. This can be explored through a sensitivity analysis.

3.6.3 *Modelling*

It is preferred that the Reference Case be based on the best quantitative estimate of “real world” effectiveness, with uncertainty about the estimate handled through a sensitivity analysis. Because of good “real world” evidence is seldom available before the intervention is used in the market, analysts are encouraged to translate efficacy data into effectiveness estimates, using the best available evidence and appropriate modelling techniques, where feasible and scientifically credible.

There should be a description and justification of the methods and assumptions used to translate efficacy data into estimates of effectiveness. Identify and evaluate those outcomes that are most feasible, given the data, and most important to the health of the patient. Where relevant, the analyst should describe the strength of the evidence for extrapolating or adjusting the data. This may involve extrapolating trial data on surrogate outcomes to final outcomes, or extrapolating data on short-term outcomes beyond the duration of the trial. It may also involve modifying the treatment effect from trials to account for “real world” factors that differ from factors in the efficacy trial (e.g., patients' adherence with treatment).

Depending on the nature of the available data, modelling may be required to transform the observed data to appropriate outcomes. Where only short-term data on outcomes are available, it may also be appropriate to extrapolate the data to long-term outcomes. Describe the strength of the evidence for extrapolating data, and justify any modelling approach or assumptions used.

One approach that is sometimes used to extrapolate short-term data involves the superimposing of estimates of baseline outcomes (e.g., probabilities of natural history survival that are derived from observational studies on estimates of treatment effect from clinical trials.³³ The validity of the extrapolation is often based on the quality of the epidemiological data, and the link between the risk factors that can be modified by the intervention and the long-term outcomes. The duration and the magnitude of the clinical benefit beyond the trial is often a critical judgment to make regarding extrapolation.

Analysts are encouraged to use modelling to link surrogate outcomes to more important patient outcomes. For studies using surrogate outcomes, the surrogate should be highly predictive of an important patient outcome. As an example, for cardiovascular disease, modelling may be used to link surrogate endpoints, such as blood pressure or cholesterol level, to clinical endpoints (e.g., the incidence of coronary heart disease), and subsequent final outcomes (e.g., all-cause and cardiovascular-related mortality and morbidity), depending on the evidence for such links. Such an analysis should be appropriately tested through a sensitivity analysis.

Several factors are often considered when estimating effectiveness using efficacy data:

- accuracy of diagnosis or screening
- patients' adherence with treatment
- health care providers' compliance or skill
- meaningful differences in how subgroups respond to treatment because of co-morbidities and the use of concomitant therapies that are not permitted in studies.

Some authors use the terms “adherence” and “compliance” as synonyms. The term “adherence” is used in the Economic Guidelines to encompass three components regarding a patient who is undertaking treatment: acceptance, the initial decision of the patient to accept treatment; persistence, long-term continuation of treatment; and compliance, the consistency and accuracy with which the patient follows the recommended treatment regimen.³⁴ Adherence is achieved if the patient meets all three components.

Correct diagnosis (or screening selection), patients’ adherence, and providers’ compliance or skill are often lower in “real life” than in RCTs. For technologies such as medical devices and surgical procedures, the outcome may depend on the skill and experience of the operator. Patients’ non-adherence can lead to a lower treatment effect than that observed in clinical trials, higher treatment costs, reduced productivity from the patient in the workplace, a greater burden on caregivers, and possible drug resistance. Non-adherence may be due to the adverse drug reactions experienced by the patient, or to other factors, such as frequency or ease of drug administration.

First, analysts should identify the “real world” factors that may modify the effect of the intervention based on the existing evidence. Second, where feasible and scientifically credible, “real world” factors should be incorporated into the Reference Case, where they are linked to important patient outcomes based on the best available evidence. For example, an evaluation in Husereau *et al.*²⁶ replaced the higher diagnostic accuracy in the clinical trial with a better estimate of diagnostic accuracy in community practice, to better reflect the effectiveness of influenza drugs when used in routine practice. The failure to account for “real world” factors in evaluations may lead to the selection of suboptimal treatment strategies.

What is key to making a judgment about incorporating “real world” factors into the analysis is the strength of the available evidence linking treatment-modifying factors with important patient outcomes. It is preferable that such evidence be based on primary, rather than secondary, outcomes from trials. Such data are usually lacking before

the launch of the intervention, and data from other sources may be useful. For example, there may be data on “real world” adherence patterns in other countries, or data from retrospective databases.

3.6.4 *Uncertainty and stratified analyses*

Before the “real world” experience with the intervention, there may be a high degree of uncertainty about the estimates of factors that have an impact on effectiveness. Where the data are unavailable, or are of low quality, uncertainty about the effectiveness estimates should be assessed through a sensitivity analysis. Unless such an analysis is based on high quality evidence, it should be identified as speculative or exploratory, and appropriate caveats should be given in the report.

For example, the duration and magnitude of the clinical benefit can be modelled based on plausible scenarios. The magnitude of effectiveness over time can also be varied in different scenarios, from “no effect” beyond the trial period, to a “diminishing effect,” to a continuing “full effect.”³⁰ Furthermore, scenarios with longer time horizons may include more important patient outcomes (e.g., serious liver complications from hepatitis C infection), and QALYs. The shorter periods may be more appropriate for surrogate outcomes (e.g., a sustained viral response to antiviral treatments for hepatitis C infection).

An alternative approach is the use of a Bayesian iterative framework for gathering and analyzing evidence. The analyst can translate efficacy data from pre-market trials into effectiveness using prior evidence (or assumptions). In some instances, this may be formed by empirical evidence when an intervention has been launched in other markets before entering the Canadian market. It may also be reasonable to examine existing comparable interventions in Canada. In other instances, this may be a subjective prior belief, based on information from expert physicians or others who might have knowledge about the parameter. After the launch of the intervention, the preliminary estimates of cost and effects can be updated, as “real world” cost and effect data are collected.

A stratified analysis should be used to assess the impact on the results of variation in the effectiveness of an intervention among subgroups in a target population. This can be used to assess the variability of effectiveness in subgroups due to differences in, for instance, risk profile (e.g., high risk versus average risk patients), condition incidence, condition progression, or access to health care. The level of certainty in parameters for subgroups is lower than that in the total population. Therefore, where a stratified analysis is undertaken, a sensitivity analysis should explicitly consider increased uncertainty.

3.7 Time Horizon

The time horizon should be long enough to capture all the meaningful differences in costs and outcomes between the intervention and comparators. It is unnecessary to extend the time horizon beyond the period where there are no meaningful differences, such as when the costs and outcomes of alternatives converge. The same time horizon should be applied to costs and outcomes for analytical consistency.

Analysts are encouraged to consider a lifetime time horizon as a default, particularly for chronic conditions (e.g., diabetes or rheumatoid arthritis), or when the alternatives have differential effects on mortality. If a shorter period is used (e.g., for acute illnesses), justification should be provided. In some cases, multiple time horizons might be appropriate for the extrapolated data. For certain chronic conditions, alternative scenarios with time horizons of one and five years, and a longer period may be appropriate.

A long-term analysis does not imply that primary data must be collected from patients over such a period. Because long-term data collection is often unfeasible or impractical, a short-term analysis based on actual data collected on intermediate (or surrogate) outcomes may be complemented by a longer term analysis of more important patient outcomes that are appropriate for the condition, based on extrapolated or modelled data. For example, in the case of highly active antiretroviral therapies for HIV infection, an analysis using a short-term time horizon on trial data for viral response and immune system surrogate markers

can be supplemented by a longer term (lifetime) time horizon to take account of serious opportunistic infections and premature death, which may occur years later.

To assess the impact of extrapolation techniques used in the analysis, the analyses of the alternative time horizons should be presented separately. The best available evidence should be used, and the causal relationships, techniques and assumptions used to extrapolate data should be explained.

3.8 Modelling

Economic evaluations of health care technologies typically involve building and then using models to synthesize evidence and assumptions from multiple sources to estimate the long-term incremental costs and outcomes of new therapies. Because the outputs (results) depend on the model structure, the data, and the assumptions used, the model should be as transparent as possible. As a result, decision makers should be critical when reviewing the results of a model-based evaluation.

3.8.1 Modelling considerations

A good model³⁵ can be defined as one that is:

- tailored for the intended purpose
- useful for informing the decisions at which it is aimed
- readily communicated.

Analysts should follow good practices to ensure the quality of their model, and their analysis. The good modelling practices summarized here are drawn from two guidance documents. For details about the guidance provided, refer to Philips *et al.*³⁰ and Weinstein *et al.*³³ In addition, consider all sections of the Economic Guidelines in the design of a model.

The scope (i.e., the boundaries) of the model should be explained and justified. The feasibility of building a model should be assessed before coding it. The study question is fundamental to developing the model. Once the question is defined, the analyst can determine whether modelling is the best approach to the problem, and define the most appropriate techniques to use.

The scope, structure, and parameters of the model should be relevant to the study question and the needs of the target audience. The model should incorporate all facets of the condition of interest that are important, and the potential impacts of the interventions considered. The model should be flexible, so that it is adaptable to the circumstances of the jurisdictions or payers (e.g., allow for variable treatment patterns).

The overall design and structure of the model determines the range of analysis that can be performed. Models should facilitate the type of economic evaluation that is relevant to the study question, and allow for an adequate assessment of uncertainty surrounding study results. The analyst should remember the iterative nature of economic evaluations and build the model to permit the updating of results as more data become available. For instance, a model may be structured to permit the incorporation of adverse event data from an ongoing clinical trial.

The model structure, values, and sources for each input parameter should be justified. The assumptions and the subjective judgments about the model structure (e.g., relationships, variables included, distributions) should be justified to enable the users to evaluate their acceptability. The model should only be as complex as required to properly address the study question and the information needs of the target audience.³⁰

The structure of the model refers to the specification of the condition or treatment pathways, the associated clinical events, and the causal relationships.

The model structure should be consistent with the underlying theory about the condition, should capture the impact of the intervention and alternatives, and should be relevant to the study question. The structure should not be defined by current patterns of practice, because the model should be able to evaluate changes to practice patterns. Clinical events that are not logically or theoretically expected to differ between the intervention and alternatives can be excluded from the structure. It is recommended that a diagram showing the condition or treatment pathways (e.g., decision tree) be included in the report.

The availability of data should not unduly constrain the design of the model, because funding or reimbursement decisions must often be made in the absence of data. Data limitations may limit the scope of the model, and can be considered when designing detailed aspects of the model. Data limitations may be handled using techniques such as surveys, expert opinion, the use of place holders in the model, or a sensitivity analysis.

When state-transition (Markov) models are used, the cycle length should be defined and justified. The length should be the minimum interval over which the pathology or symptoms are expected to change.³⁰

Data extrapolation (e.g., short-term to long-term outcomes) should be based on appropriate methods. The extrapolation relationships and techniques, and any assumptions used in the model should be explained and justified, with reference to the strength of the supporting evidence. The linear extrapolation of cost data may be inappropriate because of, for instance, economies of scale.³⁶

Models should be formally validated.^{33,35,37} Validation involves testing the model to confirm that it does what it is expected to do.

Internal validation confirms that the results generated by the model are internally consistent. This is done by testing the mathematical logic of the model, and by checking for errors in the mathematical code and for administrative errors (e.g., labelling, spelling errors). The practice may include testing extreme or zero values, examining the results of known scenarios, and examining the code. Explain any counterintuitive results. A more elaborate test involves building the model in two software packages, and cross-checking the results.³⁰

External validation confirms that the basic model structure, assumptions, and parameters are reasonable and accurately reflect the condition process, and the impact of the intervention and comparators. The results should make sense at face value, and any that do not should be explained. Sensitivity analyses can be conducted to assess the uncertainty about the structural assumptions used in the model (e.g., techniques for extrapolating data, conversion of important clinical outcomes to

final outcomes, treatment pathways). The results may be compared to those from other models (i.e., between-model validation). The intermediate results of the model (e.g., health outcomes) should be calibrated or compared against reliable independent data sets (e.g., national cancer statistics). Any differences should be explained, or used to inform adjustments to the model.

The validation process should be documented, and such documentation should be made available to decision makers, if requested. Ideally, the model validation should be undertaken by someone impartial. The analyst should state the limitations of the model, in the report and whether, and if so how, the model has been validated.

3.8.2 Data considerations

Many issues relate to the appropriate identification, collection, and analysis of data and their incorporation into a model. For guidance on such issues, which are beyond the scope of the Economic Guidelines, readers are referred to Philips *et al.*³⁰ and Weinstein *et al.*³³

Model data may come from sources such as RCTs, observational studies, administrative databases, disease registries, expert opinion, standard cost lists, and assumptions made by the analyst or specified in guidelines (e.g., discount rate). The choice of data should be appropriate for the study question and for the needs of the target audience. The data should be consistent with the design features of the model (e.g., perspective), and be relevant to the population affected by the intervention. All data should be reported, and sources identified. Details of the data, such as the population from which data were derived and to which the results apply, should be described. If data in the model are not directly comparable (i.e., they relate to different patient samples), choices or assumptions should be explained and justified.

The more reliable the data that are used to estimate model parameters, the more credible are the results of the model.³⁸ The choice of data sources (e.g., selection of studies used in a meta-analysis), the choice of methods for analyzing data inputs (e.g., handling of trial drop-outs in intention-to-treat analysis), and the subsequent incorporation of data into the model can have a bearing on the results of the evaluation uncertainty and generalizability of

the results.³⁹ Data limitations should be made explicit, and the methods for handling them described. Attempts should be made to quantify the impact of these limitations on the uncertainty of the evaluation results.

The design of a study can influence the quality of the data and the results of the evaluation. It is inappropriate to choose only favourable (or unfavourable) trials or data when estimating the outcomes and costs of the intervention. The quality of the data used in the model should be described. Different instruments can be used to rate the strength of the evidence from various types of studies.⁴⁰ Each form of evidence may add complementary information that may be useful for adjusting data for local populations, costs, and practice patterns, or as a basis for a sensitivity analysis.

Caution should be exercised when using expert opinion to establish parameter values. Justification should be given for using expert opinion (e.g., lack of appropriate data from other sources), and the source of the opinion, the method of elicitation, and the results of the exercise should be described. Uncertainty about such estimates should be appropriately assessed through a sensitivity analysis.

Systematic reviews and meta-analyses can produce high quality data for model parameters, and add to the credibility of economic evaluations.³⁸ Systematic reviews also provide useful information for analyzing uncertainty surrounding the relevant estimates. Attention should be paid to those key model parameters for which the results are most sensitive.^{30,33} Justification should be given for failure to undertake systematic reviews of key model parameters based on the adequacy and generalizability of readily obtained data.³³

Data can be incorporated into the model as point estimates or as distributions in the case of a probabilistic sensitivity analysis. The process for doing so should be explained. If data are incorporated into the model as point estimates, the mean estimates of parameters should be used in the base case.⁴¹ Data requiring transformation should follow generally accepted methods of biostatistics and epidemiology. The methods and results of these data analyses should be provided.

