CADTH

CADTH METHODS AND GUIDELINES

Guidelines for the Economic Evaluation of Health Technologies: Canada

4th Edition

Acknowledgements

Guideline Working Group Members

Working group members were responsible for: identifying and discussing key issues related to each topic, drafting topic sections, reviewing all draft topic sections, reviewing the draft consolidated report, addressing peer review and stakeholder feedback, and reviewing and approving the final version of *Guidelines*.

CADTH Members

Karen M. Lee, MA
Director, Health Economics
CADTH
Ottawa, Ontario

C. Elizabeth McCarron, MA, MSc, PhD
Health Economist
CADTH
Ottawa, Ontario

Academic Members

Stirling Bryan, PhD
Professor, School of Population and Public Health
University of British Columbia
Vancouver, British Columbia

Doug Coyle, MA, MSc, PhD
Professor, School of Epidemiology, Public Health and Preventive Medicine
University of Ottawa
Ottawa, Ontario

Murray Krahn, MSc, MD, FRCPC
Director, Toronto Health Economics and Technology Assessment (THETA) Collaborative
University of Toronto
Toronto, Ontario

Christopher McCabe, BA, MSc, PhD
Capital Health Endowed Research Chair
University of Alberta
Edmonton, Alberta

Contributors

CADTH would like to acknowledge the following individuals for their contributions:

Sheena Gosain, BHSc, MSc helped with organizing the project, provided research support to topics in the *Guidelines*, and drafted and reviewed versions of topic sections.

Carlo Marra, BSc(Pharm), PharmD, PhD was a working group member early on in the project, contributed to the Measurement and Valuation of Health section, and reviewed early drafts of topic sections.

Sandra Milev, MSc contributed to the Modelling and Effectiveness topic sections.

David Kaunelis, MLIS conducted literature searches and provided information services support.

Kim Ghosh, BA, PMP provided project management support and helped with organizing the project.

Reviewers

The reviewers reviewed either the full draft or specific sections of the *Guidelines*.

Nick Bansback , BSc, MSc, PhD
Assistant Professor, School of Population and Public Health
University of British Columbia
Vancouver, British Columbia

Ahmed Bayoumi, MD, MSc, FRCPC
Professor, Department of Medicine and Institute of Health Policy Management and Evaluation
University of Toronto
Toronto, Ontario

W.B.F Brouwer, PhD
Professor of Health Economics and Chairman
Institute of Health Policy & Management
Erasmus University Rotterdam
Netherlands

Anthony Budden, BBHSc
Health Economist
CADTH

Andy Chuck, MPH, PhD
Director of Economic Evaluation and Analytics
Institute of Health Economics
Edmonton, Alberta

Lauren E. Cipriano, PhD
Assistant Professor
Ivey Business School, Western University
London, Ontario

Philip Jacobs, PhD, CMA
Professor, University of Alberta
Edmonton, Alberta

Scott Klarenbach, MD, MSc
Professor, Department of Medicine
University of Alberta
Edmonton, Alberta

Vivian Ng, MSc, PhD
Manager, Health Economic Evaluation
Health Quality Ontario
Toronto, Ontario

Petros Pechlivanoglou, PhD
Scientist
The Hospital for Sick Children (SickKids) Research Institute
Toronto, Ontario

Sheri Pohar, BScPharm, MScPharm, PhD
Scientific Advisor
CADTH

Mohsen Sadatsafavi, MD, PhD
Assistant Professor, Faculty of Pharmaceutical Sciences
University of British Columbia
Vancouver, British Columbia

Mark Sculpher, PhD
Professor of Health Economics, Centre for Health Economics
University of York
Heslington, York, UK

Eldon Spackman, MA, PhD
Assistant Professor
University of Calgary
Calgary, Alberta

CADTH

Conflict of Interest Declaration

Anthony Budden declared co-authoring CADTH's *Guidance Document for the Costing of Health Care Resources in the Canadian Setting*, which is referred to within the *Guidelines*.

Murray Krahn declared research funding from the Canadian Liver Foundation through an unrestricted grant from Gilead, Prostate Cancer Canada through an unrestricted grant from Janssen, Pfizer Canada, Sanofi Canada, and the Hepatitis C Settlement Fund.

Carlo Marra declared consultancy for various pharmaceutical and consulting companies.

Christopher McCabe declared research funding from Johnson & Johnson Inc. through the University of Alberta Hospital Foundation, and consultancy for various pharmaceutical companies.

Mark Sculpher declared consultancy for various pharmaceutical and other life sciences companies.

Eldon Spackman declared consulting for various pharmaceutical companies.

Table of Contents

Abbreviations

CBA	cost-benefit analysis
CCA	cost-consequence analysis
CEA	cost-effectiveness analysis
CEAC	cost-effectiveness acceptability curve
CEAF	cost-effectiveness acceptability frontier
CMA	cost-minimization analysis
CUA	cost-utility analysis
EQ-5D	EuroQol 5-Dimensions Questionnaire
HRQoL	health-related quality of life
HUI	Health Utilities Index
ICER	incremental cost-effectiveness ratio
QALY	quality-adjusted life-year
SF-6D	Short Form 6-Dimensions health status classification system

Conventions

Clinical or care pathway
All health-related pathways necessary to model the costs and outcomes relevant to the decision problem.

Consistency
Uniformity of data sources across parameters.

Credibility
A perceived lack of bias, where bias refers to the systematic deviation of the estimated value from the true underlying value.

Decision problem
The decision the economic evaluation is designed to inform.

Deterministic analysis
Data parameters represented by the expected values of individual data elements (i.e., point estimates).

Expert elicitation
The formal elicitation of quantitative input from relevant experts regarding the magnitude of a given parameter and its uncertainty.

Expert input
A potential source of data within the totality of available information, comprising both expert elicitation as well as existing expert elicitation studies.

Expert judgment
Qualitative input from experts.

Fitness for purpose
Relevance to the decision problem.

Non-reference case
Alternative methods to those recommended in the reference case for the purposes of assessing methodological uncertainty. Can accompany the reference case and be provided to decision-makers, but the impact of departing from the reference case should be explicitly stated.

Probabilistic analysis
Data parameters represented by statistical distributions rather than point estimates.

Reference case
A set of recommended methods to be used for all evaluations that promote uniformity and transparency, and enable the comparison of results for different technologies and different decisions.

Scenario analysis
Alternative scenarios carried out to examine sources of uncertainty (e.g., structural) within the reference or non-reference case analysis. One complete analysis should be provided for each alternative scenario.

Social decision-making viewpoint
The premise that the health care decision-maker, acting on behalf of a socially legitimate higher authority, seeks to maximize the degree to which an explicit policy objective (e.g., improving the overall health of the population) is achieved subject to the available resources.

Foreword to the Fourth Edition

The fourth edition of the *Guidelines for the Economic Evaluation of Health Technologies: Canada* follows publications in November 1994 (first edition), October 1997 (second edition), and March 2006 (third edition). The fourth edition reflects the experience gained through using the previous editions, and takes into account the methodological advancements that have occurred in the economic evaluation of health technologies since 2006.

The development of the fourth edition of the *Guidelines* proceeded as follows:

- Guideline topics from the third edition were reviewed to determine areas where methodological advancements had occurred.
- Health economic methods literature was reviewed and health economic experts were consulted.
- Gaps within the topic areas were identified and research was commissioned.

The goals in developing the fourth edition of the *Guidelines* were as follows:

- to provide clear, concise, and practical guidance of a high standard for researchers
- to meet the needs of decision-makers for reliable, consistent, and relevant economic information
- to highlight areas where methodological issues remain unresolved and more research is required
- to allow for flexibility, innovation, and alternative approaches, particularly where methodological issues are unresolved.

Throughout the process, the inherent tensions among these goals required that compromises be made. Practical considerations included the applicability of the recommendations in terms of meeting 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 resource allocation decisions being made as a result of implementing the findings of a poor-quality evaluation.

In preparing the fourth edition of the economic guidelines, consideration was given to all the comments received from all reviewers. Decisions relating to methodological issues were achieved through consensus.

Introduction

The purpose of these *Guidelines* is to inform the conduct by providing best practices for those undertaking economic evaluations of health care technologies in Canada in order to produce credible standardized economic information that is relevant and useful for decision-makers in Canada's publicly funded health care system. Economic evaluations of health care technologies involve the assessment of the cost and effect trade-offs of any interventions, programs, or policies that impact health outcomes. These evaluations may be conducted alongside individual-level studies or through decision-analytic models that synthesize evidence from multiple sources. While the guidance contained in this document pertains primarily to model-based evaluations, many aspects apply equally to evaluations based on individual-level data.

Health economic evaluations are designed to inform decisions. There are, however, different views on how collective decisions regarding health care resource allocation should be informed. In broad terms, these views can be characterized as welfarism or extra-welfarism and social decision-making.[1] Welfarism contends that collective decisions about health care should be based on the objective of maximizing social welfare. A strict welfarist view defines social welfare based on individual preferences (expressed or revealed), whereas an extra-welfarist view expands the definition to include other social arguments.[1,2]

The social decision-making viewpoint is based on the premise that the health care decision-maker, acting on behalf of a socially legitimate higher authority, seeks to maximize the degree to which an explicit policy objective (e.g., improving the overall health of the population) is achieved subject to the available resources. The role of economic evaluation in this framework then becomes one of informing social decisions in health rather than prescribing social choice.[1,3] Considering the approach to collective health care decisions and the role of economic evaluations within the Canadian context, a social decision-making viewpoint has been adopted for these *Guidelines*.

This viewpoint involves two types of collective decisions, which can be classified according to who makes the decision, who is affected by the decision, and whether the decision yields a general rule or a judgment.[4] The first of these decisions relates to the democratic election of a socially legitimate higher authority. The second type of decision is represented by health care decision-makers, acting as the agent of the higher authority and informed by the results of the economic evaluation, and their judgment regarding the cost and effect trade-off of the interventions being compared. It is through the social legitimacy of the higher authority that the decision-maker is presumed to make decisions reflective of what the general population considers to be socially valuable. It is these social values that the researcher should endeavour to reflect in the evaluation.

The concepts of decision-making under scarcity and the efficient allocation of resources are central to the economic evaluation of health care technologies. The recommendations contained in these *Guidelines* are focused on achieving technical efficiency within a constrained budget. Technical efficiency refers to obtaining the maximum possible improvement in an outcome from a given set of resource inputs.[5] Assuming that decision-makers are likely to face exogenous budget constraints,[6] the opportunity cost of a new investment will fall on the health care budget rather than on other sectors of public expenditure or the taxpayer.[3] Accordingly, these *Guidelines* adopt a "supply-side" estimate of the cost-effectiveness threshold, which assumes that reimbursing a new technology will displace some other technology or health care service. In that way, the recommendations contained in the *Guidelines* support the management of technologies along the life cycle, from decisions informing adoption and reimbursement to potential displacement or disinvestment.

Economic evaluations produce an estimate of the cost and effect trade-off of two or more interventions, not a decision as to the cost-effectiveness of one intervention relative to another. The determination of whether an intervention represents an efficient allocation of resources depends on the decision-maker's cost-effectiveness threshold. Furthermore, the role of economic evaluation in terms of informing health care resource allocation decision-making is twofold. Firstly, economic evaluation has a role to play in informing decisions based on the currently available information. In keeping with the social decision-making viewpoint underlying these *Guidelines*, these decisions should be based on the expected cost-effectiveness, given the existing information and not statistical inference. Secondly, by identifying areas of uncertainty, economic evaluation can also contribute to informing decisions about the need for further research to help resolve these uncertainties. Both of these decisions will reflect the evolving nature of the evidence base along a technology's life cycle.