3.9 Valuing Outcomes

3.9.1 Health-related quality of life

“As a construct, health related quality of life (HRQL) refers to the impact of the health aspects of a person’s life on that person’s overall well-being. Also used to refer to the value of a health state to an individual.”⁴² Many methods have been developed to measure HRQL. They can be divided among specific measures, generic measures, and preference-based (utility) measures.⁴³⁻⁴⁵ Specific measures (e.g., Western Ontario-McMaster Osteoarthritis Index) and generic measures (e.g., Short Form 36) are generally of limited value for economic evaluations. Preference-based measures provide a summary score that numerically reflects the HRQL, and are the only approaches that are suitable for use in a cost-utility analysis (CUA).

Appropriate preference-based measures should be used where there are meaningful differences between the intervention and alternatives in terms of HRQL, and where appropriate data on preferences are available. If HRQL is being measured in a prospective study, it is advisable to include a preference-based measure where the intention is to undertake an economic evaluation. Where this has not been done, preference scores can be gathered retrospectively through a separate exercise, and then mapped onto the outcomes of the efficacy-effectiveness trial. Alternatively, preferences can be obtained from secondary sources provided that they are appropriate for the population of interest.

The terms “preference” and “utility” are generally used synonymously as a measure of HRQL in the Economic Guidelines, although technically, “utilities” are preferences obtained by methods that involve uncertainty (i.e., the standard gamble approach).¹

3.9.2 Quality-adjusted life-years (QALYs)

The preferred measure for a CUA is the QALY, which is calculated by multiplying the number of life-years gained from an intervention by a standard weight that reflects the HRQL during that time.^{1,46,47} The QALYs is preferred because of its clarity, simplicity, ease of application, and face validity.

Typically, to be suitable for calculating QALYs, preferences are measured on a cardinal (i.e., interval) scale on which states of health equivalent to immediate death are scored 0.0 and perfect health is 1.0. States that are worse than death are allowed on this scale, and would take on scores less than 0.0. Preference-based scores (i.e., the quality-weight for a CUA) can be measured directly or indirectly. Justification should be given for using alternative preference measures in a CUA, such as healthy-year equivalents (HYE) or saved-young-life equivalents (SAVE).

The direct measurement of preferences is a complex and costly task. Three methods are used for the direct measurement of preferences: standard gamble, time trade-off, and visual analogue scale.⁴⁸⁻⁵⁰ Analysts prefer the standard gamble approach because of its strong normative foundation in von Neumann-Morgenstern utility theory.^{1,51} There are arguments against the superiority of the standard gamble approach.⁵² Visual analogue scales are inappropriate to use alone because of well known biases.¹ Analysts wishing to undertake the direct measurement of preferences should select an approach that has theoretical and empirical properties to address the problem in hand, and should justify their selection. To avoid double-counting, respondents should be told, where feasible, to value lost leisure time in terms of the changes in preferences, and to assume that health care costs and income losses are fully reimbursed.

“Off the shelf” instruments are available for obtaining utilities without undertaking direct measurement. Some widely used instruments in this category are the Health Utilities Index (HUI),⁵³⁻⁵⁵ the EQ-5D,⁵⁶⁻⁵⁹ the SF-6D,⁶⁰ and the 15D.⁶¹ These instruments use preferences from the “informed” general public, which is the appropriate source to use for collective resource allocation purposes. Some of them, such as the HUI, ask survey respondents to exclude income effects when valuing health states.⁶² To use these instruments, the analyst has to classify the patient’s health status into the system provided, and compute the utility from the formula. The score represents an estimate of the mean utility that would be given to the health state by a random sample of the general public, though it may lack the sensitivity of the direct measurement approaches.

Analysts are encouraged to use indirect measurement instruments, because they are easy to obtain, compare, and interpret. The direct measurement of preferences is acceptable, and may be better suited to some situations (e.g., a condition where one area of function is predominately affected).

It is recommended that analysts study the alternative methods *a priori* and select in advance the one that is most appropriate for the condition and best suits the study question. Justify the selection and method, and explain the steps undertaken to measure preferences. It is inappropriate to try a variety of methods, and choose the one that puts the intervention in the best light. Brazier *et al.*⁶³ provides a useful checklist (in Box 3 of the report) for judging the merits of preference-based measures of health, based on the practicality, reliability, and validity of the instrument.

Regardless of the approach used, report changes in the quantity of life and quality-weights separately, and be transparent in how the two are combined. Assumptions about quality-weight changes over time (e.g., linear, curve) and the method used to estimate QALYs (e.g., change from baseline score, total area under the curve) should be reported.

A concern with using QALYs is that they do not discriminate between conditions with different severity. The conventional QALY approach focuses on absolute changes in preference scores, whereas studies have shown that societies' valuation of interventions for different groups also depends on the severity of the initial condition.²⁰ To partly address this concern, analysts can conduct a sensitivity analysis that excludes the QALYs of those outcomes that may not be considered clinically important (e.g., short-lived, self-limiting, and non-severe outcomes) and may be of limited concern to decision makers. For example, in the case of a new drug for the treatment of influenza, a sensitivity analysis could show the impact of including only influenza-related complications that are severe enough to require a visit to a physician or hospitalization, with flu-days prevented by treatment being excluded from the analysis.

Controversy exists regarding whose preferences should be used for deriving the quality-weights: the patients who experience a particular health state or a representative sample of the general public (community) who have not.⁶³ The major indirect measurement instruments are based on surveys of preferences from the general public. Direct approaches to measuring preferences for specific conditions often use the patients being studied, perhaps in the context of a trial. There is evidence that preference valuation varies by condition experience.⁶³

It is preferred that analysts measuring preferences directly use a representative sample of the general public, who are suitably informed about the health states being valued. The reasoning is that they are the ultimate payers of the publicly funded health care system and potential patients.⁶⁴ Patients who have direct experience of the relevant health states may be an acceptable source of preferences. An analyst undertaking direct measurements should describe the population from which the preferences were derived and the methods of measurement.

It would be ideal to use the preferences of the general public in the Reference Case and patients' preferences in a sensitivity analysis, although this may be impractical or unnecessary. The analyst should discuss the applicability of the estimated preferences to the Canadian population.

Some studies use expert judgment with an extensive sensitivity analysis as the source of quality-weights. This approach is not favoured. Where the results can be shown to be insensitive to the quality-weights, approximate estimates may be adequate.

Where the intervention has an impact on the quality of life of the patient's caregiver, this can be measured, though care should be taken to avoid double-counting the costs of caregiving. Changes in the quality of life of caregivers should be reported separately in the analysis, and excluded when calculating the ICER.

3.9.3 Outcomes for cost-benefit analysis

The monetary values assigned to health outcomes in a cost-benefit analysis (CBA) are usually obtained by applying a willingness-to-pay (WTP) approach. Two of the methods used in WTP

studies are contingent valuation and conjoint analysis (also known as discrete choice experiments). Contingent valuation uses a hypothetical survey to estimate an individual's maximum WTP for a good or service that usually does not have a market price.⁶⁵ A conjoint analysis uses ranking, rating, or comparison exercises to estimate the relative weights that people attach to different attributes of a good or service (including health care).^{66,67} In health care, a conjoint analysis has been used to value non-health outcomes and process attributes.

WTP methods for valuing outcomes are evolving, and several methodological and ethical issues need to be resolved.^{22,68} These methods of valuation have not been widely used in the context of resource allocation decisions in the health care sector. More research is needed to validate the use of these methods for informing health care funding decisions.

The use of a CBA in general, and WTP methods in particular, should be regarded as a secondary type of analysis. With a CBA, the evaluation should explain the steps taken to convert the outcomes into monetary terms. The key assumptions of the analysis should be validated and thoroughly tested through a sensitivity analysis.

3.10 Resource Use and Costs

3.10.1 Resource identification

In this step of the costing process, the analyst identifies those activities and resources that are likely to occur in each alternative, along with timelines. The study perspective(s) will determine which resource items to include or exclude from the analysis, some of which are outlined in Table 2.

It is recommended that costs included in the public payer perspective be classified into categories that are appropriate to the relevant decision maker. The evaluation should group costs incurred by the different sectors of the public payer (e.g., primary care, hospital, community care), and present in a disaggregated manner those costs that are the responsibility of the decision maker. For example, evaluations for a provincial drug plan should provide a breakdown of costs associated with

using the drugs. Where there is cost shifting between public payer sectors, the evaluation should quantify the relevant costs (e.g., fewer hospital bed-days and more home care visits). The public payer perspective should aggregate the costs of all public payer sectors.

Current and future costs that are a consequence of the intervention should be included in the evaluation. Identifying costs that are associated with non-drug technologies may be more complex than doing so for drugs. Costs may include start-up costs, capital costs, operating costs, costs for maintenance and repair, costs of hiring additional staff, overhead costs, and costs for professional training. These costs should be included, where appropriate. For example, when evaluating a hip prostheses, the cost of the entire episode of care and all other related costs (e.g., training costs) should be included, and not just the cost of the device. Resource items can be excluded from the analysis where there is identical use between the intervention and alternatives, though analysts should justify this.

Protocol-driven costs from a clinical trial should be excluded from the evaluation. Income transfer payments (e.g., disability and employment payments) should be excluded in the analysis, because they are not borne by the publicly funded health care system and are not real costs to society (they cancel out). Analysts may wish to report these costs when they are significant.

A cost item may be deemed to be irrelevant, because it is influenced by an event that is unrelated to the intervention being evaluated (e.g., the cost of a broken leg would not normally be counted in evaluating an acne drug). One option for determining which clinical events are related is via an adjudication committee (blinded to treatment assignment). This would allow the analyst to remove unrelated events in an unbiased manner.

A contentious, unresolved issue in the economic literature is that of unrelated health care and non-medical costs that are incurred during the life-years gained from the intervention.⁶⁹ Analysts can use their discretion about whether to include such costs, but this should only be done in a sensitivity analysis, and such costs should be identified. One

option is to exclude these costs if they have a small impact on the overall results. Include these costs in a sensitivity analysis if data are available and if their impact is substantial.

3.10.2 Resource measurement

For the purpose of transparency, resource use data for the intervention and alternatives should be reported in physical units.

There are two costing methods, though many evaluations use a combination of the two.⁶⁹ Gross costing (top down costing) uses large components as the basis for costing, such as the cost per hospital day. Detailed micro-costing (bottom-up costing) on an item by item basis can produce more precise estimates, although the time and expense of collecting such data need to be considered. The analyst must be clear about the costing method used and must justify the approach. Guidance in the use of costing methods is available in several sources, including Baladi,⁷⁰ Oostenbrink *et al.*,⁷¹ and Gold *et al.*⁴

Resource items that contribute most to total and incremental costs should be measured and valued with greater precision. This can be done a priori by conducting a sensitivity analysis of resource use and unit cost parameters in a model to determine the expected impact on total or incremental costs and the results.⁶⁹

Analysts should pay attention to the following when deriving cost estimates.

- Where possible, explain the method for allocating overhead costs, shared labour costs, and administrative costs.
- Consider the relationship between the quantity of resources used and the unit cost estimate, such as learning curve effects (e.g., for surgical procedures), which can reduce the future cost of a resource; the impact of a new program on existing infrastructure, where the scale or scope of a program is a factor; the utilization capacity used in the cost estimate (where relevant, explain the method used to adjust for normal operating capacity).

Furthermore, analysts should provide an assessment of the quality of the estimate, and where lower quality estimates are used, use a

sensitivity analysis to determine the impact of cost assumptions. Aspects of quality include how much of the total resources are included in the estimate, whether the output indicator is sufficiently refined to capture resource use, the quality of the data source, and whether the estimate of cost can be generalized to all providers.

Resource use data can be obtained from several sources, including RCTs, administrative and accounting data, clinical practice guidelines, expert opinion, and modelling exercises (combining data from a variety of sources). Data will vary considerably in terms of the quality of estimates and their applicability to Canadian practice. There are issues related to translating the resource quantities obtained from experimental practice studies and international studies to Canadian practice. The applicability of the data that has been obtained to Canadian practice should be justified.

3.10.3 Resource valuation

Resources should be valued at their opportunity cost, the value of their best alternative use.⁶⁹ The guiding principle in measuring the unit costs of resources is that these costs should measure all resources required to produce an additional unit in the long run. These resources will include capital costs, all operating costs, allocated overhead costs, and professional costs. For example, hospital costs should include the physicians' fees related to the patient's stay. In practice, many methods have been used to approximate this concept.⁷² It is recommended that analysts use the total average cost (including capital and allocated overhead costs) as the unit cost measure.⁷¹

There are several ways of valuing resources, including market prices, administrative fees, direct measurement, and calculation of shadow prices.⁷³ For many resources, there is no straightforward method of valuation (e.g., informal caregiving).

There is no consensus regarding the best method of valuation though the following points should be considered.

- Use market prices, where available, unless there is a good reason to do otherwise (e.g., excessive profits).
- For consistency and convenience, an argument can be made for using standard costs. An

updated listing⁷⁴ of standard costs for some of the more common services in various provinces was prepared in 2004 using 2002 costs. It is available from the Institute of Health Economics' web site (<http://www.ihe.ca>). While this list has been used in a variety of circumstances, economic evaluations are conducted in many settings and under different circumstances, not all of which can be covered in one cost list.

- Where costs are directly calculated or imputed, they should reflect the full economic cost of all relevant resources at normal operating levels.

Where the target audience is more than one jurisdiction, costs in the largest jurisdiction or the average Canadian costs (if available) should be used in the Reference Case, with sensitivity analyses using the highest and lowest costs from the relevant jurisdictions.

When evaluating the public payer perspective, use the full cost (i.e., contributions paid by the publicly funded health care system, private insurers, and patients) of the intervention and comparators in the Reference Case. For interventions likely to involve cost sharing between the public payer and patients (e.g., co-payments for drugs prescribed outside of hospital), analysts should use a sensitivity analysis to assess the implications of variations in the proportion of the cost of the intervention and comparator paid by the public payer. For example, the analyst may use a typical proportion of the cost borne by the public payer in a DSA, or specify a distribution of costs likely to be borne by the public payer in a PSA. Use the same proportions for the intervention and comparators, unless there is a good reason to do otherwise. This will allow decision makers to use the proportion that is most relevant to the coverage provisions in their jurisdiction. Cost sharing arrangements can be ignored where the amounts paid by the patient for the intervention and comparators cancel out (e.g., an identical flat fee paid by the patient for two alternative drug regimens), or where they have a small impact on the ICER. Analysts should be transparent in reporting how these cost sharing arrangements were handled in a sensitivity analysis and the impact on the ICER.

Unit prices may only be available for a previous time period, or for another country. Prices that were obtained from previous years should be updated. There are no price indices in Canada for hospital or physician services, so a general price index, such as the Consumer Price Index, can be used.

Analysts are cautioned about using prices or unit costs from other countries for the Canadian setting. Such data should be verified for Canada. Where such data are adjusted, appropriate methods should be used, explained, and justified.

3.10.4 *Lost time*

The condition may result in the patient giving up activities that would otherwise be undertaken, because of time spent for travel and receiving treatment, time spent being ill, and premature death. Family, friends or volunteers may devote unpaid time to caring for the patient. This results in the sacrifice of time that could be spent doing other activities. This can be important for more severe chronic conditions (e.g., migraine), or where an intervention involves a shift from formal to informal care. The impact of the intervention on lost time by patients and informal caregivers can be quantified in a separate analysis, where it is likely to have an impact on the results of the analysis.

There is debate about the appropriate approach for including lost time in an evaluation. The issue can be framed by viewing time in terms of the health-related activity that is undertaken, and in terms of activity that is given up (i.e., opportunity cost).⁷¹ The activities in the latter approach can be divided into three categories: paid work, unpaid work (e.g., housework), and leisure.

For the patient, the time affected by ill health can include health-related activities, such as time spent travelling and receiving treatment, or time lost in terms of activities that are given up.^{71,75} Lost time in paid work includes lost productivity while at work, short- or long-term absences from work, and lost productivity due to premature death of the patient. Informal caregiving by family, friends, or volunteers can be measured in terms of time devoted to caring for the patient, or time lost that could be spent doing other activities.

Two approaches for valuing lost time at paid work are the human capital approach and the friction cost approach. In the human capital approach, the employment period that is lost due to illness is the measure of the production cost. In this approach, it is assumed that there is (near) full employment, and the value of lost production may be overestimated. In the friction cost approach, it is assumed that when a person is out of work, he or she is eventually replaced with an unemployed person, so that the productivity loss to society is limited to the time before the sick person is replaced (the friction period). In the friction cost approach, lost productivity due to premature death should not extend beyond the friction period. For short-term absences from work, the patient's lost production may be partly restored by the patient when he or she returns to work, or by the company's internal labour resources. When the time lost from paid work is short, the estimates from the two methods may not be different. For longer periods, the friction cost approach will result in a lower cost estimate compared with the human capital approach.⁶⁸

Some view the friction cost approach as taking a more pragmatic view of the functioning of the labour market compared to the human capital approach.^{71,76} On balance, the friction cost approach is preferred for valuing lost time at paid work. This approach has drawbacks.⁷⁷ Specifying the length of the friction period is required to operationalize the approach. The friction period can vary by industry, occupation, macroeconomic climate (e.g. general unemployment levels), and efficiency of the matching process between job seekers and vacancies. No data are available on the appropriate length of the friction period in Canada. The length of time needed to fill job vacancies has been suggested as a proxy for the friction period. Another approach would be to use the national average friction period that is estimated for countries with a comparable economic climate and labour market as that of Canada, with the length of the friction period tested through a sensitivity analysis.⁷⁶ This approach was used by Goeree *et al.*⁷⁸ to estimate the friction costs associated with schizophrenia.

The unit cost of long-term absences from work (i.e., absences longer than the friction period) due to mortality or morbidity can be valued using age–sex adjusted wage rates plus wage supplements for employer contributions for benefits such as pensions and employment insurance.⁷⁶ This value can be adjusted for the percentage of persons employed in each population group. Population-level data for these components are available from Statistics Canada. The friction cost includes costs associated with the recruitment and training of replacement workers.

There are several options to valuing unpaid work time (including informal caregiving), but there is no preferred alternative.⁶⁹ One option is to use the replacement cost estimate based on the market value (i.e., gross wage) of the services being delivered to the patient at home, and using the time spent caregiving or the time that a professional would have spent performing the duties. Another option uses the (net) wage that the informal caregiver would have received if working for pay during the caregiving period (i.e., reservation wage). A third option is to divide the time spent caregiving into lost time from paid work, unpaid work, and leisure, and to value each period using one of the other approaches outlined in this section. It is also possible to value informal care using the contingent valuation method or a conjoint analysis, though these are less common.⁷⁹

As suggested by several authors,^{69,71,80,81} it is recommended that lost leisure time be excluded as a cost in evaluations, where feasible. Lost leisure time would be partly captured by the preference measure where the QALY is used as the health outcome measure. Where possible, subjects in exercises measuring preferences should be told to value changes to leisure in terms of preference changes and to assume that health care costs and income losses are fully reimbursed.

Care should be taken in the measurement and valuation of patient and informal caregiver time. Describe the methods and assumptions used to measure and value lost time. Because the value of lost time can depend on of the approach used, analysts should use a sensitivity analysis to evaluate alternative methods and assumptions.

3.11 Discounting

Costs and health outcomes should be discounted to present values when they occur in the future, to reflect society's rate of time preference.

Accordingly, any costs or outcomes occurring beyond one year should be discounted using standard methods.

For the comparability of results across evaluations, it is important that a common discount rate be used. The standard rate for the Reference Case is set at 5% per year. A rate of 0% should be analyzed to show the impact of discounting. In addition, a 3% discount rate must be used in a sensitivity analysis for a comparison with published evaluations in other jurisdictions using 3% as the standard discount rate.

The discount rates in the Reference Case and the sensitivity analysis are expressed in real (constant value) terms, which are consistent with valuing resources in real (i.e. constant, inflation-adjusted) dollars.

Readers are referred to Gold *et al.*,⁴ West *et al.*,⁸² and Smith *et al.*⁸³ regarding discounting practices, and Lazaro *et al.*⁸⁴ for discounting health consequences.

3.12 Variability and Uncertainty

It is important to distinguish between variability and uncertainty. Variability reflects the known differences in parameter values that are associated with identifiable differences in circumstances.⁶⁸ It is represented by frequency distributions, and cannot be reduced. Uncertainty occurs when the true value of a parameter is unknown, thus reflecting the fact that knowledge or measurement is imperfect. It can relate to parameter values in a model (e.g., resource use, utilities, effects) and model design features (e.g., model structure, analytical methods, and assumptions). These are called model inputs in the Economic Guidelines.

3.12.1 Handling variability

Variability in practice patterns must be handled through further analysis. This may involve a sensitivity or scenario analysis, a type of multi-way sensitivity analysis in which there is

simultaneous substitution of alternative values in the relevant model parameters.

Patient heterogeneity relates to different individual characteristics (e.g., "risk profiles") of the patients in the analysis. For example, treatment effects may differ for an elderly high risk population and a middle-aged low risk population. In the target population for an intervention, different groups of patients may differ in terms of cost-effectiveness.

The role of the analyst in a modelling study is to uncover heterogeneity in data relating to costs, outcomes, and preferences (utilities) in stratified analysis. Coyle *et al.*²⁵ have suggested a framework for addressing the variation in cost-effectiveness in a target population. The framework involves the stratification of the target population into more homogeneous patient groups. In this analysis, each subgroup is run through the model separately. Thus, those groups for which treatment is cost-effective (e.g., those that have a positive net monetary benefit) can be identified. Where populations are heterogeneous, it is good practice to have separate models or versions of models for each patient subgroup.

There are instances where differences between groups emerge in the process of an exploratory data analysis. Preferably subgroups should be defined at the planning stage of the outcome study. A post-hoc analysis may be unacceptable unless strong justification is given. Information derived post hoc should be interpreted cautiously, and can be viewed as introducing further research questions rather than facilitating current resource allocation.

3.12.2 Handling uncertainty

In most cases, economic evaluations use decision analytic models to estimate the cost-effectiveness of interventions. Uncertainty in the context of an economic evaluation can be broken into parameter uncertainty and model uncertainty.

A summary of recommended approaches for handling the different types of variation and uncertainty is provided in Table 3. Appendix 1 provides suggestions for graphically presenting the results of the analysis of variability and uncertainty.

Table 3: Recommended Approaches for Handling Variability and Uncertainty

Category	Type of Variability or Uncertainty	Recommended Approach
Variability	Differences in clinical practice patterns between geographic areas or settings	Sensitivity analysis
	Variability in patient population (patient heterogeneity)	Stratified analysis
Model-based uncertainty	Model uncertainty <ul style="list-style-type: none"> • analytical methods • model structure • assumptions • data sources 	DSA using alternative assumptions, one-way, multi-way, threshold, or extremes analysis; and model validation methods
	Parameter uncertainty	DSA using one-way, multi-way, threshold, or extremes analysis PSA using Monte Carlo simulation is encouraged

Parameter uncertainty refers to uncertainty in a probabilistic sense around the true value of a parameter in the model, such as health outcomes, utilities, and resource use. Where a parameter value is based on a sample, there is uncertainty in the estimate due to random chance. In most evaluations, data on costs and outcomes are not directly observed at the patient level. Instead, analysts typically use synthesized information to estimate the parameters related to costs and outcomes of the treatments under investigation. Generally, this information comes from a variety of sources, such as prospective primary data, estimated data from published and unpublished studies, expert panels, and assumptions that reflect judgment.

Model uncertainty is uncertainty that is related to other types of inputs in the model whose values are not necessarily uncertain in the probabilistic sense, but whose values are known to be different. Model uncertainty depends on the choices and assumptions made by the modeller. Model uncertainty includes:

- choice of methods and assumptions made by the analyst for the Reference Case (e.g., time horizon, discount rate or method for valuing lost production)
- structural assumptions made in model (e.g., techniques for extrapolating outcomes beyond trial follow-up, predicted duration of treatment effect)
- choice of data sources (e.g., choice of studies used in a meta-analysis, use of data pooling, data stratification).

There should be a thorough and systematic assessment of the impact of uncertainty on the results. Consider the reliability of the model at every level, including the model structure, the analytical methods, the input parameters and assumptions, and the output. Proper consideration of uncertainty is required, so that the reader can judge whether the conclusions are meaningful and robust.

Given the many assumptions in models and the difficulty knowing a priori which inputs (or combination of inputs) have the greatest impact on the results, at a minimum. DSA should be performed for all model inputs. The omission of any model input from a sensitivity analysis should be justified (for instance, based on a preliminary analysis). The objective of this analysis is to examine if the overall results are sensitive to a plausible ranges of values for the relevant model input, such that the interpretation of the results would change. Methods for conducting this analysis include using alternative assumptions, one-way and multi-way sensitivity analyses, a threshold (or break-even) analysis, and an analysis of extremes (i.e., best and worst case scenarios).

Analysts should define the plausible ranges of values for each model input, and justify the selected ranges. The ranges should reflect the full range of variability or uncertainty that is relevant for each model input, based on evidence about appropriate, credible limits. Drummond *et al.*¹ suggest that a plausible range can be determined

by reviewing the literature, consulting expert opinion, and using a specified confidence interval around the mean (for stochastic data).

The following points are provided to help guide the conduct of a DSA.

- Conduct one-way sensitivity analyses using extreme values for all model parameters.
- If the parameter is based on an assumption or expert opinion, test it over a broad range in a one-way sensitivity analysis.
- Conduct a multi-way sensitivity analysis for model parameters that are shown to be sensitive in the one-way analyses.
- If two sources of error are interdependent, conduct a two-way sensitivity analysis on each source of error.
- Use a scenario analysis to analyze methodological uncertainty (e.g., run alternative versions of the model using different inputs).³⁰
- Use a sensitivity analysis to test uncertainty around structural assumptions (e.g., techniques for extrapolating data, conversion of intermediate outcomes to final outcomes).
- If an input contains subcomponents, disaggregate and test the subcomponents separately in the model.

The internal and external validity of the model should be formally validated.

Parameter uncertainty can be handled through a DSA or a PSA. There are advantages and disadvantages associated with both these methods of analysis. Analysts are encouraged to use a PSA. A PSA can provide a more complete assessment of the uncertainty associated with all inputs in a model. This type of analysis handles interactions between inputs, and provides interpretable, quantitative results. It is more likely to produce an unbiased estimate of mean costs and effects, and lead to optimal decisions given non-linear relationships (e.g., discounting). Furthermore, it allows for a Bayesian interpretation of results and can be used to determine where research funds should be allocated.

A PSA should take the form of a Monte Carlo simulation. The assumptions regarding the range of values for key parameters, the form of probability distributions, and the number of Monte

Carlo iterations should be stated and justified. The availability of simulation software (e.g., Crystal Ball[®], an Excel add-on) has made such methods easier to implement.

A PSA can only be used to assess parameter uncertainty, and should not be used to assess model uncertainty. Model structures, methods, and assumptions should be assessed through a DSA and model validation methods.

A probabilistic sensitivity analysis can be extended to Bayesian analysis. A Bayesian framework can be used to assess the value of gathering additional information, based on comparing the costs of conducting more research and the benefits from reducing uncertainty. The framework recognizes the binary nature of the decision facing a decision-maker (accept or reject), and quantifies all forms of parameter and model uncertainty in a PSA. It focuses on identifying parameters for which it is worth obtaining more sample information to reduce risk. This information can be used to prioritize research, or increase the efficiency of study design.⁸⁵

For information about Bayesian methods for cost-effectiveness analysis and computing, and the expected value of perfect information (EVPI), readers may refer to Heitjan and Li⁸⁶ and Claxton *et al.*⁸⁷ For an overview of the concept of “updating” prior evidence as new data become available, see Briggs.⁸⁸

3.13 Equity

Equity is a concern that decision makers must often address. The identity of the beneficiary (e.g., elderly, low income groups, geographically remote communities) may be an important factor in determining the social desirability of providing an intervention or program. A comprehensive discussion of equity is beyond the scope of the Economic Guidelines. Nonetheless, the importance of equity considerations in decision making for allocating health care resources warrants that some general points be made.

“Equity,” as it relates to health, can be defined as “fairness” in the allocation of resources, treatments, or outcomes among individuals or

groups.⁸⁹ Common to most definitions of health equity is the idea that certain health differences (or inequalities) are unfair, and that subsets of these inequalities that are judged to be unfair constitute health inequities. Some definitions also include the aspect of “reversibility” (i.e., whether the health inequality can be removed). Most equity considerations can be divided into those relating to need and those relating to access to services.¹⁰

The assessing of equity requires that comparisons be made between social groups with different levels of social advantage. Given that decision makers are concerned about the different aspects of equity, economic evaluations should provide as much information on these aspects as feasible. It is then the responsibility of the decision maker to provide the necessary weights or judgments to determine the redistribution of resources in various sectors of society, if the new intervention is implemented.⁹⁰

Equity considerations should be taken into account when conducting an economic evaluation. The analysis of distributional issues should be kept separate from the analysis of economic efficiency.

3.13.1 *Equity assumptions and implications*

Every type of economic evaluation contains equity assumptions and implications. As a result, decision makers who are using this information should understand them. All equity assumptions (whether implicit or explicit) should be highlighted in the report. If possible, state the implications of the assumptions on the results of the analysis.

The choice of outcome measure can have different equity implications, and the analysis should be sensitive to this.⁹¹ The use of a particular outcome measure (e.g., lives saved, life-years gained, or QALYs gained) typically implies that each unit of measurement is considered equal (i.e., equal weights for each unit), regardless of who gains. For instance, using life-years gained can favour the young, as they may have the capacity to gain more life-years from an intervention than the elderly. By using QALYs, it is assumed that a small gain to many people is equally desirable as a large gain to a few, so long as the QALY totals are the same (e.g., a gain of 0.1 QALY to each of 1,000 people would be considered equal to a gain of 25 QALYs

each to four individuals). Some suggest that measures such as “lives saved” or “person trade-off” are associated with the notion of equality in which each life is treated with the same sense of urgency.

Valuing outcomes using human capital and willingness-to-pay approaches can favour those with higher incomes, because this approach depends on the ability to pay and the existing income distribution. For example, if the gain in productive employment from a new ovarian cancer treatment is valued using only wage rates for females in the labour force, the value will be less than if average wage rates (which include those of males) were used.⁹² If the issue is of concern, more equitable estimates of wage rates can be used in a sensitivity analysis (e.g., general versus gender-specific wage rates).¹

There are other examples of equity implications.