In terms of facilitating the decision-making process, the *Guidelines* recommend the adoption of a reference case analysis. The purpose of the reference case is to encourage comparability across evaluations and to ensure decision-makers can be confident that they are using a consistent decision framework across all decisions. The reference case does not preclude the application of other methods or jurisdiction-specific recommendations where appropriate to address the decision problem, but does recommend that any deviations be justified on the basis of the decision problem and presented as additional non-reference case analyses. Where there is uncertainty as to the appropriate methodological approach for addressing a specific component of the decision problem, the results of the reference case can be compared relative to any non-reference case analyses.

Guidance is provided on the recommended approach to the decision problem, types of evaluations, target population, comparators, perspective, time horizon, discounting, modelling, effectiveness, measurement and valuation of health, resource use and costs, analysis, uncertainty, and equity. Guideline Statements are presented first, followed by the Guidelines in Detail. Recommendations on the appropriate reporting of an economic evaluation are also provided. The *Guidelines* are written for an audience that is technically literate about the methods of economic evaluation, such that background on methods can be avoided. The *Guidelines* are neither intended to be, nor should they be, viewed as a textbook. References have been provided for readers to obtain additional information.

The guidance contained herein represents CADTH's current recommendations for the conduct of economic evaluations of health care technologies. These recommendations apply to a variety of health technologies, including those that promote health, prevent and treat conditions, or improve rehabilitation and long-term care. Economic evaluations are used to inform decisions about 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 may refer not only to individual products but also to strategies for the management or treatment of a condition. Accordingly, these *Guidelines* continue to support the information needs of this broader audience.

CADTH and the Health Economics Working Group have endeavoured to reflect current best practices, but as the science and methods continue to evolve, so too will the recommendations. Areas that lack current consensus and therefore provide opportunities for further research have been identified within the *Guidelines*. Hence, these *Guidelines* seek not only to describe recommended practices for the "doers," but also to guide the future research and methods development that will contribute to both the advancement of knowledge and the sound foundation upon which efficient health care decisions can be made.

Highlights of the Fourth 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 elements of the guidance the researcher should follow are provided at the front of the economic guidelines. Additional resources are provided in the Appendices and online.

Decision problem: Decision-makers must have information that is relevant to the circumstances under which the decision is to be made. 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.

Reference case: The reference case describes a set of recommended methods that a researcher should follow for all evaluations, to increase the comparability of results across evaluations. The purpose of the reference case is to aid decision-making by promoting uniformity and transparency, and enabling the comparison of results for different technologies and different decisions. There are instances where the analysis required to address the decision problem will differ from the recommended reference case analysis. Any non-reference case analyses should be justified and presented in addition to the reference case analysis.

Assessment of data sources: Data sources informing parameter estimates used in an economic evaluation should be assessed based on fitness for purpose (relevance to the decision problem), credibility (perceived lack of bias), and consistency with data used elsewhere in the model.

Flexibility: Although a primary objective of the economic guidelines is to encourage the use of comparable approaches for analyzing and reporting across all evaluations, it is recognized that the reference case recommendations may not be optimal in every situation. As a result, the researcher has the flexibility to undertake a non-reference case analysis in order to address the specific circumstances surrounding the evaluation. Some sections in the economic guidelines provide advice for the researcher to consider when no consensus on methodological issues has been established. A key concern is whether using alternative approaches reduces the quality of the information provided by the evaluation. Researchers should state whether the methods used in their evaluation are consistent with the recommended reference case, and justify any deviations.

Transparency: A key concept in the economic guidelines is the need for transparency in the reporting of an evaluation. Researchers 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 issues 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 the cost-effectiveness results. A standard reporting format has been included in Appendix 1 for researchers to use, to ensure thorough and consistent reporting.

The fourth edition of the *Guidelines* can be downloaded from CADTH's website (**www.cadth.ca**).

Guideline Statements

1. Decision Problem

1.1.　The decision problem addressed by the economic evaluation should be clearly stated.

1.2.　The decision problem statement should provide a comprehensive specification of the interventions to be compared, the setting(s) in which they are to be compared, the perspective of the evaluation, which costs and outcomes are to be considered, the time horizon, and the target population for the evaluation.

1.3.　The interaction among components of the decision problem requires a discrete decision problem statement to be provided for each perspective and for each proposed analysis relating to a distinct subgroup.

2. Types of Evaluations

2.1.　In the reference case, the economic evaluation should be a cost-utility analysis (CUA) with outcomes expressed as quality-adjusted life-years (QALYs). Any departure from this approach should be clearly justified.

2.2.　A cost-effectiveness analysis (CEA) with outcomes expressed in natural units is not an appropriate reference case. If convincing evidence is available to show that important patient outcomes are equivalent on virtually all measures, except for survival or quality of life, then a CUA remains the appropriate approach. This allows for the necessary comparison, using the same benefit metric, across all the technologies being considered.

2.3.　A cost-minimization analysis (CMA) is a costing exercise and not a formal economic evaluation. As such, a CMA is not an appropriate reference case analysis. A CUA remains the appropriate approach, even where convincing evidence is available to show that important outcomes are similar, as it allows for the analysis of the uncertainty in incremental effect (through probabilistic analysis), facilitating the necessary comparison across all technologies.

2.4.　A cost-consequence analysis (CCA) should be viewed as a complement to, and not a substitute for, a CUA. A CCA aids in the transparency of the reporting of an economic evaluation, as disaggregated results are presented in terms of costs and outcomes (e.g., events predicted, survival, gains in quality of life).

2.5.　Where there are important health outcomes from a technology that cannot be captured in a CUA, then these should be reported as additional components within a CCA. If such outcomes can be valued in monetary terms then, additionally, a cost-benefit analysis (CBA) can be undertaken as a non-reference case analysis, with full details provided on the derivation of monetary values for all outcomes included in the evaluation, or justification for why the outcomes were excluded.

3. Target Population

3.1.   In the reference case, the target population(s) for the intervention and its expected use should be specified, and should be consistent with the decision problem.

3.2.   Factors that may lead to different estimates of costs and outcomes associated with interventions across distinct subgroups of the population should be specified. These could be factors that affect the natural history of disease, the effectiveness of treatments, or the utilities or costs associated with the disease or treatments.

3.3.   A stratified analysis with results presented for each subgroup should be provided in the reference case if factors are identified to support the consideration of distinct subgroups. Otherwise, the analysis should be for the entire target population.

4. Comparators

4.1.   In the reference case, "current care" (i.e., the intervention[s] presently used in a Canadian context) should be considered. In many cases, this may include more than one relevant comparator.

4.2.   The choice of comparator(s) should be related to the scope of the decision problem. As such, the comparators should reflect the target population of interest and the jurisdiction for which the decision is being made.

5. Perspective

5.1.   In the reference case, the perspective should be that of the publicly funded health care payer. The perspective of the economic evaluation should be related to the decision problem.

5.2.   Both costs and outcomes should be consistent with the stated perspective.

5.3.   Where perspectives other than the reference case perspective are of interest to the decision-maker and could have a substantial impact on the results of the analysis, these should be included as additional non-reference case analyses. For these analyses:

5.3.1   Report the results separately from the reference case.

5.3.2   Clearly identify the costs and outcomes that comprise the additional perspectives and quantify and describe the impact (i.e., magnitude) of the different components on the results of the analysis compared with the reference case.

6. Time Horizon

6.1.   In the reference case, the time horizon should be long enough to capture all relevant differences in the future costs and outcomes associated with the interventions being compared. Thus, the time horizon should be based on the condition and the likely impact of the intervention.

6.2.   The time horizon of the evaluation should relate directly to the decision problem.

7. Discounting

7.1. In the reference case, costs and outcomes that occur beyond one year should be discounted to present values at a rate of 1.5% per year.

7.2. The impact of uncertainty in the discount rate should be assessed by comparing the results of the reference case to those from non-reference case analyses, using discount rates of 0% and 3% per year.

8. Modelling

8.1. Model conceptualization and development should address the decision problem.

8.2. The model should be consistent with the current understanding of the clinical or care pathway for the health condition and the interventions being compared. The scope, structure, and assumptions should be clearly described and justified.

8.3. Researchers should consider any existing well-constructed and validated models that appropriately capture the clinical or care pathway for the condition of interest when conceptualizing their model.

8.4. The choice of modelling technique should be justified. The approach should be no more complex than is necessary to address the decision problem.

8.5. Baseline natural history should be representative of the target population considered in the decision problem.

8.6. The model should be validated, including an assessment of the face validity of the model structure, assumptions, data, and results.

8.7. Models should be subjected to rigorous internal validation. This process should involve quality assurance for all mathematical calculations and parameter estimates, and these processes and their results should be reported. Models should also be evaluated for external validity.

9. Effectiveness

9.1. A comprehensive search of the available data sources should be conducted to inform the estimates of effectiveness and harms associated with the interventions. Report the included studies and methods used to select or combine the data.

9.2. The data sources should be assessed based on their fitness for purpose, credibility, and consistency. Describe the trade-offs among these criteria and provide justification for the selected source(s). Incorporate the potential impact of the trade-offs in the reference case probabilistic analysis or using scenario analysis.

9.3. Researchers should evaluate and justify the validity of any surrogate end points used for parameter estimation. Uncertainty in the association of the surrogate to the final clinical outcome should be reflected in the reference case probabilistic analysis. This uncertainty can also be explored through appropriate scenario analyses. The existence of multiple potential surrogates should be reflected in the analysis of uncertainty. When considering the use of biomarkers as surrogate end points, the researcher should evaluate and justify the validity of the biomarker and the degree to which the biomarker satisfies the criteria of a surrogate end point.

9.4. Appropriate methods for extrapolating estimated effectiveness parameters to longer-term effects should be adopted. Uncertainty in the extrapolated estimates can be considered in the reference case through a probabilistic analysis that incorporates the correlation around the parameters within the survival function. Scenario analysis exploring structural uncertainty should also be conducted.

10. Measurement and Valuation of Health

10.1. In the reference case, the QALY should be used as the method for capturing the value of the effect of an intervention.

10.2. Health preferences (i.e., utilities) should reflect the health states in the model and be conceptualized to address the decision problem.

10.3. Health preferences should reflect the general Canadian population.

10.4. In the reference case, researchers should use health preferences obtained from an indirect method of measurement that is based on a generic classification system (e.g., EuroQol 5-Dimensions questionnaire [EQ-5D], Health Utilities Index [HUI], Short Form 6-Dimensions [SF-6D]). Researchers must justify where an indirect method is not used.

10.5. The selection of data sources for health state utility values should be based on their fitness for purpose, credibility, and consistency. Describe the trade-offs among these criteria and provide justification for the selected sources.

11. Resource Use and Costs

11.1. In the reference case, researchers should systematically identify, measure, value, and report all relevant resources based on the perspective of the publicly funded health care payer. When a range of perspectives is relevant to the decision problem, researchers should classify resources and their associated costs in categories according to each perspective, reporting results separately for the reference case perspective and any additional non-reference case perspectives.

11.2. Resource use and costs should be based on Canadian sources and reflect the jurisdiction(s) of interest (as specified in the decision problem).

11.3. Where substantial variation exists in practice patterns or costs among or within the jurisdiction(s) of interest specified in the decision problem, the researcher should consider these sources of variation when conducting the evaluation.

11.4. When valuing resources, researchers should select data sources that most closely reflect the opportunity cost, given the perspective of the analysis. Fees and prices listed in schedules and formularies of Canadian ministries of health are recommended as unit-cost measures when considering the perspective of the public payer, as long as they reflect actual payments. In other instances, total average costs (including capital and allocated overhead costs) may be relevant. Where costs are directly calculated or imputed, they should reflect the full economic cost borne by the decision-maker.

11.5. When a broader societal perspective is of interest to the decision-maker, the impact of the intervention on time lost from paid and unpaid work by both patients and informal caregivers as a result of illness, treatment, disability, or premature death should be included in an additional non-reference case analysis.