- Using a higher discount rate reduces the present value of outcomes occurring in the future (e.g., prevention interventions that benefit the young).
- Clinical trial participants are often not representative of the general population, given the strict trial inclusion and exclusion criteria, or there may not be an equal opportunity to access the intervention when it is available in the “real world.”

3.13.2 *Groups that are affected*

The potential benefits, harms, and costs associated with a health technology are often unevenly distributed across the population. This may be due to differences in treatment effects, risks or incidence of conditions, access to health care, or technology uptake in population groups.

Economic evaluations should identify the subgroups that may be the primary beneficiaries if the intervention were provided. These subgroups can be defined in terms of predetermined categories that are considered to be relevant to equity, or in terms of unmet health needs. These categories can be defined by age group (e.g., elderly or young), sex, ethnic group (e.g., aboriginal groups), geographical area (e.g., rural or northern communities), socioeconomic group (e.g., low income), or condition severity and health

status. It follows that certain technologies may be performed more often for one compared to another.

Groups that are likely to be disadvantaged by the availability of the intervention should also be identified. This may occur, for example, when a change in clinical practice requires that patients be cared for at home, rather than at hospital, thereby shifting costs and burdens to patients and informal caregivers.

3.13.3 *Distributional impact*

Analysts are encouraged to provide information on how the distributional impact (e.g., benefits, harms, and costs) and cost-effectiveness results of the intervention vary across predetermined, equity-relevant subgroups. The magnitude of these effects can be reported (i.e., whether for a fixed funding level, there is a large health gain for a small number of individuals, or a small health gain for a large number of individuals).

Such information can be presented in a chart or matrix that depicts the relevant groups on one axis, and the associated distributional impacts and cost-effectiveness results on the other axis. If the distributional impact of an intervention is small, it can be mentioned in the report without the need for an elaborate presentation of information. If a large number of groups are affected, consider limiting the details in the chart to ensure that the data are clear and understandable for decision makers.⁹³

When the intervention can be provided selectively to certain subgroups, then cost-effectiveness information should be presented for each subgroup. A stratified analysis can be useful for estimating the (opportunity) cost of providing the technology to subgroups on the basis of equity. This information allows decision makers to focus on a more explicit value judgment regarding the efficiency-equity trade-off. Can the opportunity cost of equitable access be justified? Is the benefit gained from stratification worth any inequities in provision? Optimal cohorts for an intervention could be based on the incremental cost per QALY, although the net benefits framework permits a more explicit quantification of the efficiency gains obtained from stratification, and the opportunity cost of incorporating equity concerns.²⁵

Equity weights should not be applied to outcomes in the Reference Case. This implies that health gains are weighted equally regardless to whom they accrue (e.g., one QALY is valued as one QALY regardless of who experiences the gains). The analysis and results should be presented in a disaggregated and transparent manner to allow decision makers the ability to evaluate how outcomes were combined or converted, assess distributional impacts, and undertake any additional analysis, if desired. To avoid confusion, the analysis of the distributional issues should be kept separate from the economic-efficiency analysis.

3.14 Generalizability

Generalizability refers to “the problem of whether one can apply or extrapolate results obtained in one setting or population to another.”⁹⁴ The term “generalizability” may also be called transferability, transportability, external validity, relevance, or applicability. Generalizability raises issues regarding the conduct of an economic evaluation and the interpretation of its findings. Two useful publications on topic are those by Sculpher *et. al.*⁹⁵ and Welte *et. al.*⁹⁶.

The key question about the generalizability of results is whether there are differences in an intervention’s impact on effectiveness and costs across settings or locations that produce meaningful differences in cost-effectiveness.⁹⁴ In-country differences can be relevant in Canada, given that regional differences may exist, and the likelihood that an evaluation will be used to inform decision making in more than one jurisdiction. Similar issues can arise regarding the generalizability of results across different populations (e.g., men versus women, young versus elderly) in the same jurisdiction.

Three aspects of generalizability may be an issue in economic evaluations, and warrant the attention of the analyst:

- distinction between the efficacy and the effectiveness of the intervention
- handling of data on costs and health state preferences derived from another setting
- handling of data from trials that are undertaken on a multinational basis.

3.14.1 Effectiveness

The main issue is whether the efficacy data obtained from a controlled (trial) setting reflect the effectiveness that might be achieved in a broader “real world” setting (e.g., routine clinical practice in Canada). This is most relevant to the decision maker (i.e., the external validity of the clinical trial).

It can be argued that clinical relevance should be the first criterion that is addressed in determining study generalizability. If one cannot verify that relevant clinical evidence exists, then there may be little benefit (and substantial work) involved in proceeding with the verification of other clinical and cost data to complete the transfer.

3.14.2 Economic data

The second aspect of generalizability is the handling of economic data that have been collected in a trial conducted in a country other than Canada. Cost data vary from country to country, reflecting differences in resources use patterns and relative unit cost levels. An example of this is the variation in length of stay for surgical procedures that vary from place to place. Preferences for health states often depend on cultural factors that vary among countries. It follows that cost and preference (utility) data from other countries may not be generalizable to Canada.

In light of this, analysts may seek to adjust the data from outside Canada to apply to the Canadian setting. Methods for doing so differ, but the principle is that the adaptation of cost data is more complex than simply changing the price weights (e.g., converting from US to Canadian dollars). It may be appropriate to use modelling techniques to adjust observed data about practice patterns from other countries or settings to apply to local circumstances and then use local unit costs in the model. Attempts have been made to adapt observed practice patterns to local circumstances using modelling techniques.^{97,98} Where the evaluation relies on data based on practice patterns from outside Canada, information on how the condition is defined and treated (e.g., dosage regimens, hospitalization rates, length of stay) should be reported, so that jurisdiction-specific information can be identified and used in the evaluation. Where clinical practice has been

adjusted in the analysis, the methods for doing so should be transparent.

3.14.3 Multinational trials

The third aspect of generalizability relates to the selection of appropriate methods to analyze economic and clinical data that have been collected in prospective trials involving several countries. A central issue is whether to pool data collected from all countries or to use data from the centres or countries that are most applicable to the decision maker’s setting. The appropriate approach will vary from case to case, and will depend on the statistical evidence and informed judgment.

It is often assumed that clinical data can be pooled, although this may not be the case for all interventions, and tests of homogeneity should be performed to confirm this. The choice may depend on the selection of data from those trial centres using a comparator that is relevant to the decision maker’s setting. In contrast, it is generally assumed that economic data will differ systematically between multinational centres, and therefore pooling will be impossible. At issue then is the adequacy of the sample size for centre- or country-specific ICER estimates.

If the analyst uses cost data that are obtained from a multinational trial, the differences between centres or countries in terms of quantities of resources used and unit prices, and the method used to estimate the overall cost-effectiveness should be reported. Report the source, place, and year of data for resource quantities and unit costs, and rates used for inflation and currency conversion. The appropriate ranges of parameters should be tested through a sensitivity analysis.⁹⁹

Methods are being developed to address the issues of handling economic data from multinational trials, including the use of multi-level modelling,¹⁰⁰ empirical Bayesian methods,¹⁰¹ multivariate regression analysis,¹⁰² and net benefit regression analysis.¹⁰³

3.14.4 Sensitivity analysis

It is unfeasible to perform primary economic evaluations that are tailored to the specific circumstances of each jurisdiction that uses the evaluation for decision making. Nonetheless,

incorporating local parameter values into a sensitivity analysis will permit the interpretation of evaluation findings for local application. If there is regional (i.e., in country) variation in clinical practice or other model parameters, the Reference Case analysis can be performed at a national (or aggregate) level using the most widespread or best available practice or data, with a sensitivity analysis performed using regional (or local) practice or data.

3.14.5 *Transparency*

Present the analysis in a transparent manner, to help decision makers judge the generalizability of the findings to their setting. Describe the intervention and alternatives, ensure the methods of analysis are easy to follow, and present the cost and outcomes in disaggregated detail before aggregation. Physical quantities (e.g., length of stay in hospital) and unit costs should be reported separately rather than reporting total costs only. This information will allow decision makers to make a more informed judgment about how applicable data from other countries are to their jurisdiction.

3.14.6 *Discussion in report*

The report should include a discussion on the relevance of the data and model to the jurisdictions and populations of interest, and the generalizability of the overall results (see Appendix 3). If economic data from outside Canada have been used, there should be discussion about the validity of such data for the target audience. Regional or setting differences for the target audience in terms of disease epidemiology, population characteristics, effectiveness of the intervention, clinical practice patterns, resource use patterns, unit costs, and other relevant factors should be discussed. If differences exist, discuss the impact on the results (i.e., direction and expected magnitude) and conclusions.

Because it is unfeasible to tailor evaluations to the specific circumstances of every jurisdiction using the information, decision makers may request a copy of the model to conduct their own analysis. This would allow them to input their data, make any necessary model adjustments, and generate results that are more relevant to their setting.

3.15 Reporting

3.15.1 *General considerations*

The reports of economic evaluations should be clear and detailed, and the analysis should be presented in a transparent manner. The report should provide enough information to enable the audience to critically evaluate the validity of the analysis, including information on each element of the economic evaluation, as outlined in the Economic Guidelines, and estimates of costs and outcomes. The report should be written in a manner that is easily understood by the target audience. Wherever possible, use plain language, and define jargon or technical terms that may be unfamiliar to the target audience.

The report format should be well structured and easy to follow. To enhance clarity and facilitate the comparison of economic evaluations, analysts can use the structured report format in Appendix 3.

An executive summary should be included at the beginning of the report. The executive summary should be no longer than two pages, and written in language that is understood by a non-technical reader.

The report should also address how each element of the economic evaluation, as outlined in the Economic Guidelines, has been handled.

3.15.2 *Presentation of results*

All results should first be reported in the detailed steps of the analysis, with aggregations and use of value judgments (e.g., preference scores) introduced into the presentation of information as late as possible. A stepped approach is useful for presenting the perspectives and type of evaluations that may have been used in the analysis (e.g., presenting the results in terms of a CCA, then a CEA, and finally a CUA and a CBA, as relevant). Each aggregation step should be presented in enough detail to allow for independent verification of the results.

Intermediate results for each alternative should be disaggregated into undiscounted totals of costs and outcomes before aggregation and discounting. Totals should be shown in gross and net terms. The estimates of the individual components of total

cost should be presented. The total costs may not represent “true” totals, because costs that are common to the alternatives may have been disregarded in the analysis. Nonetheless, the totals enable decision makers to better appreciate the magnitudes involved, which can be masked by a ratio of two increments. The totals also allow future users to make comparisons with new or different comparators, to bring studies up to date, or to transfer study results across geographic or practice boundaries.

Final results should be based on increments (i.e., differences between the intervention and alternatives) of expected costs and expected outcomes. The results for a CEA or CUA should be reported as ICERs. The results should not be reported as average cost-effectiveness ratios (i.e., total costs divided by total outcomes) of the alternatives, as this can mislead decision makers.¹⁰⁴ The net benefit measure may be used as an additional (but not alternative) measure to the ICER. The willingness-to-pay threshold and the associated ICER should be stated for each net benefit estimate.

When there are more than two alternatives being compared, the expected costs and outcomes of the alternatives can be reported in a table in increasing order of cost, starting with the lowest cost alternative at the top. The incremental costs, incremental outcomes, and ICERs of the alternatives can then be calculated sequentially. Any alternatives that are ruled out by strong (or strict) dominance (i.e., they are more costly and less effective than another alternative), are noted in the table and disregarded in the calculation of the initial ICERs. After the initial ICERs have been calculated, any alternative that is ruled out by weak (or extended) dominance (i.e., when an alternative is more costly and more effective, and has a lower ICER), are noted in the table, and are disregarded in calculation of the final ICERs.⁶⁸

To facilitate understanding, analysts are encouraged to present the results of the analysis in graphical or visual, and tabular, forms. Appendix 1 provides suggestions for the graphic presentation of results for variability and for the uncertainty

analysis. All graphics should be appropriately discussed, and not used to replace a written description or interpretation of results.

3.15.3 *Disclosure of relationships*

Funding and reporting arrangements should be stated in the report, or in a letter of authorship accompanying the report. Disclosure should include a list of all key participants in the study with their contributions, and the sponsor of the study. It should also indicate whether the sponsor had any review or editing rights regarding the analysis plan and report.

Declarations of any conflicts of interest by the authors, or a declaration that no conflict exists, should accompany the report. Conflicts of interest can be considered to be a financial or a non-financial interest. Guidelines for declaration of conflicts of interest and a declaration template can be found in the Guidelines for Authors of CADTH Health Technology Reports.¹⁰⁵

3.15.4 *Quality assurance*

Analysts should assure the users of economic evaluations about the quality of the process underlying the study. This can be achieved by a thorough delineation of the conduct of the study, including how it was documented to ensure consistency and quality in the process. Documents specific to the quality assurance process should be made available to users.

If requested, documentation describing the model and the model validation process in detail should be made available to decision makers. A description of the statistical analysis (i.e., data sources, methods, and results) should be made available if used in the economic evaluation. An operable copy of the model with an adequate user interface should be made available to decision makers upon request (under conditions of strict confidentiality and protection of property rights) for review, and to permit a sensitivity analysis to be undertaken using the decision maker’s data and assumptions.

4 REFERENCES

1. Drummond MF, O'Brien BJ, Stoddart GL, Torrance GW. *Methods for the economic evaluation of health care programmes* [Oxford medical publications]. 2nd ed. New York: Oxford University Press; 1997.
2. Contandriopoulos AP, Champagne F, Denis JL, Avargues MC. L'évaluation dans le domaine de la santé: concepts et méthodes [Evaluation in the health sector: concepts and methods]. *Rev Epidemiol Sante Publique* 2000;48(6):517-39.
3. Getzen TE. *Health economics: fundamentals and flow of funds*. New York: John Wiley and Sons; 1997.
4. Gold MR, Siegel JE, Russell LB, Weinstein MC. *Cost-effectiveness in health and medicine*. New York: Oxford University Press; 1996.
5. Incorporating health economics in guidelines and assessing resource impact. In: *The guideline developments methods: information for National Collaborating Centres and guideline developers*. Rev.ed. London: National Institute for Clinical Excellence; 2005. Available: http://www.nice.org.uk/pdf/GDM_Chapter8_0305.pdf (accessed 2005 Sep 29).
6. Donaldson C, Mooney G. Needs assessment, priority setting, and contracts for health care: an economic view. *BMJ* 1991;303(6816):1529-30.
7. Mitton C, Donaldson C. *Priority setting toolkit: a guide to the use of economics in healthcare decision making*. London: BMJ; 2004.
8. Drummond M, Brown R, Fendrick AM, Fullerton P, Neumann P, Taylor R, et al. Use of pharmacoeconomics information--report of the ISPOR Task Force on use of pharmacoeconomic/health economic information in health-care decision making. *Value Health* 2003;6(4):407-16.
9. McDaid D, Mossialos E, Mrazek MF. Making use of economic evaluation. *Int J Risk Saf Med* 2002;15:67-74.
10. Hauck K, Smith PC, Goddard M. *The economics of priority setting for health care: a literature review* [HNC discussion paper]. Washington: The International Bank for Reconstruction and Development/The World Bank; 2003. Available: http://www-wds.worldbank.org/servlet/WDSContentServer/WDSP/IB/2004/05/18/000265513_20040518173341/Rendered/PDF/288780Hauck1Th1cs0of0PriorityIwhole.pdf.
11. 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(16):2116-21.
12. Hoffmann C, Graf von der Schulenburg JM, EUROMET Group. The influence of economic evaluation studies on decision making. A European survey. *Health Policy* 2000;52(3):179-92.
13. Hjelmgren J, Berggren F, Andersson F. Health economic guidelines--similarities, differences and some implications. *Value Health* 2001;4(3):225-50.
14. Drummond MF, McGuire A, editors. *Economic evaluation in health care : merging theory with practice*. New York: Oxford University Press; 2001.
15. Muennig P. *Designing and conducting cost-effectiveness analyses in medicine and health care*. San Francisco: Jossey-Bass; 2002.

16. *Common drug review* [database online]. Ottawa: Canadian Coordinating Office for Health Technology Assessment; 2004. Available: http://www.ccohta.ca/entry_e.html (accessed 2004 Oct 29).
17. Garber AM, Phelps CE. Economic foundations of cost-effectiveness analysis. *J Health Econ* 1997;16(1):1-31.
18. Bucher HC, Guyatt G, Cook DJ, Holbrook A, McAllister F, Evidence-based Medicine Working Group. *Applying clinical trial results: part A, how to use an article measuring the effect of an intervention on surrogate end points*. Edmonton: Centre for Health Evidence; 2005. Available: <http://www.cche.net/text/usersguides/surrogate.asp> (accessed 20050822).
19. Prentice RL. Surrogate endpoints in clinical trials: definition and operational criteria. *Stat Med* 1989;8(4):431-40.
20. Nord E. Severity of illness versus expected benefit in societal evaluation of healthcare interventions. *Expert Rev Pharmacoecon Outcomes Res* 2001;1(1):85-92.
21. Briggs AH, O'Brien BJ. The death of cost-minimization analysis? *Health Econ* 2001;10(2):179-84.
22. Cookson R. Willingness to pay methods in health care: a sceptical view. *Health Econ* 2003;12(11):891-4.
23. Ryan M. Using conjoint analysis to take account of patient preferences and go beyond health outcomes: an application to in vitro fertilisation. *Soc Sci Med* 1999;48(4):535-46.
24. O'Brien B, Viramontes JL. Willingness to pay: a valid and reliable measure of health state preference? *Med Decis Making* 1994;14(3):289-97.
25. Coyle D, Buxton MJ, O'Brien BJ. Stratified cost-effectiveness analysis: a framework for establishing efficient limited use criteria. *Health Econ* 2003;12(5):421-7.
26. Husereau DR, Brady B, McGeer A. *Oseltamivir for the treatment of suspected influenza: a clinical and economic assessment*. Ottawa: Canadian Coordinating Office for Health Technology Assessment; 2001.
27. O'Brien BJ, Goeree R, Bernard L, Rosner A, Williamson T. Cost-effectiveness of tolterodine for patients with urge incontinence who discontinue initial therapy with oxybutynin: a Canadian perspective. *Clin Ther* 2001;23(12):2038-49.
28. Noorani HZ, Brady B, McGahan L, Teasell R, Skidmore B, Doherty TJ. *Stroke rehabilitation services: systematic reviews of the clinical and economic evidence* [Technology report no 35]. Ottawa: Canadian Coordinating Office for Health Technology Assessment; 2003.
29. Shih YC, Han S, Cantor SB. Impact of generic drug entry on cost-effectiveness analysis. *Med Decis Making* 2005;25(1):71-80.
30. Philips Z, Ginnelly L, Sculpher M, Claxton K, Golder S, Riemsma R, et al. Review of guidelines for good practice in decision-analytic modelling in health technology assessment. *Health Technol Assess* 2004;8(36):1-172.
31. Rawson NS. "Effectiveness" in the evaluation of new drugs: A misunderstood concept? *Can J Clin Pharmacol* 2001;8(2):61-2.
32. Boutron I, Tubach F, Giraudeau B, Ravaud P. Methodological differences in clinical trials evaluating nonpharmacological and pharmacological treatments of hip and knee osteoarthritis. *JAMA* 2003;290(8):1062-70.
33. Weinstein MC, O'Brien B, Hornberger J, Jackson J, Johannesson M, McCabe C, et al. Principles of good practice for decision analytic modeling in health-care evaluation: report of the ISPOR Task Force on Good Research Practices--Modeling Studies. *Value Health* 2003;6(1):9-17.

34. Wahl C, Gregoire JP, Teo K, Beaulieu M, Labelle S, Leduc B, et al. Concordance, compliance and adherence in healthcare: closing gaps and improving outcomes. *Healthc Q* 2005;8(1):65-70.
35. Decision analytic modelling in the economic evaluation of health technologies. A consensus statement. Consensus Conference on Guidelines on Economic Modelling in Health Technology Assessment. *Pharmacoeconomics* 2000;17(5):443-4.
36. Grenier W, Schoffski O, Schulenburg JM. The transferability of international economic health-economic results to national study questions. *Eur J Health Econ* 2000;1:94-102.
37. Brennan A, Akehurst R. Modelling in health economic evaluation. What is its place? What is its value? *Pharmacoeconomics* 2000;17(5):445-59.
38. Jordan HS, Lau J. Linking pharmacoeconomic analyses to results of systematic review and meta-analysis. *Expert Rev Pharmacoecon Outcomes Res* 2004;3(4):441-8.
39. Coyle D, Lee KM. Evidence-based economic evaluation: how the use of different data sources can impact results. In: Donaldson C, Mugford M, Vale L, editors. *Evidence-based health economics: from effectiveness to efficiency*. London: BMJ; 2002. p.55-66.
40. Lohr KN. Rating the strength of scientific evidence: relevance for quality improvement programs. *Int J Qual Health Care* 2004;16(1):9-18.
41. Briggs AH. Handling uncertainty in economic evaluation and presenting the results. In: Drummond MF, McGuire A, editors. *Economic evaluation in health care: merging theory with practice*. New York: Oxford University Press; 2001. p.172-214.
42. Society of Medical Decision Making. *Glossary of medical decision making terms*. Philadelphia: The Society; 2005. Available: <http://www.smdm.org/Glossary.htm> (accessed 2005 Mar 8).
43. Guyatt GH, Feeny DH, Patrick DL. Measuring health-related quality of life. *Ann Intern Med* 1993;118(8):622-9.
44. Guyatt GH, Veldhuyzen Van Zanten SJ, Feeny DH, Patrick DL. Measuring quality of life in clinical trials: a taxonomy and review. *CMAJ* 1989;140(12):1441-8.
45. Feeny DH, Torrance GW, Labelle R. Integrating economic evaluations and quality of life assessments. In: Spilker B, editor. *Quality of life and pharmacoeconomics in clinical trials*. 2nd ed. Philadelphia: Lippincott-Raven; 1996. p.85-95.
46. Weinstein MC. Principles of cost-effective resource allocation in health care organizations. *Int J Technol Assess Health Care* 1990;6(1):93-103.
47. Gold MR, Patrick DL, Torrance GW, Fryback DG, Hadorn DC, Kamlet MS, et al. Identifying and valuing outcomes. In: Gold MR, Russell LB, Seigel JE, Weinstein MC, editors. *Cost-effectiveness in health and medicine*. New York: Oxford University Press; 1996. p.82-134.
48. Bennett KJ, Torrance GW. Measuring health state preferences and utilities: rating scale, time trade-off and standard gamble techniques. In: Spilker B, editor. *Quality of life and pharmacoeconomics in clinical trials*. 2nd ed. Philadelphia: Lippincott-Raven; 1996. p.253-65.
49. Feeny DH, Torrance GW, Furlong WJ. Health utilities index. In: Spilker B, editor. *Quality of life and pharmacoeconomics in clinical trials*. 2nd ed. Philadelphia: Lippincott-Raven; 1996. p.239-52.
50. Torrance GW, Furlong W, Feeny D. Health utility estimation. *Expert Rev Pharmacoecon Outcomes Res* 2002;2:99-108.

51. Von Neumann J, Morgenstern O. *Theory of games and economic behaviour*. Princeton (NJ): Princeton University Press; 1944.
52. Richardson J. Cost utility analysis: What should be measured? *Soc Sci Med* 1994;39(1):7-21.
53. Feeny D, Furlong W, Torrance GW, Goldsmith CH, Zhu Z, DePauw S, et al. Multiattribute and single-attribute utility functions for the health utilities index mark 3 system. *Med Care* 2002;40(2):113-28.
54. Furlong WJ, Feeny DH, Torrance GW, Barr RD. The Health Utilities Index (HUI) system for assessing health-related quality of life in clinical studies. *Ann Med* 2001;33(5):375-84.
55. Horsman J, Furlong W, Feeny D, Torrance G. The Health Utilities Index (HUI(R)): concepts, measurement properties and applications. *Health Qual Life Outcomes* 2003;1(1):54.
56. EuroQol Group. EuroQol--a new facility for the measurement of health-related quality of life. *Health Policy* 1990;16(3):199-208.
57. Essink-Bot ML, Stouthard ME, Bonsel GJ. Generalizability of valuations on health states collected with the EuroQolc-questionnaire. *Health Econ* 1993;2(3):237-46.
58. Brooks R. EuroQol: the current state of play. *Health Policy* 1996;37(1):53-72.
59. Kind P. The EuroQol instrument: an index of health-related quality of life. In: Spilker B, editor. *Quality of life and pharmacoeconomics in clinical trials*. 2nd ed. Philadelphia: Lippincott-Raven; 1996. p.191-201.
60. Brazier J, Roberts J, Tsuchiya A, Busschbach J. A comparison of the EQ-5D and SF-6D across seven patient groups. *Health Econ* 2004;13(9):873-84.
61. Sintonen H. The 15D instrument of health-related quality of life: properties and applications. *Ann Med* 2001;33(5):328-36.
62. Torrance GW, Feeny DH, Furlong WJ, Barr RD, Zhang Y, Wang Q. Multiattribute utility function for a comprehensive health status classification system. Health Utilities Index Mark 2. *Med Care* 1996;34(7):702-22.
63. Brazier J, Deverill M, Green C, Harper R, Booth A. A review of the use of health status measures in economic evaluation. *Health Technol Assess* 1999;3(9):1-164.
64. Gafni A. Willingness-to-pay as a measure of benefits. Relevant questions in the context of public decisionmaking about health care programs. *Med Care* 1991;29(12):1246-52.
65. Olsen JA, Donaldson C. Helicopters, hearts and hips: using willingness to pay to set priorities for public sector health care programmes. *Soc Sci Med* 1998;46(1):1-12.
66. Ryan M. A role for conjoint analysis in technology assessment in health care? *Int J Technol Assess Health Care* 1999;15(3):443-57.
67. Ratcliffe J, Buxton M. Patients' preferences regarding the process and outcomes of life-saving technology. An application of conjoint analysis to liver transplantation. *Int J Technol Assess Health Care* 1999;15(2):340-51.
68. Berger ML, International Society for Pharmacoeconomics and Outcomes Research. *Health care cost, quality, and outcomes : ISPOR book of terms*. Lawrenceville (NJ): ISPOR; 2003.
69. Brouwer W, Rutten F, Koopmanschap M. Costing in economic evaluations. In: Drummond M, McGuire A, editors. *Economic evaluation in health care: merging theory with practice*. Oxford: Oxford University Press; 2001. p.68-93.
70. Baladi JF. *A guidance document for the costing process*. Version 1.0. Ottawa: Canadian Coordinating Office for Health Technology Assessment; 1996.
71. Oostenbrink JB, Koopmanschap MA, Rutten FF. Standardisation of costs: the Dutch

- manual for costing in economic evaluations. *PharmacoEconomics* 2002;20(7):443-54.
72. Adam T, Koopmanschap MA, Evans DB. Cost-effectiveness analysis: can we reduce variability in costing methods? *Int J Technol Assess Health Care* 2003;19(2):407-20.
 73. Jacobs P, Ohinmaa A, Brady B. Providing systematic guidance in pharmaco-economic guidelines for analysing costs. *PharmacoEconomics* 2005;23(2):143-53.
 74. Jacobs P. *A guide to the 2002 national list of provincial costs for Canada*. Edmonton: Institute of Health Economics; 2004. Available: <http://www.ihe.ca/publications/phcc/index.cfm> (accessed 2004 Dec 1).
 75. Liljas B. How to calculate indirect costs in economic evaluations. In: Mallarkey G, editor. *Economic evaluation in healthcare*. Auckland: Adis; 1999. p.71-7.
 76. Pritchard C, Sculpher M. *Productivity costs: principles and practice in economic evaluation*. London: Office of Health Economics; 2000.
 77. Sculpher M. The role and estimation of productivity costs in economic evaluation. In: Drummond M, McGuire A, editors. *Economic evaluation in health care*. New York: Oxford University Press; 2001. p.94-112.
 78. Goeree R, O'Brien BJ, Blackhouse G, Agro K, Goering P. The valuation of productivity costs due to premature mortality: a comparison of the human-capital and friction-cost methods for schizophrenia. *Can J Psychiatry* 1999;44(5):455-63.
 79. van den BB, Brouwer WB, Koopmanschap MA. Economic valuation of informal care. An overview of methods and applications. *Eur J Health Econ* 2004;5(1):36-45.
 80. Johannesson M. Avoiding double-counting in pharmaco-economic studies. In: Mallarkey G, editor. *Economic evaluation in healthcare*. Auckland: Adis; 1999. p.155-7.
 81. Sendi P, Brouwer WBF. Leisure time in economic evaluation: theoretical and practical considerations. *Expert Rev Pharmacoecon Outcomes Res* 2004;4(1):1-3.
 82. West RR, McNabb R, Thompson AG, Sheldon TA, Grimley EJ. Estimating implied rates of discount in healthcare decision-making. *Health Technol Assess* 2003;7(38):1-60.
 83. Smith DH, Gravelle H. The practice of discounting in economic evaluations of healthcare interventions. *Int J Technol Assess Health Care* 2001;17(2):236-43.
 84. Lazaro A. Theoretical arguments for the discounting of health consequences: where do we go from here? *PharmacoEconomics* 2002;20(14):943-61.
 85. Coyle D, Buxton MJ, O'Brien BJ. Measures of importance for economic analysis based on decision modeling. *J Clin Epidemiol* 2003;56(10):989-97.
 86. Heitjan DF, Li H. Bayesian estimation of cost-effectiveness: an importance-sampling approach. *Health Econ* 2004;13(2):191-8.
 87. Claxton K, Sculpher M, Drummond M. A rational framework for decision making by the National Institute For Clinical Excellence (NICE). *Lancet* 2002;360(9334):711-5.
 88. Briggs AH. A Bayesian approach to stochastic cost-effectiveness analysis. *Health Econ* 1999;8(3):257-61.
 89. The World Bank Group. *Health economics equity*. Washington: The Group; 2001. Available: http://www1.worldbank.org/hnp/hsd/healthcon_equity.asp (accessed 2004 Jul 13).
 90. Baltussen R, Leidl R, Ament A. The impact of age on cost-effectiveness ratios and its control in decision making. *Health Econ* 1996;5(3):227-39.
 91. Gafni A, Birch S. Equity considerations in utility-based measures of health outcomes in