12. Analysis

12.1. In the reference case, the expected values of costs and outcomes (as defined by the publicly funded health care payer perspective) for each intervention should be estimated.

12.2. The economic evaluation should be assessed based on the incremental cost-effectiveness ratio (ICER). Estimates of net monetary benefit may also be provided.

12.3. For analyses with more than two interventions, a sequential analysis of cost-effectiveness should be conducted following standard rules for estimating ICERs, including the exclusion of dominated interventions.

12.4. In the reference case, expected values of costs and outcomes should be derived through probabilistic analysis, whereby all uncertain parameters are defined probabilistically:

12.4.1 In most cases, the probabilistic analysis will take the form of a Monte Carlo simulation.

12.4.2 An appropriate form of probability distribution should be employed that is based on standard rules that reflect the nature of each variable.

12.4.3 Correlation among parameters should be incorporated, as it can affect both expected values and their degree of uncertainty.

13. Uncertainty

13.1. In the reference case, uncertainty regarding the value of each parameter should be examined through probabilistic analysis.

13.2. Methodological uncertainty should be explored by comparing the reference case results to those from a non-reference case analysis that deviates from the recommended methods in order to examine the impact of methodological differences.

13.3. The impact of uncertainty on the estimated costs and outcomes for each intervention should be presented using cost-effectiveness acceptability curves (CEACs) and cost-effectiveness acceptability frontiers (CEAFs).

13.4. When the decision problem includes the option of commissioning or conducting future research, value-of-information analysis may be helpful to characterize the value of these options and design future research and should be included in the reference case analysis.

13.5. Structural uncertainty should be addressed using scenario analysis. Probabilistic analyses should be presented for each scenario.

14. Equity

14.1. In the reference case, all outcomes should be weighted equally, regardless of the characteristics of people receiving, or affected by, the intervention in question.

14.2. In support of a social decision-making viewpoint, a full description of the relevant populations should be provided, to allow for subsequent consideration of any distributional or equity-related policy concerns by the decision-maker. Researchers should approach any equity concerns by acknowledging the potential implications of both horizontal equity (equal treatment of equals) and vertical equity (unequal treatment of unequals).

14.3. Any stratified analysis of subgroups motivated by vertical equity considerations should be defined in the decision problem and, as such, fully explained, justified, and reported. When justifying stratified analyses, particular attention should be paid to respecting horizontal equity associated with any proposed vertical equity positions.

15. Reporting

15.1. The economic evaluation should be reported in a transparent and detailed manner with enough information to enable the reader or user (e.g., decision-maker) to critically assess the evaluation. Use a well-structured reporting format (Appendix 1).

15.2. A summary of the evaluation written in non-technical language should be included.

15.3. Results of the economic evaluation should be presented in graphical or visual form, in addition to tabular presentation.

15.4. Details and/or documents describing quality assurance processes and results for the economic evaluation should be provided. An electronic copy of the model should be made available for review with accompanying documentation in adequate detail to facilitate understanding of the model, what it does, and how it works.

15.5. Funding and reporting relationships for the evaluation should be described, and any conflicts of interest disclosed.

Guidelines in Detail

1. Decision Problem

Economic evaluations are designed to inform decisions. As such, they are distinct from conventional research activities, which are designed to test hypotheses. A comprehensive account of the decision problem that the economic evaluation will address is a necessary prerequisite for designing an appropriate economic evaluation.

Specifying a decision problem entails identifying the perspective from which the problem is to be addressed, and specifying the interventions (such as drug treatments, surgical procedures, diagnostic tests) to be compared, as well as the measures (e.g., costs, outcomes) that will be used to compare them.

In the context of economic evaluations, the measures that are used to evaluate the interventions are "costs" and "outcomes." The costs and outcomes to be included in the evaluation depend on the perspective adopted by the decision-maker whom the evaluation is intended to inform. The perspective for the evaluation should be specified by the intended decision-maker(s). When an economic evaluation is undertaken to inform multiple decision-makers, analyses from a series of increasingly inclusive perspectives may be appropriate. Detailed consideration of the appropriate perspective is provided in the Perspective section.

The specification of the interventions being compared needs to be comprehensive, providing clarity as to what will be provided, to whom and by whom, in what setting, and for what purpose. A health care intervention may be more than just the technology for which reimbursement is sought; variation in the other health care components associated with the technology can and frequently will affect both costs and outcomes. The choice of interventions for comparison should reflect the variety of interventions that are relevant to the decision problem. For example, if the decision problem relates to reimbursement of a new technology, then comparators should include interventions that could be substituted for the new technology. When the decision problem relates more generally to the most efficient practice with respect to a specific population, comparators should include all currently available, relevant therapeutic options. In addition, when there is reason to believe that current technologies are of poor or uncertain value compared with best supportive care, best supportive care should be considered for inclusion as a comparator. This will highlight to decision-makers whether a technology is benefiting inappropriately from being compared with a historically adopted technology of poor value. The statement of the purpose of the intervention should be sufficiently detailed to allow researchers to identify which outcome measures need to be captured in the economic evaluation.

The decision problem should support the clear identification of the costs to be included in any analysis. While costs incurred by the decision-maker should always be included, the decision-maker's objectives may lead it to consider a broader range of costs (accruing to other budgets). These should be presented in disaggregated form and could include out-of-pocket costs incurred by the patient (or household), costs falling on health care budgets outside of the decision-maker's area of responsibility, costs falling on non–health care public budgets (such as social services, education, or judicial systems), and costs falling on private-sector budgets (such as employment costs). The decision problem should identify the costs to be included in the analysis, and this should be consistent with the stated perspective.

For both costs and outcomes, the decision problem must specify the time horizon over which they will be considered, as the estimated value of an intervention is likely to be sensitive to the time horizon. This is of particular importance to interventions such as preventive treatments, like vaccines, that continue to produce desired outcomes in the distant future. This may also be important if long-term harms (e.g., adverse events) are considered an important aspect of a technology.

Within the target population, the value of an intervention will often differ among subgroups. Whether and how the value of the intervention varies by subgroup may be of interest to the decision-maker. Where this is relevant, a decision problem should be specified for each subgroup. This will ensure that any differences in the other components of the decision problem are appropriately captured, or confirm the assumption that all other factors in the decision problem are unaffected.

It is good practice to specify the decision problem in consultation with clinicians, members of the target population, and the decision-maker(s) to ensure that all relevant comparators are included; the most relevant outcomes for each stakeholder are taken into account; and the assessment is founded on a thorough understanding of all available evidence.

2. Types of Evaluations

Common forms of economic analysis are:

a) Cost-utility analysis (CUA): outcomes expressed as quality-adjusted life years (QALYs).
b) Cost-effectiveness analysis (CEA): outcomes expressed in natural units (e.g., life-years gained, lives saved, or clinical event avoided or achieved).
c) Cost-minimization analysis (CMA): interventions being compared are considered equivalent in terms of all relevant outcomes.
d) Cost-consequence analysis (CCA): costs and outcomes are presented in disaggregated form.
e) Cost-benefit analysis (CBA): outcomes expressed in monetary terms.

A CUA is the recommended type of economic evaluation and should be used in the reference case analysis. The use of a generic outcome measure allows decision-makers to make broad comparisons across different conditions and interventions. This feature facilitates the allocation of resources based on maximizing health gains.

A CUA is not without limitations. For instance, the methods and instruments for measuring health-related quality of life (HRQoL) and/or preferences can produce very different utilities for the same health state.[7] While potential limitations may speak to the need for further methodological advances, CUA remains the recommended reference case approach.

A CEA refers to an 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 disadvantage of a CEA is that the results can be compared only with the results of other evaluations that have used the same outcome measure; it does not facilitate the broad comparison of technologies and the allocation of resources across different conditions. Furthermore, a CEA, by definition, offers only a partial description of the benefit profile of

an intervention and is likely to omit some important aspects (e.g., preferences for clinical outcomes). Even if convincing evidence is available to show that the important outcomes are equivalent on virtually all measures, except for survival or quality of life, the CUA remains the appropriate approach. This allows fair comparison to be made across all the technologies being considered by using the same benefit measure. Results in terms of a CEA may be reported in addition to those of the CUA.

In a CMA, the interventions being compared are considered equivalent in terms of all important outcomes; thus, the lowest-cost intervention is considered the preferred technology. A CMA can be regarded as an extension of a CUA or CEA where the outcomes are demonstrated to be equivalent; therefore, only the costs associated with the interventions are compared. The critical issue with the use of a CMA is that it does not facilitate consideration of the uncertainty with respect to whether there are differences among the interventions in terms of important outcomes (including adverse events). Uncertainty in differences in outcome measures should be considered through probabilistic analysis within a CUA, again facilitating the necessary comparison across all interventions.

In a CCA, the costs and outcomes of the interventions are listed separately in a disaggregated format (e.g., intervention costs, hospital costs, clinical outcomes, and adverse events). A CCA has particular value in aiding transparency and can be used to present results for analyses conducted for different perspectives. This type of evaluation can also be useful for understanding the wider implications of an intervention. It may be helpful to present results in this manner for interventions involving public health, interventions with implications outside of health (e.g., crime, social services, education) as well as interventions with implications for informal caregivers. A CCA should be viewed as a complement to, and not a substitute for, a CUA.

In a CBA, costs and outcomes are valued in monetary terms, and the values are usually obtained through a willingness-to-pay approach, such as contingent valuation or discrete choice experiments. The difficulties with CBAs in a health context relate primarily to the challenges of measuring health outcomes in monetary terms and the ethical concerns associated with resource allocation decisions driven by willingness-to-pay data.[8] If there are important benefits from an intervention that are not captured in the CUA, then these should be reported separately. If such benefits can be valued in monetary terms, then a CBA might also be undertaken as a non-reference case analysis. Full details would need to be provided on the derivation of monetary values, which, in keeping with the social decision-making viewpoint adopted in these *Guidelines*, should be determined by the socially legitimate higher authority that funds the health care system and, consequently, be reflective of the monetary values of society.

3. Target Population

The cost-effectiveness of an intervention depends on the population for which it is being evaluated. The decision problem for the study should specify the target population(s) for which the interventions are to be used. Where applicable, this could include a description of the patient population for which the treatment is approved by Health Canada, as well as any anticipated off-label use of the product. The target population may include patients as well as their informal caregivers (i.e., unpaid caregivers), or the evaluation may be focused on the impact of an intervention (e.g., respite care) on informal caregivers. In other broader, population-based interventions (e.g., immunization programs), the target population may be the Canadian population at large.

Based on the target population specified in the decision problem, researchers should consider any potential spillover impacts beyond those individuals for whom the interventions are being targeted. For example, an intervention aimed at patients may have spillover impacts on informal caregivers due to changes in the level of care required by patients. Depending on the target population(s) specified in the decision problem, any associated spillover beyond the targeted population(s), in terms of either costs or effects, should be addressed in a non-reference case analysis.

A description of the target population should be detailed and include pertinent information on factors such as the health condition and its severity (e.g., not just angina, but Canadian Cardiovascular Society grading of angina pectoris), the distribution of comorbid illnesses likely to be present in the population, and the age and gender distribution of the population.

The economic evaluation should reflect the entire target population as defined by the decision problem. Researchers should, however, examine any potential sources of heterogeneity that may lead to differences in parameter-input values across distinct subgroups.[9,10] Note that heterogeneity may result from differences in the natural history of the disease, effectiveness of the interventions, health state preferences, or costs of the interventions. Heterogeneity may result in different decisions with respect to cost-effectiveness among different subgroups. The responsibility of the researcher, therefore, is to establish whether important heterogeneity exits in parameter estimates. A stratified analysis will allow decision-makers to identify any differential results across subgroups.