- economic appraisals: an adjustment algorithm. *J Health Econ* 1991;10(3):329-42.
92. Donaldson C, Birch S, Gafni A. The distribution problem in economic evaluation: income and the valuation of costs and consequences of health care programmes. *Health Econ* 2002;11(1):55-70.
93. Treasury Board of Canada. *Benefit cost analysis guide: draft july 1998*. Ottawa: The Board; 1998. Available: http://www.tbs-sct.gc.ca/fin/sigs/revolving_funds/bcag/bca2_e.asp?printable=True (accessed 2004 Aug 14).
94. Willke RJ. Tailor-made or off-the-rack? The problem of transferability of health economic data. *Expert Rev Pharmacoecon Outcomes Res* 2003;3(1):1-4.
95. Sculpher MJ, Pang FS, Manca A, Drummond MF, Golder S, Urdahl H, et al. Generalisability in economic evaluation studies in healthcare: a review and case studies. *Health Technol Assess* 2004;8(49):1-192. Available: <http://www.ncchta.org/fullmono/mon849.pdf>.
96. Welte R, Feenstra T, Jager H, Leidl R. A decision chart for assessing and improving the transferability of economic evaluation results between countries. *Pharmacoeconomics* 2004;22(13):857-76.
97. Schulman K, Burke J, Drummond M, Davies L, Carlsson P, Gruger J, et al. Resource costing for multinational neurologic clinical trials: methods and results. *Health Econ* 1998;7(7):629-38.
98. Drummond M, Pang F. Transferability of economic evaluation results. In: Drummond M, McGuire A, editors. *Economic evaluation in health care: merging theory with practice*. New York: Oxford University Press; 2001. p.256-76.
99. Halliday RG, Darba J. Cost data assessment in multinational economic evaluations: some theory and review of published studies. *Appl Health Econ Health Policy* 2003;2(3):149-55.
100. Rice N, Jones A. Multilevel models and health economics. *Health Econ* 1997;6(6):561-75.
101. Gould AL. Multi-centre trial analysis revisited. *Stat Med* 1998;17(15-16):1779-97.
102. Willke RJ, Glick HA, Polsky D, Schulman K. Estimating country-specific cost-effectiveness from multinational clinical trials. *Health Econ* 1998;7(6):481-93.
103. Hoch JS, Briggs AH, Willan AR. Something old, something new, something borrowed, something blue: a framework for the marriage of health econometrics and cost-effectiveness analysis. *Health Econ* 2002;11(5):415-30.
104. Leach J. Children undergoing treatment with botulinum toxin: the role of the physical therapist. *Muscle Nerve Suppl* 1997;6:S194-S207.
105. *Guidelines for authors of CCOHTA health technology assessment reports*. Rev.ed. Ottawa: Canadian Coordinating Office for Health Technology Assessment; 2003. Available: http://www.ccohta.ca/misc/authors_guidelines_files/AuthorsGuidelines.doc.
106. Siebert U, Sroczynski G, Rossol S, Wasem J, Ravens-Sieberer U, Kurth BM, et al. Cost effectiveness of peginterferon alpha-2b plus ribavirin versus interferon alpha-2b plus ribavirin for initial treatment of chronic hepatitis C. *Gut* 2003;52(3):425-32.
107. Black WC. The CE plane: a graphic representation of cost-effectiveness. *Med Decis Making* 1990;10(3):212-4.
108. Glick HA, Briggs A, Polsky D. Quantifying stochastic uncertainty and presenting results of cost-effectiveness analyses. *Expert Rev Pharmacoecon Outcomes Res* 2001;1(1):25-36.
109. Mandelblatt JS, Lawrence WF, Womack SM, Jacobson D, Yi B, Hwang YT, et al.

- Benefits and costs of using HPV testing to screen for cervical cancer. *JAMA* 2002;287(18):2372-81.
110. Mark DH. Visualizing cost-effectiveness analysis. *JAMA* 2002;287(18):2428-9.
111. Khan KS, ter Rief G, Glanville J, Sowden AJ, Kleijnen J, editors. *Undertaking systematic reviews of research on effectiveness: CRD's guidance for those carrying out or commissioning reviews*. 2nd ed. York (UK): NHS Centre for Reviews and Dissemination, University of York; 2001. CRD report no 4. Available: http://www.york.ac.uk/inst/crd/index_guidance.htm
112. Khan KS, Kunz R, Kleijnen J, Antes G. *Systematic reviews to support evidence-based medicine: how to review and apply findings of healthcare research*. Lake Forest (IL): Royal Society of Medicine Press; 2003.
113. Napper M, Varney J. Health economics information. In: *Etext on health technology assessment (HTA) resources*. Bethesda (MD): National Information Center on Health Services Research & Health Care Technology, National Library of Medicine; 2004. Available: <http://www.nlm.nih.gov/nichsr/ehta/chapter11.html> (accessed 2004 Dec 12).
114. Moher D, Cook DJ, Eastwood S, Olkin I, Rennie D, Stroup DF. Improving the quality of reports of meta-analyses of randomised controlled trials: the QUOROM statement. Quality of reporting of meta-analyses. *Lancet* 1999;354(9193):1896-900.
115. Drummond MF, Jefferson TO, BMJ Economic Evaluation Working Party. Guidelines for authors and peer reviewers of economic submissions to the BMJ. *BMJ* 1996;313(7052):275-83. Available: <http://www.bmj.com/cgi/content/full/313/7052/275>.
116. 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;21(2):240-5.

APPENDICES

APPENDIX 1: Presenting Results of the Analysis

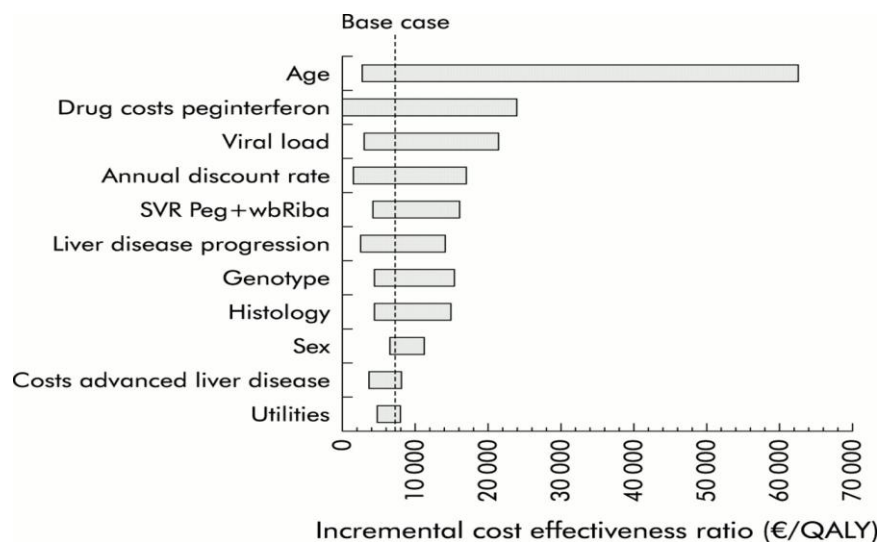
Suggestions are provided here for presenting the results of the analysis of variability and uncertainty in an evaluation. Sensitivity analysis results can be shown in tabular form, with results grouped into categories such as health outcome, cost, and modelling assumptions. The results can be reported as actual incremental cost-effectiveness ratio (ICER) figures, or as the percentage change from the Reference Case.

Wherever possible, analysts are encouraged to use a graphical or visual presentation of results, to aid understanding by the audience. This can be done using a variety of approaches, depending on the nature of the analysis including tornado diagrams, scatter plots, or cost-effectiveness acceptability curves (CEAC).

Tornado diagram

A tornado diagram is a useful way of displaying the results of the subgroups and one-way sensitivity analysis in one graph. Figure 1 is an example from an economic evaluation of an antiviral treatment for hepatitis C infection.¹⁰⁶ The horizontal axis depicts the ICER results, and the vertical axis shows the subgroups and parameters analyzed. The dotted line represents the result for the base case, and the bars represent the results for subgroups or parameters tested over the full range of values in a one-way sensitivity analysis. The ends of the bars can be labelled with the upper and lower values of the ranges tested for each parameter. Bars are ordered from widest to narrowest starting at the top of the figure. In this example, parameters with the greatest impact on the results are age, drug costs peginterferon, and viral load. Bars that reach the vertical axis (i.e., drug costs peginterferon) indicate where the intervention is cost saving.

Figure 1: Tornado Diagram



Source: Gut, 2003, vol. 52, issue 3, pp 425-432;¹⁰⁶ reproduced with permission from the BMJ Publishing Group

Cost-effectiveness plane

An alternative graphical presentation of results has come from Black.¹⁰⁷ An example is shown in a paper by Glick *et al.*¹⁰⁸ In this example, incremental costs and effects are displayed on the cost-effectiveness plane. The incremental effects and incremental costs of the intervention are plotted in the four quadrants denoted by the X and Y axes, respectively, with the comparator forming the origin (Figure 2). The slope of the line joining the origin to any cost-effect combination is the ICER result. The results of cases that are more effective and more costly than the comparator are located in the northeast quadrant, and those that are less effective and less costly than the comparator appear in the southwest quadrant. ICER results cannot be calculated for cases that are located in the northwest quadrant (i.e., where the intervention is dominated by the comparator), and

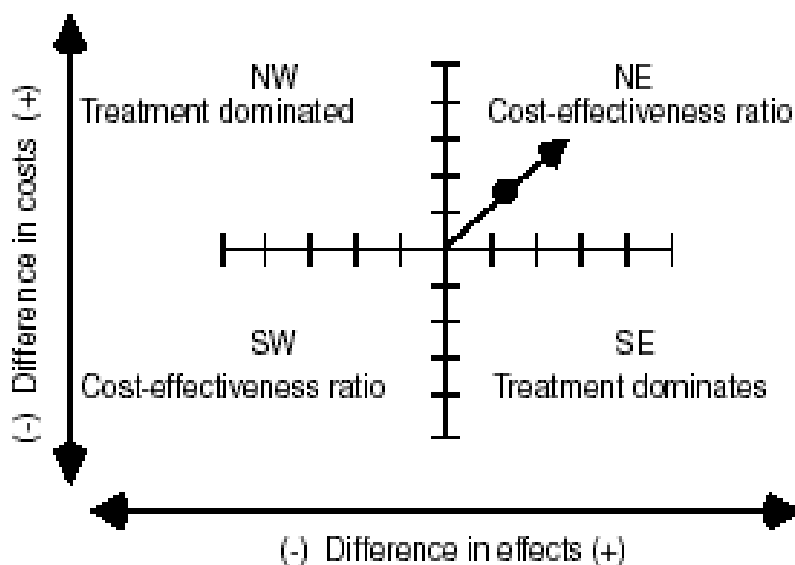
the southeast quadrant (i.e., where the intervention dominates the comparator).

The cost-effectiveness plane can be used to simultaneously compare (visually) the joint distribution of incremental costs and incremental effects of alternative strategies (e.g., different screening intervals for different subgroups). An example is provided in Mandelblatt *et al.*,¹⁰⁹ and this method is described in Mark *et al.*¹¹⁰

Scatter plot

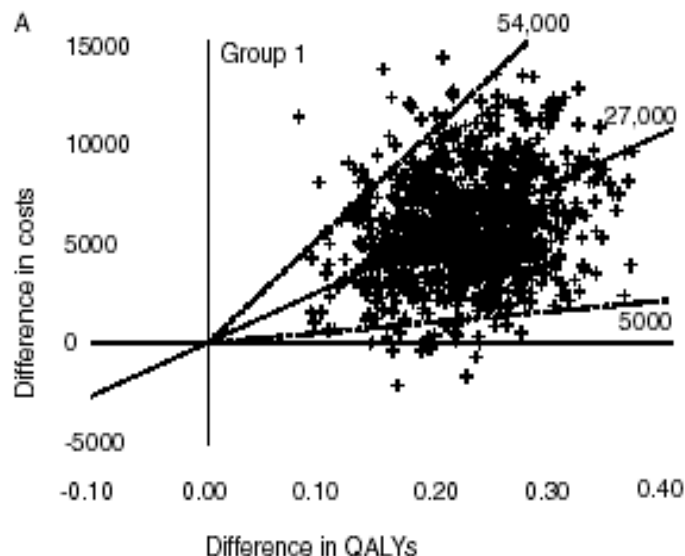
For a PSA, analysts are encouraged to present the ICER results using the scatter plot on the cost-effectiveness plane, as depicted in Figure 3. In this example, replicates are plotted in the plane using the bootstrap method to quantify uncertainty. The mean result (\$27,000 /QALY) and 95% confidence interval (\$5,000/QALY and \$54,000/QALY) are shown.

Figure 2: Cost-effectiveness plane



Source: Expert Review of Pharmacoeconomics and Outcomes Research, 2003, vol. 1, issue 1, pp 25-36;¹⁰⁸ reproduced with permission from Future Drugs Ltd.

Figure 3: Scatter plot



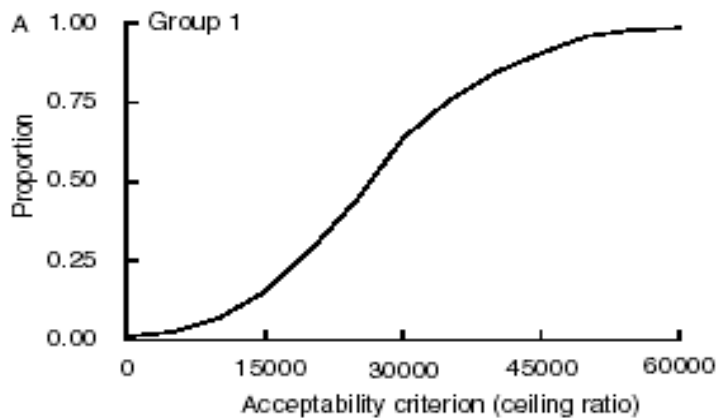
Source: Expert Review of Pharmacoeconomics and Outcomes Research, 2003, vol. 1, issue 1, pp 25-36;¹⁰⁸ reproduced with permission from Future Drugs Ltd.

Cost-effectiveness acceptability curve (CEAC)

In instances where the net benefit measure is used, the results of a PSA can be shown as a CEAC (Figure 4). The CEAC shows the probability that the intervention is cost-effective at each ceiling ratio (or willingness-to-pay threshold), given the data available. The curve illustrates that at the 50% point, the ICER is \$27,000/QALY. It also shows

that there is a 95% probability the intervention falls below \$54,000/QALY. The CEAC space can simultaneously present the results for uncertainty and variability using multiple CEACs for different patient strata (e.g., by risk of clinical event) and display the impact of varying other (non-sampling) parameters, such as the discount rate.

Figure 4: Cost-effectiveness acceptability curve



Source: Expert Review of Pharmacoeconomics and Outcomes Research, 2003, vol. 1, issue 1, pp 25-36;¹⁰⁸ reproduced with permission from Future Drugs Ltd.

APPENDIX 2: Review of Existing Economic Evidence

There are key steps for undertaking a review of existing economic evidence, and quality control measures that may minimize bias in the review process. The standard reporting format (Appendix 3) includes a section on reviewing the existing economic evidence of the intervention.

Refer to the UK National Health Service's (NHS) Centre for Reviews and Dissemination (CRD) publication: *Undertaking systematic reviews of research on effectiveness: CRD's guidance for those carrying out or commissioning reviews*, by Khan *et al.*¹¹¹ (<http://www.york.ac.uk/inst/crd/report4.htm>). A useful textbook is *Systematic reviews to support evidence-based medicine: how to review and apply findings of healthcare research*, by Khan *et al.*¹¹²

Steps in the Review Process

- Define the study question.
- Perform a literature search.
- Select the relevant studies.
- Extract data from the selected studies.
- Assess the quality of the selected studies.
- Synthesize and analyze the extracted data.
- Interpret and report the results.

To minimize bias, a review can incorporate the following elements of good practice:

- literature search by a qualified librarian or information specialist.
- transparent management process for selecting relevant studies.
- predetermined eligibility criteria for selecting relevant studies.
- criteria list to assess the quality of the studies.
- predetermined data extraction form.
- independent involvement of two reviewers in study selection, data extraction, and quality assessment.

The purpose of reviewing existing economic studies is to summarize the available knowledge in a systematic way that will be useful for decision

makers and researchers. To minimize bias, methods similar to those used for a systematic review of clinical studies can be applied to reviews of economic studies. The following guidance is provided to ensure quality control and minimize bias when undertaking the review process.

Protocol

A protocol should be written for a systematic review.¹¹¹ It should include a clear study question, and specify the eligibility criteria for the selection of relevant studies to be included in the review. Eligibility criteria can be framed in terms of the relevant PICOS, i.e., population, intervention, comparators, outcomes (for a cost-effectiveness analysis), and study design (i.e., the types of economic studies considered). Exclusion criteria may also be identified (e.g., abstracts only, studies set in developing countries).

Literature search

A crucial component in the review process is the design and execution of unbiased and comprehensive literature searches. Sources of economic evidence include electronic bibliographic databases, statistical data, reference lists from relevant articles, grey literature, research registers, and reliable web sites.^{111,113} Specialized economic databases such as the NHS Economic Evaluation Database (<http://www.york.ac.uk/inst/crd/nhsdhp.htm>) and the Health Economic Evaluation Database provide information on specific aspects of published economic evaluations. It may also be useful to contact the manufacturers of the technology, content experts, and appropriate regulatory agencies. It is recommended that an information specialist or librarian working with the review team design and execute the search strategies.

Study selection

Eligibility criteria are used to select the material retrieved through the literature search. During the first stage, potentially relevant citations are

selected, and this is followed by the selection of relevant articles. At each stage, two independent reviewers can perform the selection of material to minimize bias. Any disagreements between reviewers can be resolved by an agreed upon a priori method (e.g., consensus after discussion or involvement of a third party). It is recommended that a flow chart depicting the management of studies be included, such as that described in Moher *et al.*¹¹⁴

Data extraction

For studies identified for review, two independent reviewers can use a predetermined data extraction form to extract relevant information. A sample template of a data extraction form can be found in the CRD Guidance document.¹¹¹ The information extracted may include study characteristics, methods, key parameter values, sources of data, results, study conclusions, and source of study funding.

To inform judgments relating to transferability of study results to the setting or jurisdiction of interest, there should be transparent reporting of the estimates of costs, effectiveness, and preferences (utilities) in the studies. It is useful to report the type and quantities of key resources used and unit costs, and the year and currency of the costs.

Assessing study quality

Criteria lists can be used to assess the quality of selected studies in terms of their methodological elements.¹¹¹ Two possible criteria lists are the British Medical Journal Guidelines for economic submissions¹¹⁵ (<http://www.bmj.com/cgi/content/full/313/7052/275>) and the Consensus on Health Economic Criteria (CHEC) list¹¹⁶ (<http://www.beoz.unimaas.nl/chec>). To minimize bias,

two independent reviewers can apply the criteria list. The method of quality assessment should be described.

Data synthesis and analysis

Methods that are used to synthesize the information from selected studies should be described. A qualitative approach is most often used. This involves summarizing the extracted information in two tables: one for information about the characteristics of the selected studies, and another for the results. The latter may include a table showing the direction and magnitude of effects and total costs in the individual studies. The use of a permutation plot of results is another approach that may be used (see the CRD Guidance document).¹¹¹

Interpreting and reporting results

The main findings should be summarized, and any limitations (e.g., methods used, data availability, relevance of results) should be noted. Factors to consider when drawing conclusions include:

- number of studies selected
- consistency of results in (e.g., in a sensitivity analysis) and across studies
- magnitude of results
- quality or limitations of studies
- generalizability or relevance of studies to the target audience or jurisdiction.

Analysts can try to resolve the differences in results across studies by comparing the methods and inputs. Analysts should investigate whether the clinical evidence in a study is at odds with that used in the analyst's evaluation.¹¹¹ The comparison of the results with those of other studies and systematic reviews can be reported.

APPENDIX 3: Standard Reporting Format

A structured reporting format for the preparation of reports of economic evaluations ensures that studies are thoroughly presented, and organized consistently to facilitate review and comparison by decision makers.

It is suggested that reports follow this format as much as is practical, though in some instances, deviation from the format may be appropriate. For example, the report sections could be reordered or certain sections excluded if they are irrelevant to the evaluation. The study should be presented in a clear and transparent manner with enough information provided to enable the audience to critically evaluate the validity of the analysis. The Executive Summary and Conclusions should be written so that they can be understood by a non-technical reader.

Helpful hints on writing style and conventions are available in the Guidelines for Authors of CADTH Health Technology Reports.¹⁰⁵

Preface

- List of authors, affiliations, and a description of contributions
- Acknowledgements
- Disclosure of funding and reporting relationships, study sponsor, contractual arrangements, autonomy of analysts, and publication rights; declaration of conflicts of interest (guidelines and a declaration template can be found in Guidelines for Authors of CADTH Health Technology Reports¹⁰⁵ at <http://www.cadth.ca>)

Executive Summary

The Executive Summary should be one to two pages long and written in non-technical language.

- Issue: a statement about the economic issue or reason for evaluating the technology
- Objectives
- Methods
- Results: a numerical and narrative summary of the findings
- Discussion: study limitations, relevance of findings, health services impact
- Conclusions: state the bottom-line findings of the evaluation, uncertainty about the results, and caveats

Table of Contents

Abbreviations

Glossary

1. OBJECTIVES

- 1.1 Description of issue(s) addressed in the report
 - Set the scene for the reader, and include reasons for the analysis (e.g., funding or costs implications, issues of competing technologies).
- 1.2 Statement of study question
 - Define the study question, state it in an answerable form, make it relevant for the target audience.
 - Define the patients and population(s), intervention, and comparators.
 - State the primary perspective of the study and related secondary questions (e.g. impact of the intervention on subgroups).
 - Identify the primary target audience and possible secondary audiences.
- 1.3 Objectives of the study

2. BACKGROUND

- 2.1 General comments on condition
- State the condition and population-specific patient group(s) being studied.
 - List the etiology, pathology, diagnosis, risk factors, prognosis (if relevant).
 - Describe the epidemiology (i.e., incidence or prevalence), burden of the condition in Canada.
 - Describe the economic impact and burden of the condition in Canada.
 - Describe the current clinical practice in Canada. Refer to clinical practice guidelines (if relevant). Include a description or comparison of alternatives for the indication.
- 2.2 Technology description
- For drugs, state brand and generic names, dosage form, route of administration, recommended dosage, duration of treatment, therapeutic classification, mechanism of action.
 - For non-drug technologies, state basic features, underlying theory or concept.
 - List advantages and disadvantages (e.g., relating to clinical use).
 - State adverse events, contraindications, cautions, warnings.
 - Describe setting for the technology if relevant (e.g., hospital-based)
 - Give unit cost of the intervention and comparators.
- 2.3 Regulatory status
- List the approved indication(s) in Canada that is the topic of the study, including applicable population and subgroups, and date of approval.
 - Give additional approved indication(s) in Canada.
 - Include the regulatory status and approved indications in other countries.

3. REVIEW OF ECONOMIC EVIDENCE

- Discuss existing economic studies that address the same technology, and similar study question(s). Include a summary of methods and results of reviewed studies (can be summarized in a table and placed in appendices).
- If a systematic review has been undertaken, identify and discuss the steps of the review (Appendix 2). Include as appendices: literature search strategy, flow chart of included and excluded studies, data extraction form, and quality assessment criteria list.
- Comment on the relevance and generalizability of the results of the reviewed studies to the target audience.

4. METHODS

Report how each element of the economic evaluation, as outlined in the Economic Guidelines, has been handled.

- 4.1 Types of economic evaluation
- Describe the CUA, CEA, CMA, CBA, or CCA, and justify the type(s) conducted.
- 4.2 Target population
- Describe target population(s) and the care setting for the intervention or expected use. Where appropriate, describe and justify the population subgroups analyzed.
- 4.3 Comparators
- Describe and justify selected comparators; relate choice of comparators to the study population, and the local context or practice.
- 4.4 Perspective
- State and justify the perspective(s) used in the analysis (e.g., public payer, society).

- 4.5 Effectiveness
- a) Evidence of efficacy and effectiveness
- Give details about the evidence on efficacy and effectiveness used in the analysis (if lengthy, place in a separate section preceding section 3 or in an appendix).
 - For clinical studies, report on PICOS (participants, intervention, comparator or control, outcomes, study design).
 - Describe adverse events, where important and relevant.
 - Indicate sources of information (e.g., trials, a meta-analysis of individual trials, literature, expert opinion).
- b) Modelling effectiveness
- Identify factors that are likely to have an impact on effectiveness (e.g., adherence, diagnostic accuracy), and describe how these were factored into the analysis. Explain causal relationships and techniques used to model or extrapolate the data (e.g., short-term to long-term outcomes, surrogate to final outcomes). Describe the strength of the evidence for the relationships and links.
- 4.6 Time horizon
- Indicate the time horizon(s) used in the analysis, and its justification.
- 4.7 Modelling
- a) Modelling considerations
- Describe the study design: modelling, trial-based, prospective or retrospective analysis, or combination of methods.
 - Describe the model structure: description of the scope, structure, and assumptions made with justification; inclusion of a model diagram or decision tree is recommended.
 - Describe how the model was validated. This can involve validating different aspects of the model (e.g., model structure, data and assumptions, model coding) and using different validation methods (e.g., top-down versus bottom-up methods for checking of model results, comparison with other models). If relevant, results from validation exercises can be attached as appendices.
- b) Data considerations
- List other data or assumptions with sources and justification. This may include details about epidemiological factors, such as prevalence or incidence of the condition.
 - Where a statistical analysis has been conducted, describe how censored data were handled.
- 4.8 Valuing outcomes
- Identify, measure, and value all relevant outcomes, including important adverse events, for each alternative.
 - Give the sources of information and data, assumptions, and justification
 - Give the HRQL measurement approach used, with the justification (a copy of the instrument may be included in an appendix). Describe the methods of eliciting preferences and the population measured.
 - Include other outcomes that were considered but rejected (with rationale).
- 4.9 Resource use and costs
- Identify, measure, and value all resources included in the analysis.
 - Report the costing methods used (e.g., gross or micro-costing).
 - Classify resources into categories relevant to the perspective (e.g. relevant agencies comprising the public payer).
 - Report resource quantities and unit costs separately.
 - Distinguish additional costs from averted costs.
 - Report the method used for costing lost time, including productivity losses. Identify, measure, and value

- lost time. Provide justification when time costs are not considered.
- Report all sources of information and data and assumptions.
- 4.10 Discount rate
- Indicate the discount rates used for costs and outcomes, and the justification.
- 4.11 Variability and uncertainty
- a) Handling variability
- Describe and justify the basis for stratification of the target population. State whether there are a priori identifiable subgroups for which differential results might be expected (e.g., based on effectiveness, preferences and utilities, costs).
 - Describe how other types of variability (e.g., variation in costs or practice patterns) were analyzed, and provide the justification.
- b) Handling uncertainty
- Identify sources of uncertainty (parameter or modelling) in the analysis.
 - Describe methods used to analyze uncertainty (e.g., a DSA or a PSA).
 - For a DSA, state the range of parameter values and assumptions tested; provide sources and justification for each.
 - For a PSA, state the form of probability distributions and the number of Monte Carlo iterations, with sources and justification.
- 4.12 Equity
- State equity assumptions (e.g., a QALY is equal for all).
 - Identify equity-relevant characteristics of the main subgroups that may benefit, or be adversely affected by, the technology, and describe how they were analyzed.
- 5. RESULTS**
- 5.1 Analysis and results
- Present all analyses in a step-by-step fashion, so the calculations can be replicated if desired. Present the analysis first in a disaggregated fashion showing all perspectives separately. If relevant, show separately the analysis of different time horizons and types of economic evaluations performed.
 - Show undiscounted totals (gross and net) before aggregation and discounting.
 - Show the components of the ICER numerator (cost of each alternative), and ICER denominator (outcomes of each alternative).
 - For outcomes, express in natural units first, and then translate into alternative units, such as QALYs or monetary benefits.
 - Provide tables of results in appendices; a visual display of results is encouraged.
- 5.2 Results of variability analysis
- Give the results for all subgroups analyzed.
 - Indicate the distribution impacts (i.e., benefits, harms, costs) and ICER results for any subgroups that are relevant for equity purposes.
 - Indicate the results of analysis for other types of variability (e.g., variation in costs or practice patterns).
- 5.3 Results of uncertainty analysis
- State the results of sensitivity analysis.
 - Identify the greatest sources of uncertainty (i.e., key drivers).
- 6. DISCUSSION**
- 6.1 Summary of results
- Critically appraise and interpret the main findings of the analysis in the context of all reasonable alternatives.

- Address the place of the intervention in practice, based on the evidence.
 - Discuss the uncertainty of the results and the key drivers of results.
 - Discuss the trade-off between benefits, harms, and costs.
- 6.2 Study limitations
- Discuss key limitations and issues concerning the analysis, including methodological limitations and issues, validity of assumptions, strength of the evidence for data, relationships or links used in the model. Describe whether the data and methods used may bias the analysis in favour of one alternative.
- 6.3 Other economic studies
- Where other economic studies have been reviewed, compare the methods and results of these studies with the present study.
- 6.4 Generalizability
- Comment on the generalizability or relevance of results, and the validity of the data and model to the relevant jurisdictions and populations.
 - Comment on regional differences in terms of disease epidemiology, population characteristics, clinical practice patterns, resource use patterns, unit costs, and other factors of relevance. Where differences exist, discuss the impact on the results (expected direction and magnitude), and the conclusions.
- 6.5 Equity considerations
- Indicate the distributional considerations (e.g., primary beneficiaries and those adversely affected)
 - List other ethical and equity implications or issues. For example, are there likely to be variations in patients' access to the intervention, as defined geographically or by patient characteristics? Does the technology address the unmet needs of certain disadvantaged groups (e.g., telehealth for those in remote locations)? Is the technology responsive to those with greatest need, for whom there is no alternative treatment (e.g., "rule of rescue")?
- 6.6 Health services impact
- Comment on the possible shift in health care resources or impact on the health services of adopting the technology. How does this change current practice?
 - Practical considerations include health service planning concerns (e.g., increased or decreased need for related health care services); implementation mechanisms (e.g., changes to the physician fee schedule, development of new clinical practice guidelines); human resource implications (e.g., training, workload); legal and regulatory issues; and psychosocial issues (e.g., social acceptability of the intervention).
 - With regards to budget impact, identify the potential funders of the technology.
 - Describe the factors likely to determine the budget impact (e.g., epidemiological factors, current comparator baseline use, intervention uptake rate).
 - Estimate the resource use and budget impact for various scenarios (e.g., base case, high and low ranges); disaggregate the latter into gross expenditure, savings, and net expenditure over the relevant period.
- 6.7 Future research
- Identify knowledge gaps and areas for further research that is relevant to Canada.
- 7. CONCLUSIONS**
- Address the research objective(s) and question(s).
 - Summarize the bottom-line findings of the study, aggregate impact, uncertainty about the results,

appropriate uses for the intervention (e.g., population subgroups), and any caveats.