A stratified analysis requires the population to be parsed into smaller, more homogeneous subgroups, with an analysis conducted for each distinct subgroup.[11,12] Subgroups may be defined by baseline demographics (e.g., age, gender, socioeconomic status), disease severity, disease stage, comorbidities, risk factors, treatment-related factors (e.g., community or hospital setting), geographic location, usual adherence rates, or typical patterns of treatment. As far as available data allow, subgroup analyses should be based on mutually exclusive categories that combine all characteristics found to be heterogeneous with respect to a given parameter. For example, if disease severity depends on whether a patient has diabetes (yes/no) and is a smoker (yes/no), there would be four mutually exclusive subgroups. When data are unavailable, or considered unreliable, estimates for the broader target population should be used, although any limitations of this approach should be stated.

A stratified analysis requires the use of parameter estimates pertinent to each subgroup under consideration. Researchers must consider whether the data provide robust evidence of differences across subgroups. This is especially important in order to avoid the potential for post-hoc data dredging. To avoid this, the assessment of the face validity of differences in data inputs based on specific characteristics is recommended (further discussed in the Modelling section).

When a stratified analysis is conducted, but a decision-maker cannot implement decisions by subgroups, rather than calculating the mean result (i.e., the ICER) over the entire population, the appropriate estimate of the overall result is determined by weighting the estimates for each subgroup by their respective prevalence.

4. Comparators

Economic evaluations involve the comparison of two or more interventions; these interventions are referred to as comparators. This section describes how to identify, select, and describe the comparators within an economic evaluation.

Identification of Comparators

The comparators should be directly related to and stated as part of the decision problem. Comparators should be identified based on the components defined by the decision problem. It is crucial to identify all 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. All interventions currently used and potentially displaced should be identified, in addition to interventions likely to be available in the near future. Researchers should consider the complete clinical or care pathway for the condition as it relates to the decision problem and a broad range of possible comparators, including individual interventions as well as management or intervention strategies. The identification of comparators should not be limited to a specific type or class of interventions (e.g., when considering a biopharmaceutical, determine whether a non-pharmaceutical intervention such as medical management might be currently among the options used in the management of patients).

The inclusion of best supportive care should be assessed for its appropriateness as a comparator where there is reason to believe that current technologies are of poor or uncertain value in comparison with best supportive care. This will allow decision-makers to note whether a technology appears more cost-effective as a result of being compared with a historically accepted technology of poor value.

Selection of Comparators

Based on a comprehensive list of identified comparators, a starting point for selecting the appropriate comparators for the analysis is to determine what represents current care, or which technologies are likely to be displaced by the intervention(s) under investigation. These should be technologies that the decision-maker is currently funding and are commonly used. In addition, consideration of best supportive care should be assessed when new technologies have not been fully adopted by the decision-maker(s), or newer technologies represent uncertain (or poor) value. The selection of comparators should be conceptually driven and should not be determined by the availability of data. Justification for the chosen comparators should be provided.

In some cases, comparators may be management strategies (e.g., codependent technologies, adjunctive use of interventions — given to maximize the effectiveness of the primary therapy) rather than individual interventions (i.e., a single drug or device). When dealing with management strategies, researchers should ensure that uncertainty in the data informing all parts of the strategy is appropriately characterized. For example, if the management strategy consists of both a test and a treatment component, researchers should account for the costs and effects of both false-positive and false-negative test results. Where interventions are used concomitantly, consideration should be given to the possible combinations. Where interventions are used sequentially (e.g., as a result of treatment failure or intolerance, testing algorithms), consideration should be given to the sequence, as the results of the analysis may be sensitive to alternative sequence pathways.[13] Where treatment is based on events along the pathway, consideration should be given to the sequence of events. Uncertainty related to the choice and sequence of comparators should be explored using scenario analyses.

Selecting comparators may be complicated when there are a number of appropriate comparators identified, or where comparators vary among the jurisdictions for which the analysis is being conducted. In these cases, all comparators should be considered and the choice to remove any comparator(s) from the decision problem clearly justified. Should there be a paucity of clinical evidence to support the evaluation of comparators, this should be addressed in scenario analyses or discussed with respect to potential implications for decision-making.

Description of Comparators

Comparators should be clearly described to allow for the identification of all relevant costs and outcomes. This should include a description of how interventions differ (e.g., dosing, route and frequency of administration, use in combination with other interventions, use in sequence with other interventions, placement along the clinical or care pathway, and any relevant starting and/or stopping rules).

Where comparators are management strategies (e.g., codependent technologies, adjunctive use of interventions), the researcher must distinguish among situations where the intervention 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. Strategies should be explained (e.g., when, under what circumstances, and for whom), and the elements of the alternative strategies defined.

5. Perspective

The perspective chosen for the economic evaluation should be directly related to, and stated as part of, the decision problem. The costs and outcomes included in the analysis would then be defined by the selected perspective, as detailed in Table 1. In the reference case, the publicly funded health care payer perspective should be adopted (see Appendix 2) and the included costs should be those incurred by the Canadian public payer, and the included outcomes should reflect all meaningful health effects for patients and their informal caregivers.

Having identified the target population, any spillover costs or effects that fall outside the target population should be incorporated in a non-reference case analysis (see Target Population). That is, if patients are the target population, then the impact of including any meaningful health effects for informal caregivers would be assessed using a non-reference case analysis. When incorporating costs or effects associated with either patients or informal caregivers, researchers should be mindful of the potential for double-counting and avoid situations where the same elements are valued as part of both the incremental costs and outcomes (e.g., time sacrifices being considered in changes in health outcomes).[14,15]

Any relevant non-health effects for either patients or informal caregivers would fall outside the perspective of the publicly funded health care payer and should be examined in non-reference case analyses. When exploring a payer perspective, the researcher should clearly define the payer, and the relevant costs and outcomes that will be included in the analysis.

Where the perspective is that of the private payer, particular attention should be placed on determining any proportion of health care services that may be covered by the public payer. For example, a standard hospital stay may be covered by the public payer, but an upgrade to a private room may be paid for by the private payer; or if individuals with public drug plan coverage have to pay an annual deductible prior to receiving drug coverage, this deductible may be covered by private payers in cases in which individuals are covered by both private and public plans. Private insurers may also cover a number of services not paid for by the public payer

Table 1: Examples of Different Costs and Outcomes by Perspective

	Reference Case	Non-Reference Case			Examples
	Public Health Care Payer[a]	Private Payer[b]	Broader Government Payer	Societal	
Types of Costs					
Costs to publicly funded health care payer	✓		✓	✓	○ Drugs, medical devices, procedures ○ Equipment, facilities, overhead ○ Health care providers ○ Hospital services ○ Diagnostic, investigational, and screening services ○ Informal caregivers' health care costs ○ Rehabilitation in a facility or at home[c] ○ Community-based services, such as home care, social support[c] ○ Long-term care in nursing homes[c]
Costs to private insurer		✓		✓	○ Drugs, medical devices (falling outside of public payer) ○ Aids and appliances ○ Alternative care (e.g., chiropractic services, massage therapy, homeopathy) ○ Rehabilitation in a facility or at homec ○ Community-based services, such as home care, social support[c] ○ Long-term care in nursing homes[c]
Costs to government payer (beyond health care)			✓	✓	○ Social services, such as home help, meals on wheels[c] ○ Affordable housing ○ Education
Costs to patients and informal caregivers				✓	○ Out-of-pocket payments (e.g., copayments for drugs, dental, assistive devices) ○ Cost of travel, paid caregivers ○ Premiums paid to private insurers ○ Patient's time spent for travel and receiving treatment
Productivity costs				✓	○ Lost productivity due to reduced working capacity, or short-term or long-term absence from work ○ Lost time at unpaid work (e.g., housework) by patient and family caring for the patient ○ Costs to employer to hire and train replacement worker
Types of Outcomes					
Health effects relevant to patients and informal caregivers	✓	✓	✓	✓	○ Health-related quality of life ○ Life-years gained ○ Clinical morbidity
Non-health effects relevant to patients and informal caregivers			✓	✓	○ Information available to patients ○ Reduction in criminal behaviour ○ Better educational achievements

[a] Any spillover impacts should be handled in a non-reference case analysis.

[b] Researchers should consult the private payer to determine costs and outcomes of relevance.

[c] Some of these costs may be incurred by the publicly funded health care payer, depending on the precise nature of these costs and the relevant jurisdiction.

(e.g., alternative health services, such as chiropractic services, acupuncture, massage therapy, and dental services). Where the perspective is that of the broader government payer, in which not only health services are considered, the costs and effects of government programs and services beyond health care, such as affordable housing or education, may be relevant.

Where there are multiple decision-makers and, therefore, multiple decision problems, the perspective should be presented for each decision problem, which may require considering multiple perspectives. This may occur where the intervention(s) under investigation are paid for by different payers or a combination of payers.

Perspectives other than that considered in the reference case analysis may also be of interest if they are expected to have an impact on the results. The alternative perspectives and their relevance to the decision problem should be noted and should be included in additional non-reference case analyses. The costs and outcomes associated with different perspectives should be reported separately. Where quantification is difficult, the likely magnitude of such costs and outcomes and their impact on the results of the analysis should be discussed and disaggregated results may be presented as part of the analysis. An example of a case in which other perspectives may be considered is when the decision problem is from the perspective of the publicly funded health care payer, but the intervention permits patients to return to work sooner, which may shift costs away from patients and their informal caregivers. In such cases, a societal perspective may be evaluated in a non-reference case analysis that allows for the full consideration of all costs and outcomes associated with the evaluation of the intervention.

6. Time Horizon

In the reference case, the time horizon should be long enough to capture all potential differences in costs and outcomes associated with the interventions being compared.[16,17] The same time horizon must be applied to costs and outcomes for analytical consistency.

The time horizon of the analysis should be conceptually driven, based on the natural history of the condition or anticipated impact of the intervention (e.g., public health promotion), and must reflect all states of the health condition. A longer-term analysis allows for the exploration of uncertainty; this does not, however, imply that primary data must be collected from patients or affected populations over such a period. When modelling chronic conditions, or when the interventions have differential effects on mortality, a lifetime horizon is most appropriate. For decision problems involving the dynamic evolution of the target population (i.e., individuals enter and exit the population over time), the time horizon may extend beyond the lifetime of a single cohort and should relate to the maximum expected lifetime of future patients (e.g., vaccination programs). Shorter time horizons might be considered where there are no meaningful differences in the long-term costs and outcomes of interventions (e.g., convergence of clinical pathways for the remainder of patients' lifetimes), or the condition affects the individual only over a defined period (e.g., acute illnesses). In these cases, justification should be provided for the duration of the time horizon selected.

In some cases, multiple time horizons might be appropriate to consider how the cost-effectiveness of interventions differs in various phases of the condition, as well as in the overall condition. When there is uncertainty in the choice of time horizon, the implications of this should be assessed by comparing the results based on the time horizon used in the reference and non-reference case analyses. This is of special relevance in instances when the majority of QALY gains from therapy occur long after treatment has been curtailed.

7. Discounting

Economic evaluations that involve costs and outcomes that occur beyond one year require the application of a discount rate that reflects society's preferences over time (i.e., the social discount rate). Economic efficiency necessitates that the social discount rate measure the marginal social opportunity cost of resources allocated to government investment, which may be approximated by the real rate of interest on government bonds.[3,18]

In keeping with the social decision-making viewpoint adopted in these *Guidelines*, this rate should reflect the real rate of interest on government bonds faced by the higher authority (i.e., the level of government) that funds the health care system. In Canada, health care is funded primarily by the provincial, territorial, and federal governments. Therefore, both provincial and federal government bonds may be considered as sources for estimating the real rate of interest. Taking into account the proportion of public health care spending by the provinces and territories relative to the federal government[19,20] and the observed uniformity between the historical returns of provincial and Canadian federal bonds,[21] the recommended discount rate for the reference case is based on provincial bond rates.[22]

It is therefore recommended that costs occurring beyond one year be discounted using the real rate of interest on provincial government bonds. Assuming that decision-makers are likely to face exogenous budget constraints,[6] outcomes should be discounted using the same real rate of interest on provincial bonds, minus the growth rate of the cost-effectiveness threshold (i.e., the estimated health expected to be forgone as a result of any new costs that must be accommodated within a budget-constrained system).[3] In practice, however, given the uncertainty in the value of the cost-effectiveness threshold and how it is anticipated to change over time, these *Guidelines* recommend that researchers discount costs and outcomes at the same rate.

The recommended rate for the reference case is set at 1.5% per year for both costs and outcomes, reflecting recent empirical evidence on the long-term cost of borrowing for Canadian provinces.[3] The discount rate is expressed in real (i.e., constant, inflation-adjusted) terms, which is consistent with valuing resources in real dollars. Nominal provincial bond rates were adjusted for inflation using the Bank of Canada's target inflation rate (currently 2% per year),[23,24] and a weighted average of the real provincial bond rates was calculated based on the relative proportion of the population represented by each province.[25]

In principle, the appropriate discount rate to use depends upon the inter-temporal distribution of costs and outcomes over the time horizon of the analysis. In the absence of robust empirical evidence on this distribution, the Guideline recommendation is based on the real long-term cost of borrowing. The potential impact of applying non-constant discount rates (i.e., rates that vary according to when the costs and outcomes are accrued) may be investigated in a non-reference case analysis. When deciding whether to explore the use of non-constant discount rates, the researcher should be guided by the magnitude of the observed differences between short-term rates (i.e., less than 10 years) and long-term rates (i.e., greater than or equal to 10 years), as well as the degree of divergence in the timing of costs and outcomes for the interventions being compared. The smaller the differences, the less impact non-constant discounting will have on the results.[26]

In order to incorporate potential uncertainty and to assess the sensitivity of the results to changes in the discount rate, a non-reference case analysis using a rate of 0% should be conducted to show the impact of discounting. In addition, a 3% discount rate (double the recommended reference case rate) should be used as an upper bound. If desired, an additional non-reference case analysis based on a discount rate of 5% may be conducted to allow for comparison with previous analyses.

8. Modelling

This section is intended to identify the key considerations that should guide the model-development process. This is not intended to be a step-by-step guide to building a model; researchers should consult modelling guidelines and adhere to good modelling practices when constructing models for use in economic evaluations.[27-33]

Model Conceptualization

Modelling in the context of health economic evaluations provides an important framework for synthesizing available data and assumptions from multiple sources to generate estimates of the expected costs and outcomes for the interventions of interest. A decision-analytic model uses mathematical relationships to define a series of possible consequences that would result from the set of interventions being evaluated.[34]

The conceptualization of the model is a critical component of the model-development process. This process involves the development of a model structure that is defined by specific states or events and the relationships among them that together constitute the clinical or care pathway for the condition of interest and the interventions being compared. Depending on the condition and the interventions being modelled, the clinical or care pathway will vary and could include, for example, doctor's visits, hospitalizations, screening, or certain risk factors, as well as various health states. The conceptualization of the model should incorporate the potential for changes along the clinical or care pathway (e.g., reflecting changes in how individuals might progress or regress among health states), and the model should be structured in a way that can accommodate these changes.

The model structure must be clinically relevant, and close collaboration with and input from those able to provide clinical expert judgment is necessary. The model must be consistent with the current knowledge of the biology of the health condition, the causal relationships among variables that constitute the clinical or care pathway, and the expected effects of the interventions.[34] In addition, the model must be conceptualized to address the decision problem, including the setting(s) in which the interventions are to be compared, the perspective of the evaluation, the costs and outcomes to be considered, the jurisdiction of interest, and the time horizon of the evaluation.

The process of model conceptualization should not be dictated by the availability of data. In practice, however, the availability of data may constrain the options in model development, and the model structure will need to be revisited accordingly and modified within an iterative process. Where there are well-constructed, validated models that appropriately capture the clinical or care pathway and reflect all of the components of the decision problem, researchers should consider these models when trying to conceptualize their own model structure. This

could involve contacting and potentially collaborating with researchers who have developed previous models, or simply attempting to make similar structural assumptions, or, where they remain relevant, using similar natural history data to populate the model. Researchers' awareness of comparable models and their relative similarities, differences, strengths, and weaknesses should help to avoid duplication of efforts to address the same decision problem, and may provide a basis for model validation (as subsequently described).

Determining what constitutes an adequate level of detail in terms of the required states or events and the clinical or care pathway is one of the most difficult challenges a researcher faces when conceptualizing a decision model.[28] The model need not reflect every possible aspect associated with the health condition or pathway, but should be detailed enough to capture those factors that are likely to result in differences in costs or outcomes among the interventions being compared. That is, the model should be conceptualized in such a way that it provides a representation of reality that captures the elements and relationships that are essential to address the decision problem,[35] but should not be more complex than is required.

Modelling Techniques

There are many decision-modelling techniques available to researchers when conducting economic evaluations, including decision trees,[28] cohort-level state-transition models (i.e., Markov models),[29] patient-level state-transition models (i.e., microsimulations or first-order Monte Carlo simulations),[28,29] system dynamic models,[28,31,36] discrete event simulation models,[28,30,36] and agent-based models.[28,31,36] A number of papers provide guidance on the selection of appropriate modelling techniques.[28,35,37,38]

Most decision problems can be addressed with a wide variety of modelling techniques. The choice of model type should be related to the characteristics of the decision problem, with justification provided regarding the choice of modelling approach. For any type of modelling approach chosen, the model must be methodologically sound and transparent, and researchers are encouraged to follow good modelling practice guidelines.[27] When choosing among various techniques, the researcher should seek to identify a modelling technique that addresses the decision problem and that appropriately reflects the conceptualization of the clinical or care pathway for the health condition and the interventions being compared.[39]

Researchers should think about the decision problem and whether the model is intended to address a single question, on one occasion, or if the model will be used on an ongoing basis to address multiple questions.[35] Researchers should also consider requirements for addressing the decision problem, such as the potential heterogeneity or dynamic evolution of the target population.[40] Whether to model time as being continuous (i.e., using exact times) or discrete (i.e., using defined time intervals) would be another consideration. Researchers should also anticipate the likelihood of interaction among individuals, or the possible impact of the intervention on the spread of the condition. The decision problem may also require the ability to incorporate competition for constrained resources and the development of waiting lists or queues.

Incorporating Parameter Estimates

Regardless of the chosen modelling technique, data are required to inform the various states and events and the movement along the clinical or care pathway. The conceptualization and incorporation of data into the model will vary depending on how the model is specified for the different modelling techniques.

Natural History

Economic evaluations will often require estimates of both relative clinical effects of interventions (Effectiveness section) and information on the outcomes at baseline (i.e., the start of the analysis time horizon).[2,41] Outcomes at baseline reflect the natural history of the health condition, while the effects of interventions are typically reported relative to a comparator such as current care.[41,42] There are a number of ways of selecting parameter estimates for the baseline natural history, which will depend on the decision problem.[41,42]

Where the baseline outcomes and the relative effects data are estimated from the same source (e.g., studies in a network meta-analysis), the researcher should consider whether the data are sufficiently representative of the target population (see Target Population section) under current circumstances, and justify the fitness of the data for the purposes of the evaluation.[42] The baseline data, in particular, should be as specific as possible to the Canadian context. Consequently, where there exist baseline natural history data that are independent of the relative effects data (e.g., from administrative databases or observational cohort studies) and these data are more representative of the Canadian context than those used to estimate the relative effects, it may be reasonable to use only those data in the estimation of the baseline outcomes.[41,42]

In all cases, researchers should define what constitutes relevant data based on the decision problem and then seek to identify data sources using a comprehensive and transparent approach that can be replicated by others. The researcher should then determine whether a synthesis of the data sources is appropriate or possible. Where a synthesis cannot be conducted, the researcher should employ judgment in assessing the individual sources in terms of their relative fitness for purpose, credibility, and consistency in order to determine which data source represents the best trade-off among these criteria (see Effectiveness section).

When informing and incorporating parameter estimates for natural history, probability distributions should be derived and the associated uncertainty propagated through the model. Where more than one data source is judged to be appropriate, this should be reflected in the reference case probabilistic analysis or using scenario analyses that consider the alternative sources for the estimates. Where Canadian data are lacking, and there is reason to believe that the available data may differ substantially from what would likely be observed in the jurisdiction of interest, this should also be incorporated into the analysis of uncertainty by examining the impact of different scenarios. These scenarios should consider alternative estimates that may be more representative of the Canadian context. The sources of the alternative estimates (e.g., expert elicitation) must be described and justified.

Incomplete natural history data or inconclusive knowledge about the underlying condition will lead to structural uncertainty within the model. In such cases, the limitations of the natural history data should be acknowledged and addressed by building more than one plausible framework and examining the impact of the alternative model structures using scenario analysis. Where the results of the analyses vary in ways that are substantive enough to potentially have an impact on the cost-effectiveness, the reasons for these differences should be identified and critically examined and, on the basis of this, researchers should recommend a particular model structure (see Uncertainty section).

Incorporating Effectiveness Estimates

The incorporation into the model of parameter estimates for effectiveness, in addition to baseline natural history information, defines the movement along the clinical or care pathway of the model. Estimates of relative effect can be combined with baseline natural history data to derive the movement through the model for each intervention.[34,41] Researchers may also refer to Cooper et al.[42] regarding methods for deriving absolute estimates for effect probabilities based on combining estimates of relative effects and baseline data. Measures of effect are not limited to binary outcomes and can include other types of outcome scales (e.g., change in HRQoL). If the effectiveness of the intervention changes over time, this should also be captured in the model.

Incorporating Harms Estimates

Researchers should be explicit about how the adverse events included in the economic evaluation were identified, and what methods were used to incorporate them. Where adverse events have a negligible impact on health effects, or no impact on costs and resources, it is often appropriate to exclude these events from the model. Where adverse events are not included, a clear justification must be provided.

Adverse events should be incorporated into the model by combining both the health condition and the associated adverse effects. In the case of utilities, the utility for a specific health state can then be adjusted by applying a disutility for an adverse event to allow the utility for the health state with an adverse event to be estimated.[43]

If effects are transitory (i.e., short-term), they should be incorporated through appropriate refinement of the states or events within the model. Where data are available on the prevalence, costs, and disutility associated with each adverse event by intervention, this facilitates greater transparency.

Outcomes (and Health Costs) Not Attributable to the Intervention(s) or Condition of Interest

There is some debate regarding the relevance of including clinical outcomes and health-related costs that originate not with the intervention(s) or condition of interest, but may be realized as a result of the impact of the intervention on an individual's health. Specifically, interventions that extend life may result in individuals experiencing future clinical events or costs associated with aging or other health conditions that are incurred as a result of their lives being extended. Where health effects are projected over the individual's lifetime, costs should be similarly projected.[2] The modelling of projected outcomes and costs remains an area of ongoing research and debate.[44] In the meantime, researchers may consider a scenario analysis that incorporates modelling assumptions that project potential future health-related costs and effects that are not directly attributable to the intervention(s) or condition of interest, but are likely to occur over the time horizon.

Expert Elicitation

In the absence of sufficient data for informing parameter estimates, the elicitation of quantitative input from relevant experts may be useful. For the purposes of these *Guidelines*, the formal elicitation of quantitative input from relevant experts regarding the magnitude of a given parameter and its uncertainty (i.e., expert elicitation) is distinguished from obtaining expert judgment on issues such as the structure of the model, the face validity of the results, or the potential relevance of clinical events.