8. REFERENCES

9. APPENDICES

- Include in appendices (depending on practical considerations, and amount of material) a table of data and sources; data collection forms,

questionnaires, instruments; diagram of model decision tree; step by step details of analyses, including intermediate results; tables of results; visual presentation of results (e.g., figures, graphs).

- If a systematic review was conducted, include literature search strategy, flow chart of included and excluded studies, data extraction form, quality assessment criteria list, summary of methods, and results of studies.

APPENDIX 4: Glossary

Adverse event – an undesirable effect of a health technology.

Adherence (sometimes referred to as “compliance”) – adherence is achieved if the patient achieves the following three components of treatment: acceptance, the initial decision of the patient to accept treatment; compliance, the consistency and accuracy with which the patient follows the recommended treatment regimen, and persistence, long-term continuation of treatment³⁴

Analysis of extremes – type of sensitivity analysis that involves changing all parameters or other model inputs to extreme values (either best or worst case) simultaneously

Analyst – someone who conducts an economic evaluation

Bayesian method – a branch of statistics that uses prior information on beliefs for estimation and inference

Budget impact analysis – application of methods to estimate planned resource use and expenditure of a budget over a period of time

Comparator – alternative to which the intervention is compared

Condition – medical condition that includes disease

Conjoint analysis (or discrete choice experiment) – technique for valuing the benefits of health technologies by asking respondents to make discrete choices between alternative bundles of attributes that form the technology. If the cost of the technology is one of the included attributes, this technique allows one to determine willingness to pay indirectly

Contingent valuation – technique for valuing the benefits for health technologies, typically by determining individuals’ maximum willingness to pay for the availability of that technology, or the minimum amount that they would accept as compensation for not having that technology available

Cost – the value of opportunity forgone (strictly the best opportunity forgone), as a result of engaging resources in an activity (see opportunity cost); there can be a cost without the exchange of money; range of costs (and benefits) included in a particular economic evaluation depends on perspective taken; average costs are average cost per unit of output (i.e., total costs divided by total number of units produced); incremental costs are extra costs associated with intervention compared to alternative; marginal cost is cost of producing one extra unit of output

Cost-benefit analysis (CBA) – type of economic evaluation that values costs and outcomes in monetary terms

Cost-consequence analysis (CCA) – type of economic evaluation in which costs and outcomes are listed separately in a disaggregated format, without aggregating these results (e.g., usually in incremental cost-effectiveness ratio)

Cost-effectiveness acceptability curve (CEAC) – graphical representation of probability that intervention is cost-effective at various willingness-to-pay thresholds (or ceiling ratios), given the data available

Cost-effectiveness analysis (CEA) – type of economic evaluation in which outcome is measured in natural (health) units, such as life-years gained or clinical event avoided; term is also sometimes used to refer to all types of economic evaluations

Cost-minimization analysis (CMA) – type of economic evaluation in which intervention and alternatives are considered equivalent in terms of factors relevant to decision (other than cost), and so, lowest cost is selected

Cost-utility analysis (CUA) – type of economic evaluation in which outcome is measured as health-related preferences, often expressed as quality-adjusted life-years (QALYs)

Decision tree – graphical representation of decision, incorporating alternative choices, uncertain events (and their probabilities), and outcomes

Deterministic sensitivity analysis (DSA) – method of decision analysis that uses one-way and multi-way sensitivity analyses; involves substituting different values (or processes) for one or more model inputs; one-way sensitivity analysis involves changing value of one model input through range of plausible values while keeping other model inputs constant

Direct costs – value of all health care resources that are used in provision of intervention or in dealing with adverse events or other current and future consequences linked to it; typically include medications, physician office visits, laboratory tests, and hospitalizations

Direct preference measurement – direct measurement of preferences by using techniques such as standard gamble and time trade-off

Discounting – process by which streams of future costs and benefits occurring in future (typically beyond one year) are converted to equivalent present values using discount rate

Distributional impact – distribution of benefits, harms, and costs (and cost-effectiveness) of technology across population subgroups

Dominance (simple, strong or strict) – state when intervention is more effective and less costly than alternative

Economic evaluation – application of analytical methods to identify, measure, value, and compare costs and consequences of alternatives being

considered;¹ addresses issue of efficiency to aid decision making for resource allocation

Effectiveness – extent to which intervention produces benefit in defined population in routine or “real world” circumstances; compare with efficacy

Effect modification – change in magnitude of effect measure according to value of third factor; this factor is called an effect modifying factor

Efficacy – extent to which intervention produces benefit in defined population in controlled or ideal circumstances; compare with effectiveness

Efficiency – extent to which maximum possible benefit is achieved out of available resources (i.e., good value for money); two (related) types of efficiency are technical and allocative

Equity – as it relates to health, “fairness” in allocation of resources, interventions, or outcomes among individuals or groups⁸⁹

Extended (or weak) dominance – state when intervention is more costly and more effective, and has lower incremental cost-effectiveness ratio, than alternative

External validity – extent to which one can generalize study conclusions to populations and settings of interest outside study

Extrapolation – prediction of value of model parameter outside measured range or inference of value of parameter of related outcome (e.g., extrapolation of reduction in rate of progression to AIDS from improvement in HIV viral load)

Final outcome – health outcome that is related directly to length and quality of life, examples include life-years gained and quality-adjusted life-years

Friction cost approach (FCA) – method of estimating productivity costs by calculating value of production losses during friction period (i.e., between start of absence from work and replacement); compare with human capital approach

Generalizability – problem of whether one can apply or extrapolate results obtained in one setting or population to another;⁹⁴ term may also be referred to as “transferability,” “transportability,” “external validity,” “relevance,” or “applicability”

Gross (or top down) costing – costing approach that uses large components as basis for costing, such as cost per hospital day; compare with micro-costing costing

Guideline Statement – in context of Economic Guidelines, key point of guidance for analyst to follow when conducting an economic evaluation

Health technology – application of scientific or other organized knowledge (including any tool, technique, product, process, method, organization, or system) to practical tasks that promote health; prevent, diagnosis, and treat conditions; or improve rehabilitation and long-term care. It includes drugs, vaccines, medical devices, medical and surgical procedures, disease prevention and screening activities, health promotion activities, and organizational and managerial systems such as telemedicine, technologies can be different strategies for management or treatment of condition

Health-related quality of life (HRQL) – physical, social, and emotional aspects that are relevant and important to an individual’s well-being; can be assessed using a disease-specific, generic, or a preference-based measurement tool

Human capital approach (HCA) – method of estimating productivity costs based on individual’s entire employment period that is lost because of illness; compare with friction cost approach

Important patient outcome – in context of Economic Guidelines, valid outcome of importance to health of patient; outcomes include disease-specific events (e.g., avoidance of stroke and HIV infection), final outcomes (e.g., life-years gained), and validated surrogate outcomes

Incremental cost-effectiveness ratio (ICER) – ratio of difference in costs of intervention and alternative to difference in outcomes

Indirect preference measurement – use of instruments (e.g., Health Utilities Index and ⁵⁵EQ-5D)^{58,59} to measure preferences, without undertaking direct measurement

Intervention – health technology of interest for assessment in economic evaluation

Leisure time – time that is not spent at paid or unpaid (e.g., informal caregiving) work

Limited use criteria – category of reimbursement for drug in formulary, such that drug is reimbursed for more restrictive subgroup of patients (e.g., based on specific clinical criteria) than that indicated in drug’s licensing

Micro-costing (or bottom-up) costing – costing approach based on detailed resources used by patient on item by item basis; compare with gross costing

Meaningful (or important) difference – in context of Economic Guidelines, difference or impact that is likely to have a substantial impact on main results of analysis, or important bearing on decision facing decision maker (e.g., is likely to change decision or conclusions of the report)

Model inputs – parameters (e.g., outcomes, resource use, utilities) and model design features (e.g., model structure, analytical methods, and assumptions) of model

Model uncertainty – uncertainty relating to model design features (e.g., model structure, analytical methods and assumptions); model uncertainty depends on choices and assumptions made by modeller.

Monte Carlo simulation – type of simulation modelling that uses random numbers to capture effects of uncertainty; multiple simulations are run, with value of each uncertain parameter in analysis selected at random from probability distribution for each simulation; simulation results are compiled, providing probability distribution for overall result

Net benefit – refers to method of reporting results of economic evaluations (versus incremental cost-effectiveness ratio) in terms of monetary units

(called net monetary benefit) or units of outcome (called net health benefit); in cost-benefit analysis, (incremental) net benefit is difference in total benefit and total cost of intervention less the difference in total benefit and total cost of alternative

Opportunity cost – costs of resources consumed expressed as value of next best alternative for using resources

Outcome – consequence of condition or intervention; in Economic Guidelines, outcomes most often refer to health outcomes, such as surrogate outcome or important patient outcome

Parameter uncertainty – uncertainty about true numerical values of parameters (e.g., health outcomes, utilities, and resource use) of model

Perspective – viewpoint from which economic analysis is conducted (e.g., public payer, society, individual); defines which costs will be examined

Point estimate – estimate of parameter of interest

Preference – desirability of particular outcome or situation; terms preference and utility generally used synonymously as measure of HRQL in Economic Guidelines; utilities are preferences obtained by methods that involve uncertainty (e.g., standard gamble approach), whereas values are preferences derived by methods that do not involve uncertainty (e.g. time trade-off approach); both utilities and values are preferences

Present value – value of future cost or benefit after adjusting for time preferences by discounting

Probability – expression of degree of certainty that event will occur, on scale from zero (certainty that event will not occur) to one (certainty that event will occur)

Probability distribution (or probability density function) – numerical or mathematical representation of relative likelihood of each possible value that parameter may have

Probabilistic sensitivity analysis (PSA) – method of decision analysis in which probability distributions are specified for uncertain parameters

(e.g., outcomes, costs, utilities); a Monte Carlo simulation is performed; and resulting probability distribution of expected outcomes and costs is displayed

Productivity costs – the costs associated with lost or impaired ability to work because of morbidity or death

Protocol-driven costs – resource use that is required as part of clinical trial protocol; is usually not part of usual care

Public payer – publicly funded health care system

Quality-adjusted life year (QALY) – measure of health outcome that combines effect of intervention on length of life and quality of life (usually expressed in terms of utilities); common measure of outcome in cost-utility analysis

Recommended (or appropriate) care – high quality, clinically appropriate care determined by reference to recommendations in evidence-based clinical practice guidelines or by clinical experts

Reference Case – set of preferred methods for analyst to follow when conducting base case analysis in economic evaluation

Scenario analysis – form of multi-way sensitivity analysis, which involves simultaneously substituting model inputs associated with identifiable subgroup of interest; variability can be assessed through scenario analysis

Sensitivity analysis – method of analysis to determine how and whether changes in uncertain model inputs affect main results and conclusions of analysis

Serious adverse event – adverse event that results in death; is life threatening; requires in-patient hospitalization or prolongation of existing hospitalization; results in persistent or significant disability or incapacity, or is congenital anomaly or birth defect; adverse events that require significant medical intervention to prevent one of these outcomes also considered to be serious

Standard gamble – technique used to assess a person's utility for outcomes or health states that

differ in quality or length of life; done by asking person to choose given health state, or gamble between ideal health and immediate death; probability of ideal health versus immediate death systematically changed until person no longer prefers either gamble or health state; compare with time-tradeoff

Stratified analysis – process of analyzing smaller, more homogeneous subgroups according to specified criteria such as age groups, socioeconomic status, where there is variability (heterogeneity) in population.

Surrogate outcome – laboratory measurement or a physical sign used as substitute for clinically meaningful endpoint that measures directly how patient feels, functions, or survives; examples for cardiovascular disease include blood pressure or cholesterol level; surrogate outcomes can be validated or unvalidated

Systematic review – form of structured literature review that addresses question formulated to be answered by analysis of evidence; involves application of explicit methods to search literature, apply predetermined inclusion and exclusion criteria to literature, critically appraise relevant literature, and extract and (qualitatively or quantitatively) synthesize data from evidence base to formulate findings

Threshold analysis – type of sensitivity analysis in which model input is varied over a range to determine value of input that would lead to major changes in conclusions

Time costs – relates to time patient spends seeking care (e.g., for travel) or participating in or undergoing an intervention; time costs can also be related to productivity costs, including unpaid caregiver time off from work

Time horizon – period of time over which costs and outcomes are measured in economic evaluation

Time-tradeoff – technique of preference assessment in which subject is asked to determine length of time in ideal health that he or she would find equivalent to longer period of time with specific condition; compare with standard gamble

Tornado diagram – diagrammatic display of results of one-way sensitivity analysis; each bar in diagram represents range of change of model inputs resulting from analysis of maximum and minimum values

Transfer (or income transfer) payment – payment made to individual (usually by government body) that does not perform any service in return; examples are social security payments and employment insurance benefits

Uncertainty – a state in which true value of parameter or structure of process is unknown

Unvalidated (or unproven) surrogate outcome – surrogate outcome that has not been proven to be predictive of important patient outcome

User – user of economic evaluations, most often decision and policy makers in Canada’s publicly funded health care system

Usual care – most common or most widely used alternative in clinical practice for specific condition; also referred to as “existing practice,” “current practice,” “typical care,” or “status quo”

Utility – cardinal measure of preference for specific health outcome; technically, utilities are preferences obtained by methods that involve uncertainty (e.g., standard gamble approach)

Validated surrogate outcome – surrogate outcome proven to be predictive of important patient outcome, based on meeting certain conditions

Validity – extent to which technique measures what it is intended to measure

Valuation – process of quantifying desirability of outcome in utility or monetary terms or of quantifying cost of resource or individual’s productivity in monetary terms

Variability – reflects known differences in parameter values associated with identifiable differences in circumstances;⁶⁸ is represented by frequency distributions; variability can be attributed to diverse clinical practice patterns in different geographical areas or settings, or inherent

variability in patient population (i.e., patient heterogeneity)

Willingness-to-pay (WTP) approach – evaluation method used to determine maximum

amount of money individual is willing to pay for particular outcome or benefit (e.g., receive health care service); method is often used in cost-benefit analysis to quantify outcome in monetary terms

A large, stylized, semi-transparent maple leaf logo is centered on the page. The leaf is composed of several overlapping, jagged shapes that form its characteristic shape. The background of the page is a gradient of blue and purple, with a solid green horizontal band at the top and a solid blue horizontal band at the bottom.

Publications can be requested from:

Canadian Agency for Drugs and Technologies in Health
600-865 Carling Avenue, Ottawa ON Canada K1S 5S8

Telephone (613) 226-2553 • Fax (613) 226-5392 • E-mail pubs@cadth.ca

www.cadth.ca www.acmts.ca