Expert elicitation involves key steps that researchers should implement and clearly describe. These include determining the information to be elicited, identifying a representative sample of experts, applying specific elicitation methods (reflecting uncertainty in parameter estimates between and within experts), and synthesizing the elicited information across experts. For details regarding these steps, researchers are referred elsewhere.[45,46] When identifying experts, researchers should justify their choice based on the parameter information being sought, the perspective of the analysis, and the demonstrated expertise of the individual.

It is recommended that researchers continue to focus on identifying appropriate data sources for informing parameter estimates and, as elicitation methods continue to evolve, consider expert elicitation as a potential source of data for filling in gaps in the available information.

Calibration

Model calibration methods involve the estimation of unknown model parameters by achieving agreement between the model outputs and other sources of data that are external to the model and not used to parameterize the model.[47,48] Calibration provides a method for estimating model parameters when available data are insufficient. Calibration should be distinguished from other sources of parameter estimation, which involve separate processes from the model itself and do not consider the overall similarity of the model outputs to the external data.

The use of calibration for informing parameter inputs is not routine and, as methods evolve, researchers should continue to focus on identifying appropriate data sources for informing parameter estimates, where calibration may be used to fill in gaps in the existing evidence base.

Validation

Models should be formally validated in order to judge their accuracy. Validation involves testing the model to confirm that it does what it is expected to do. Validation as a concept should apply to particular settings and not to the model itself;[33] therefore, researchers should assess model validity in the context of the specific setting being addressed in the decision problem.

Researchers should evaluate face validity during model conceptualization in order to ensure the validity of the model. The assessment of face validity is mostly qualitative in nature and is intended to assess whether the model structure, assumptions, and parameters accurately reflect the clinical or care pathway of interest and the potential impact of the interventions. This should be informed based on the expert judgment of those with content expertise,[33,49] and done early and iteratively throughout the analysis.

Formal internal validation of the model should be performed as a quality assurance measure for all mathematical calculations and parameter estimates.[33] This process should be undertaken by researchers who are not directly involved in model development. It should include testing the mathematical logic of the model and checking for errors. This involves verifying both the individual equations and their implementation in the code,[33] and ensuring that the parameters accurately reflect the sources used. Coding accuracy can be validated by techniques such as double programming (i.e., the independent programming of sections of a model by two programmers), testing extreme or zero values, and examining the results of scenarios that lead to known results.[33,50,51] Counterintuitive results of the internal validation may reflect either errors or new insights, and must be explored and explained.

If there are other models addressing the same or similar decision problems, the researcher should assess the degree of similarity among the model results (i.e., cross-validation).[33,52-56] However, the meaningfulness of this type of validation depends on the degree of similarity of the model structures, methods, and data sources. A high degree of dependency among models will reduce the value of cross-validation.[33]

External validation of the model, with actual data, should be undertaken to determine whether the model estimates are consistent with other reliable and preferably independent data sources (i.e., data not used to populate the model), such as natural history or mortality data. This is achieved by running the model to simulate outcomes and examining how well the model results correspond with the independent data sources.[33] The researcher should consider and provide explanations for any significant inconsistencies. Where there is discordance, researchers should justify which data source represents the best trade-off in terms of fitness for purpose, credibility, and consistency, and this data source should dictate the specification of the final model. If the objective of the analysis is to make projections about the future, the validity of the model and the associated projections should be assessed as data become available.

Validation processes should be documented and provided as part of the technical description of the model. These processes should be reported in sufficient detail to enable a clear understanding of how the model has been validated such that someone using or reviewing the model can have confidence in the results.

Transparency

Transparency and clarity of presentation are necessary to allow the model to be critically appraised and reproduced. Researchers should provide detailed information on how the model structure was developed. Diagrammatic representation of the model structure is highly recommended to facilitate the understanding of the model structure and function, and to highlight the most important assumptions.[29-31] A detailed description explaining the flow of data through the model is recommended.

Technical transparency requires documentation that describes the model in sufficient detail, including its structure and all components, to enable those with the necessary expertise and resources to reproduce the model.[33] Additionally, the model structure, including the assumptions and any subjective judgments, should be justified in sufficient detail so as to enable the decision-maker to evaluate its acceptability. The choice of data sources and of methods for analyzing data inputs, and the subsequent incorporation of data into the model can have a bearing on the results of the evaluation and therefore must be clearly stated and justified. All steps in the analysis should be presented in a disaggregated manner before being aggregated into the cost-effectiveness results. The researcher should also report any known limitations of the model.

9. Effectiveness

Efficacy refers to the performance of a health intervention under controlled, optimal circumstances, often in the context of randomized controlled trials (RCTs), while effectiveness refers to the performance of an intervention in the real world (i.e., routine clinical practice).[17] Decision-makers are generally concerned with the impact of interventions on patients treated in routine practice. In the reference case analysis, this would entail the need for clinically meaningful outcomes to inform the duration and quality of life.

A key issue is the extent to which the data obtained from an RCT reflect the effectiveness likely to be achieved in a real-world setting (i.e., the external validity of the trial). For the evaluation to be relevant to the decision-maker, the effects and costs in an economic evaluation should reflect the effectiveness of the intervention rather than its efficacy. When seeking to derive estimates of effectiveness from efficacy data researchers should identify the factors that may differ between the real-world setting and the trial environment (e.g., adherence, implementation). Critical to making a judgment about incorporating real-world factors into the analysis is the strength of the available data linking potential intervention effect-modifying factors with important patient outcomes. Researchers should present these linkages in a transparent manner and provide justification.

Health interventions should also be assessed in terms of the potential for harms (adverse events). It is important to consider the adverse events associated with the interventions being compared in the economic evaluation, as they may affect the expected costs and outcomes. In particular, depending on their nature, frequency, duration, and severity, adverse events may have an impact on patients' adherence, mortality, morbidity, preferences (utilities), or resource use. Researchers should focus on harms that are clinically meaningful and therefore likely to have an impact in routine clinical practice on the expected costs and effects of the interventions being compared.

Types of Parameter Estimates

Parameter estimates for effectiveness and harms can take different forms. In conceptualizing the clinical or care pathway (see Modelling section), economic models will often incorporate estimates of relative effects such as the relative risk, odds ratio, or risk difference. Relative estimates of effectiveness or harms can also be estimated in terms of hazard ratios based on time-to-event data and derived from parametric or non-parametric survival analyses.[42] The use of reported hazard ratios to inform parameter estimates in an economic evaluation often requires the assumption that the relative effect was constant over the duration of the study data. Researchers should consider whether this represents a reasonable approximation of the likely behaviour of the relative intervention effects in the context of their analysis. When assessing the reasonableness of this approximation, other variables that change with time and their possible impact on the hazard ratio should also be considered. Of critical importance is the degree to which any such simplification may be at odds with the real world, and the implications this may have in terms of producing potentially misleading results.

In situations where the estimation of a comparative effect is not practical, or where strong therapy preferences (i.e., lack of equipoise) render a comparative study unethical, researchers will have to rely on data measured in absolute terms (i.e., separate estimates for each intervention of interest) rather than relative terms. This could occur in cases where a researcher requires information on the risk of an adverse event, or the effectiveness of an intervention in the face of rapid and fatal disease progression.[57] Similarly, when an evaluation is focused on the acquisition of a new technology (e.g., magnetic resonance imaging machine, computed tomography scanner) where no comparative technology was previously available, the effect of interest may be expressed in absolute terms; for example, as the number of cases identified.

Data in absolute terms may also be presented as rates. The terms "probabilities" and "rates" are often used interchangeably. When seeking to inform probability estimates in a decision model, an instantaneous event rate must be converted to a probability over a specific period of time.[2,34,58] In the absence of data to establish this empirically, the conversion requires that

researchers are able to assume that the rate is constant over the time period of interest.[2,34] Probabilities may be used, for example, to describe how individuals transition among health states or when considering the accuracy of a diagnostic test (i.e., sensitivity, specificity).

Data Sources and Assessment

Potential sources for informing parameter estimates for effectiveness (e.g., clinical effects, detection, harms) could include RCTs, observational studies, administrative databases, non-comparative studies, or expert input. Based on the decision problem, the researcher should define the relevant data (e.g., reflecting the target population, comparators, and jurisdiction of interest) for the economic evaluation. A search for information on clinical effectiveness and harms should be conducted, as defined by the decision problem. The search should be undertaken in a way that is both comprehensive and can be replicated by others.

When assessing data sources, the researcher should consider the context of the evaluation. It is important that the data used to inform the parameters are relevant to the decision problem and consequently fit for purpose. Assessing fitness for purpose involves the examination of differences between the data source and the decision problem being addressed. This would involve consideration of issues such as the types of patients likely to receive a certain intervention or to experience a specific adverse event, as well as differences in the dose or delivery method of the intervention or differences in the standard of care being used. If a previous search or review has already been undertaken, its fitness for addressing the decision problem can be assessed in terms of its scope and inclusion and/or exclusion criteria. Similarly, the fitness for purpose of an existing meta-analysis or network meta-analysis should also be assessed in terms of its scope and inclusion and/or exclusion criteria and how they relate to the decision problem.[42]

The credibility of parameter estimates (defined, for the purposes of these *Guidelines*, as a perceived lack of bias, where bias refers to the systematic deviation of the estimated value from the true underlying value) must also be considered. When assessing credibility, researchers should determine whether the methodological rigour of the data source is such that it provides a believable estimate of the true value. This would involve consideration of the extent of baseline imbalances in patient characteristics between study arms, as well as other potential sources of bias such as performance and detection bias.

Another criterion that researchers should consider when assessing a data source is that of consistency with data used elsewhere in the model.[59] The assessment of consistency depends on the fitness for purpose and credibility of the data sources used to estimate the various parameters, but it also involves considerations such as how the data for a particular source are collected and measured relative to the other parameters.

Parameter Estimation

When estimating effectiveness and harms parameters for an economic evaluation, researchers should ideally synthesize data from all available sources using methods that take account of potential differences in the fitness for purpose, credibility, and consistency among the various sources,[60-64] thus reflecting the totality of available information. Probability distributions for effectiveness and harms should be derived, and the associated uncertainty propagated through the model (Uncertainty and Modelling sections).

The development of methodological approaches capable of synthesizing evidence from all available sources remains an area of ongoing research. In the absence of appropriate methods for combining all of the available data sources, researchers may instead focus on synthesizing data from sources that are assessed to be fit for purpose, credible, and consistent using meta-analytic or network meta-analytic methods. In the context of an economic evaluation designed to inform health care decision-making, the decision problem may involve more than two comparators, or two comparators that have never been compared directly; consequently, a network meta-analysis that allows for the simultaneous comparison of multiple interventions as well as the incorporation of data based on indirect comparisons may be preferred to the standard pairwise meta-analysis framework. In situations involving a relevant previously conducted review or synthesis, researchers may seek to update the existing data source with additional studies; as with other data sources, this too should be done based on a consideration of the additional studies' fitness for purpose, credibility, and consistency. This could involve combining new studies with the summary estimate from a prior synthesis, or including estimates from the individual studies to update the prior synthesis.

Where only a single data source is identified, or where concerns exist regarding differences among the data sources such that it would be inappropriate to synthesize the information, the individual source(s) should be assessed in terms of their respective fitness for purpose, credibility, and consistency with data used elsewhere in the model. When selecting a data source from among multiple sources, researchers should employ judgment regarding which data source represents the best trade-off among these criteria. For example, when informing decisions at the population level, comparing data from a non-randomized observational study or administrative database with data from an RCT featuring an idealized environment, truncated horizon, or restrictive inclusion and exclusion criteria may require that a trade-off be made in terms of the relative fitness for purpose and credibility of the respective data sources. Similarly, researchers may have concerns about the credibility of parameter estimates for effectiveness or harms from a non-comparative single arm study, although the characteristics of the patients in the study may be consistent with the sources used to estimate other parameters in the model. A trade-off between this study and a methodologically sound expert elicitation study of physician input that may be considered credible but not fit for purpose, if the decision problem requires estimates of patient-reported adverse events, may then be required. The judgment of which data source represents the best trade-off among the criteria of fitness for purpose, credibility, and consistency and the resulting choices regarding the data sources and associated parameter estimates must be described in detail and justified in the economic evaluation report.

The potential implications of these trade-offs should be considered in the context of the reference case probabilistic analysis, or using scenario analysis. That is, an estimate that is deemed to be credible, but for which there are concerns regarding its fitness for the purpose of the evaluation (e.g., efficacy estimates), may be accounted for in the probabilistic analysis by shifting the effect and interval estimate accordingly. The magnitude of the estimated impact of the lack of fitness on the parameter estimate may be informed by experts (ideally through expert elicitation) or empirical data.[62] The interval around the estimated value should be widened to reflect the degree of uncertainty associated with the magnitude of the affect. Alternatively, the data source judged to represent the best trade-off among the criteria may be used in the initial reference case analysis and the sensitivity of the results to other potential sources of estimates investigated in subsequent scenario analyses.

In situations where researchers are faced with limited information for informing parameter estimates, the analysis of uncertainty will play a particularly important role. The implications of sparse data and the associated uncertainty should be reflected in the width and shape of the distributions used in the reference case probabilistic analysis, and assessed through scenario analyses incorporating alternative structural assumptions or sources of estimates. For a discussion of some of the challenges of synthesizing sparse data (e.g., handling study groups with zero events), researchers are referred to Cooper et al.[42]

Other Key Considerations

Surrogate Outcomes and Biomarkers

A surrogate outcome is used in clinical trials as a substitute for the direct measurement of patients' preferences, function, and survival.[65] In the absence of data on the parameter of interest, the surrogate outcome is intended to predict the effectiveness or harms associated with the interventions being compared.[65]

Researchers should evaluate and justify the validity of surrogate outcomes used to estimate parameters in the economic evaluation.[66] When selecting a single data source from among multiple sources reporting different candidate surrogates, these surrogates should be assessed in terms of the previously defined criteria of fitness for purpose, credibility, and consistency. As before, researchers should employ judgment regarding which data source (i.e., candidate surrogate) represents the best trade-off among these criteria in terms of informing the parameter estimation for the economic evaluation. In the reference case, the true association between the surrogate and the final outcome should be treated as unknown and this uncertainty reflected in the probabilistic analysis. Uncertainty in the association of the surrogate with the final outcome can also be explored through appropriate scenario analyses. Where multiple potential valid surrogates exist, this should also be reflected in the analysis of uncertainty.

While researchers may refer to biomarkers to support the analysis of subgroups that are more likely to benefit from a given intervention (see Target Population section), their usefulness as valid surrogate outcomes can be more challenging to justify. As a result, a biomarker capable of distinguishing between patients who are more or less likely to benefit from a given intervention may not necessarily have value as a valid surrogate outcome to measure effectiveness.[65] To ensure the intervention has the expected effect, the researcher should evaluate and provide justification for the validity of the biomarker and the degree to which the biomarker satisfies the criteria for a surrogate outcome. This would involve consideration of whether substantive data are supported by a clear mechanistic rationale, and whether the data provide strong evidence that an effect on the surrogate is predictive of an effect on final outcomes.[65]

A determination of the appropriateness of the approach used to translate the surrogate outcome to the final outcome will ultimately require the availability of suitable real-world data sources in order to verify the results. In the meantime, researchers should adhere to the preceding guidance to minimize the chance of producing misleading results. Where the value of collecting additional real-world information is of interest to the decision-maker, this should be explored in the reference case analysis of uncertainty using value-of-information analysis (see Uncertainty section).

Extrapolation

Depending on the nature of the data available to inform parameter estimates and the extent to which they are censored (i.e., the time at which the event of interest occurs is not observed), extrapolation may be required to derive estimates for long-term effects (e.g., where only short-term data relating to effectiveness are available). Extrapolation of long-term effects requires both estimation of the long-term natural history of the condition and the effectiveness of the intervention beyond the time horizon for which data are available.

Time-to-event (also referred to as survival) analysis using parametric models can be used to extrapolate from shorter-term parameter estimates to longer-term effects.[67] Systematic approaches to survival analysis based on individual-level data have been developed. These analyses should follow the Survival Model Selection Process Algorithm developed by the Decision Support Unit commissioned by the National Institute for Health and Care Excellence (NICE).[68]

Two issues of central importance when considering the appropriateness of an extrapolation method relate to assumptions regarding the effects of treatment beyond the observed data and the effects of treatment beyond the treatment period. Researchers should report and justify the percentage of the estimated effect that occurs beyond the observed data. This can be measured as simply the ratio of the estimated incremental QALYs over the period of time that clinical effectiveness data are available to the estimated incremental QALYs over the entire time horizon of the model. Researchers should also report and justify the percentage of the estimated incremental benefit that is accumulated after treatment is stopped. This can be measured as simply the ratio of incremental QALY gains during the period on treatment to the estimated incremental QALYs over the entire time horizon of the model. Considering whether both these values are clinically realistic will help researchers to assess the suitability of the extrapolation methods. Expert judgment may be helpful in this regard.

In all situations, the researcher must describe the strength of the evidence for extrapolating parameter estimates. Where researchers have access to patient-level data, they should adopt appropriate techniques as defined by the Survival Model Selection Process Algorithm.[68] Researchers must justify all the approaches and assumptions used. A full consideration of parameter uncertainty can be incorporated through a probabilistic analysis incorporating correlation with respect to the parameters within the survival function. Scenario analysis exploring structural uncertainty should be conducted using alternative plausible parametric forms, as well as comparing results with and without an assumption of proportional hazards in order to assess how well the distributions fit the observed data.

Where researchers have access only to summary-level data, they may consider the use of methods to recreate patient-level data,[69,70] rather than relying on summary measures. For methods related to the synthesis of time-to-event (survival) data based on either summary or individual participant data, researchers are referred to Cooper et al.[42]

The duration and magnitude of the clinical effect beyond the study is often a critical judgment to make regarding extrapolation. Justifying the plausibility of the extrapolation may involve reference to external data sources, biology or clinical expert judgment.[67] In the reference case, the researcher must consider the best estimate of the duration and magnitude of clinical effect beyond the period for which data are available. It is not acceptable to assume that the relative effectiveness will be maintained for the duration of the intervention without adequate justification. In all situations, scenario analyses should be conducted that consider alternative assumptions relating to the waning of effectiveness. Although not an exhaustive list, reasonable alternatives to consider include assuming that the intervention effect continues

only for the duration for which data are available; that effectiveness declines over time; and that effectiveness continues for the duration of the intervention. With respect to the second of the three options, researchers may consider that the relative reductions in the risk of events may be reduced after a specific time point, either stabilizing at the reduced effectiveness or declining over time until the intervention is no longer effective. Particular care should be taken when extrapolation is required for parameter estimates obtained early in a technology's life cycle.[71] This is due to the potentially substantial extrapolation required to estimate the mean time-to-event over the entire life cycle of the technology.

Novel Trial Designs

Adaptive trial designs are the most commonly used novel approaches. An adaptive study design is one in which the statistical methodology allows for modification of design elements (e.g., sample size, randomization ratio, number of treatment groups) and an interim analysis with full control of the type I error rate.[72] As part of the researcher's identification of data sources, the designs of clinical trials should be assessed to identify any features that may affect the use or interpretation of the data (considering fitness for purpose, credibility and consistency). Given the evolving nature of such study designs, it is recommended that where other data sources exist, these should be used to inform parameter estimates in the reference case and the impact of information from novel designs considered in scenario analyses.

10. Measurement and Valuation of Health

HRQoL is a multidimensional measure of the effect of a health condition or intervention on an individual's overall well-being, encompassing physical and occupational function, psychological state, social interaction and somatic sensation.[73] Many methods have been developed to measure HRQoL or aspects of it; however, preference-based measures, which provide a single, overall summary score that numerically reflects the value of a given HRQoL state, are the only approaches suitable for use in a CUA.

Within these *Guidelines*, the terms "preference" and "utility" are generally treated as synonymous when referring to measures of HRQoL. Technically, however, "utilities" refer specifically to preferences obtained by methods that involve choices made under uncertainty (i.e., standard gamble approach).[2] For the purposes of these *Guidelines*, the term "utilities" is broadly applied to describe the numeric weight that quantifies the value of specific HRQoL states. These "utilities" are based on preferences for a particular health state, where a higher "utility" equates to a more preferable health state and a better HRQoL.[74] A utility of 1 represents perfect health, while a utility of 0 represents dead or a health state equivalent to being dead. Utilities less than 0 are possible for health states that are considered worse than dead. By definition, utilities are based on a cardinal scale (i.e., interval) such that changes can be compared (i.e., increases in utilities and the magnitude of the increases are reflective of preferences).

Given the social decision-making viewpoint adopted in these *Guidelines*, estimated utilities that reflect the preferences of the general population should be used in the reference case (refer to Introduction). As it has been shown that preferences can differ among individuals from different countries,[75] ideally, preferences that are representative of the Canadian general population should be used. Where there may be variation in preferences within or across jurisdictions (e.g., provinces or regions), and researchers have access to information specific to the provinces or regions of interest, these preferences should be considered.

Data Sources

As part of conceptualizing the model (see Modelling section) researchers should identify health states for which utilities will be required. Parameter estimates for utilities may be obtained from either direct or indirect methods of measurement. Direct methods of measurement refer to methods that enable the direct translation of preferences onto a utility scale, while indirect methods require a conversion scale to derive utilities. The most common direct methods include time-trade-off, standard gamble, and discrete choice experiments. Direct approaches are complex to design and implement and, unless individuals' own health states are being measured, likely to result in estimates that are highly dependent on the validity of the descriptions of the health states.[76,77]

Given the complexity and cost associated with directly measuring health utilities, pre-scored multi-attribute health status classification systems have been developed.[2] These classification systems can be generic (i.e., not focused on the health impacts of a specific condition) or condition-specific. The most common generic classification systems include the EQ-5D, the HUI, and the SF-6D.[78-80] While the various classification systems are unique, each one consists of a series of domains (e.g., mobility, self-care, physical activity, social activity) and each domain has a defined number of levels that correlate with a level of impairment.[75] To translate scores from an indirect preference-based measure to a meaningful utility score, established preference sets (algorithms) are used. These sets are generated by using direct preference measures (e.g., time-trade-off, standard gamble) for a sample of unique health conditions within the classification system to elicit preferences from members of the general population. Preferences for all health states described by the instrument are then predicted based on these findings. When seeking to interpret the meaning of these preferences, researchers should be aware that while they may reflect the general public's preferences for a health state, they may not necessarily reflect the preferences of a particular individual in a health state. Examples of Canadian-specific preference sets include the EQ-5D, which has been measured using the time-trade-off technique and a discrete choice experiment for both the EQ-5D-3L and EQ-5D-5L (three-level and five-level versions, respectively),[81,82] and the HUI Mark 2 and Mark 3, which have been measured using a combination of the standard gamble and visual analogue scale from a sample of adults from Hamilton, Ontario.[2]

To use an indirect generic instrument, depending on the version (self-administered or interviewer-administered), individuals or interviewers classify the health state by completing the questionnaire to determine the level of functioning on each domain. The scoring function is then applied to the health state in order to compute the associated utility. Where the scoring function is derived from a random sample of the general population, the resulting utility would represent an estimate of the mean utility for the health state based on a random sample from the general population.

Based on their ease of use, comparability, and interpretability, it is recommended that, in the reference case, researchers use utilities from an indirect method of measurement that is based on a generic classification system. The selection of a particular indirect generic method should be based on its fitness in terms of reflecting the health states of interest and their associated attributes, and in terms of capturing potentially important changes within and among states. The researcher should also assess the credibility of the instrument in terms of whether it represents an established instrument with demonstrated psychometric properties including feasibility, reliability, and validity.[83] Consistency with respect to the data used for estimating utility values is strongly recommended; in particular, it is recommended that utility data from the same instrument and population be used to estimate all the utilities in an economic evaluation.

When considering Canadian-specific preference sets, researchers must weigh their fitness for purpose against any issues of credibility and consistency with the data used to inform other parameter estimates, to ensure that the specific instrument addresses the decision problem.

There may be instances where utilities for the applicable health states have already been estimated based on indirect generic methods of measurement. These may have been gathered as part of routine data sets, published in the literature, or collected alongside a study. In such cases, these utilities should be critically assessed according to the previously discussed criteria of fitness for purpose, credibility, and consistency, to ensure that they reflect the health states of interest in the model and that the preferences reflect those of the general population. When searching for utilities, it is important that the search methods are comprehensive and presented in a transparent manner such that they can be replicated by others (i.e., similar to the approach for searching for data on clinical effects, as detailed in the Effectiveness section).

Often, there will be multiple utility estimates for a given health state.[2] Although meta-analysis has been used to combine utility estimates,[84] methodologies for the synthesis of health utilities continue to evolve and should be revisited as advancements in the understanding and application of these methods are made.[43] Similar to the selection of other data inputs, when selecting utility estimates from among multiple indirect generic measures, researchers should employ judgment regarding which data source represents the best estimate, weighing trade-offs among the criteria of fitness for purpose, credibility, and consistency. Based on a consideration of these criteria, the researcher must justify the selection of the utilities used in the reference case analysis. Regardless of the data source(s) used to inform utility estimates, probability distributions for each utility value should be derived and the associated uncertainty propagated through the model. The potential implications of any trade-offs among utility estimates should be considered in the context of the probabilistic analysis or using scenario analysis.

In circumstances where data from an indirect generic utility measure may not be optimal — for example, where they may not capture changes in health attributes of importance to a specific condition[85] — researchers should justify the lack of suitability of these measures for addressing the decision problem. Consideration of the use of a disease- or condition-specific outcome measure should involve the evaluation of its measurement properties, including content validity, construct validity, reliability, and responsiveness relative to a generic measure.[86] Researchers must ensure that the disease- or condition-specific measure provides a preference-based measure of health status. In order to assess the implications of using a disease- or condition-specific alternative, a comparison between the results of the reference case analysis based on an indirect generic utility measure relative to a non-reference case analysis based on an indirect disease- or condition-specific measure should be undertaken.

In the context of the social decision-making viewpoint adopted in these *Guidelines*, the valuation of health state utilities should be based on the preferences of the general population. When there is, however, a concern that general population preferences may not fully represent the experiences or outcomes of those who are affected by the intervention (both in terms of new interventions to be funded, and those that would potentially be defunded), alternative sources of preferences may be considered. Preferences other than those of the general population would constitute a non-reference case analysis. Justification for the use of preferences other than those of the general population, as well as the methods for measuring and valuing the utilities should be provided and clearly described so that the implications can be assessed relative to the reference case results. This would be particularly important when outcomes of the analysis are sensitive to preference weights.

Statistical mapping functions have been used to convert from one measure to another (e.g., HUI to EQ-5D) in an effort to maintain consistency, but mapping as a means of deriving health utilities is not recommended. "Mapping" refers to the development of algorithm(s) to predict health utilities using data on other indicators, or measures of health.[87] The key concern is the explanatory power that exists between the new "target" HRQoL and the "source" HRQoL measures. Utilities may be more highly correlated when mapping is conducted between two generic preference-based measures, compared with an exercise between specific and generic measures, which may be the result of lower degrees of overlap between the dimensions of HRQoL captured by specific and generic measures.[88] As such, the predictive value can vary dramatically depending on the instruments being mapped, the algorithm used, and the severity of the health states included and, therefore, mapping is unlikely to successfully capture the utility relationship.[88] To add to the complexity, there also exist several mapping algorithms to convert between the same HRQoL measurement tools, resulting in different estimates.[89] When there is no alternative but mapping to obtain utility estimates, empirical data are required. The mapping should be undertaken using a validated algorithm. The full uncertainty in the mapped estimates, reflecting parameter uncertainty in the source algorithm, parameter uncertainty in the mapping algorithm, and uncertainty regarding which HRQoL states are represented in the clinical health states included in the model, should be incorporated into the evaluation.

Combining Health Utilities

Economic models may involve states of health defined by a combination of health states (i.e., joint health states). Ideally, utilities may be obtained for these joint health states. However, it is often not possible to identify utilities that fully reflect the combination of health states. Three standard approaches to combining health utility information are the additive, multiplicative, and minimum methods.[43] These methods, however, have not been validated.[43]

The additive approach involves taking the utility for a single health state and applying the marginal disutility for each additional state. Marginal disutility refers to the difference in utility between patients in a specific state of health and those not in that health state. Use of this approach relies on the assumption that there are no interactions among the health states. The multiplicative approach multiplies individual health state utilities to derive a single health utility. This approach assumes that the individual utility for each of the states applies, and that the result of being in multiple health states is a joint health state with lower utility than any of the individual states. The minimum approach involves assessing the utilities for individual health states and using the lowest utility to represent the joint health state.[90]

Each of these approaches approximates a joint utility, but none is free of bias or necessarily captures the true value. Further research is required to strengthen and validate these methods, as well as to explore other emerging methods.[90-92] Given the uncertainty with these approaches, researchers should justify the approach undertaken and adopt appropriate scenario analyses to assess the impact of alternative plausible assumptions.

Health States That Are Challenging to Estimate

There may be health states for which the estimation of utilities is particularly challenging, due to both limited data and the lack of consensus on methods (e.g., health states for individuals with disabilities, states affecting vulnerable populations, temporary health states, states with spillover effects on informal caregivers).[14,15,93-95] Given the dearth of information with which to estimate utilities for such health states, the analysis of uncertainty will be especially important. Researchers should categorize uncertainty in their estimates in terms of both parameter and methodological uncertainty. Parameter uncertainty stemming from a lack of data should be

handled in the reference case using probabilistic analysis that incorporates increases in the width of the interval estimates. Sources of methodological uncertainty (i.e., the use of methods that deviate from those recommended in the reference case) should be addressed through additional non-reference case analyses that accompany and can be compared to the reference case.

Quality-Adjusted Life-Years

A QALY is the recommended outcome to capture health effects when conducting a CUA. The QALY takes into account the impact of interventions on both length of life (e.g., life-years gained) and quality of life.

The QALY is calculated by multiplying the number of life-years within a given health state with a utility that reflects the HRQoL in that state, aggregated over the various health states encountered over the course of the analysis. For transparency, the utility weights and discounted time in each health state should be reported in a disaggregated manner before being combined to calculate the QALY, as detailed in the Reporting section.

The incorporation of utilities into health effects allows preferences for health states and effectiveness measures to be brought together into a single metric. This single metric reflects the valuation for different health effects and enables comparisons across health care interventions.

Valuing Non-Health Effects

Where the decision problem requires a perspective other than that of the publicly funded health care payer (e.g., broader government payer, societal), the evaluation may involve the consideration of non-health effects (see Perspective section, Table 1). Non-health effects should be accommodated using a CCA to complement the health effects captured in a CUA. The incorporation of non-health effects could also be achieved by conducting a non-reference case CBA. Other instruments, such as time-trade-off or standard gamble, could also be used. In the context of the social decision-making viewpoint underpinning these *Guidelines*, the value of non-health effects (e.g., criminal activity, levels of education) should be based on being traded off against health in order to be incorporated into the economic evaluation. As noted in the section on Types of Evaluations, these valuations should reflect those of society.

11. Resource Use and Costs

This section is intended to guide the identification, measurement, and valuation of resource use and costs that are typically required for model-based economic evaluations. The information contained in this section should be used in conjunction with the CADTH document *Guidance Document for the Costing of Health Care Resources in the Canadian Setting: Second Edition*.[96]

Resource Identification

The researcher should identify all activities and resources that are likely to occur within the context of the decision problem (e.g., accounting for the target population, perspective, and time horizon). The conceptualization of the clinical or care pathway for the health condition will provide the basis for identifying relevant resources. The structure of the pathway will dictate how resource use and the associated costs are included in the model (e.g., whether they are determined by health state or event). Researchers must consider all resources that occur along the pathway and that are attributable to the interventions of interest. For instance, future resource use should be included where it is understood that the clinical or care pathway includes resource-intensive health states, or events that occur over long

periods. When considering multiple perspectives as part of addressing the decision problem, researchers should identify all relevant resources incurred by the different perspectives, and present them in a disaggregated manner based on the reference case perspective and any additional non-reference case perspectives.[97,98]

The associated resources may differ, depending on the type of intervention being evaluated (e.g., whether the intervention is intended to prevent, diagnose, or treat). When the intervention is being defined, resources associated with implementation, operation, maintenance and repair, staffing, training, and overhead should be identified and included in the analysis.[2] For example, when considering a new intervention that has infrastructure requirements associated with its use (such as building requirements), requires that staff are trained prior to use (or they have a set retraining schedule or learning curve to become proficient), has a specific maintenance schedule to ensure performance, and/or requires a specified dose-titration schedule to achieve therapeutic doses, these should be considered in the identification of resources associated with the intervention. The impact of a new program on existing infrastructure should also be taken into account where the scale or scope of the program has implications; for example, in terms of capacity issues.

Resource items may be excluded where there is identical use among the interventions being evaluated over the time horizon of the analysis,[99] but researchers must justify this assumption. Resource use associated with an event that is deemed irrelevant to the intervention being evaluated (i.e., there is no causal association between the outcome(s) of interest and the event, such that the resource use or event would have occurred regardless of whether the intervention is used; for example, use of acne medication and breaking a leg while playing sports) need not be included. The distinction between what constitutes a relevant and an irrelevant event requires careful consideration. This may be particularly true in the context of chronic health conditions in which the presence and treatment of the chronic condition may decrease the likelihood of receiving treatment for other disorders.[100]

One option for determining which clinical events are relevant is through an adjudication committee (blinded to treatment assignment). This would allow the researcher to remove irrelevant events in an unbiased manner. Another option would be to conduct a scenario analysis to assess the impact on the cost and effect results of including all events compared with including only those events deemed relevant. Where the results of these analyses are the same, the excluded events would be irrelevant.

Resource Measurement

Resource use should be measured and quantified in as disaggregated a form as necessary for the economic evaluation. The required level of precision for quantifying resource use will guide the resource inputs for the economic evaluation. Where greater precision is required, individual-level costing may be appropriate; however, where resource use is unlikely to vary based on individual characteristics, resource quantities measured by event or case may be sufficient.[96]

When informing parameter estimates for resource use, similar to estimating other model inputs (e.g., effectiveness or harms), researchers should first define what constitutes relevant data, based on the decision problem. In the context of resource use, the jurisdiction of interest will be of particular importance when seeking to identify potential data sources.[2] As such, a search for information on resource use may be more narrowly defined compared with a search for information on other model parameters (e.g., clinical effectiveness), but researchers must still adopt a transparent and justified search approach.

Given the importance of using resource estimates that reflect the jurisdiction(s) defined in the decision problem (e.g., as a result of concerns regarding the generalizability of the information between and within jurisdictions), the synthesis of resource use data from different locations is not recommended, although the synthesis of data from a single location would be an option. Researchers should consider the potential for variations in resource use both between and within jurisdictions.

When faced with resource use data that reflect different locations, researchers should select a single data source based on an assessment of the relative trade-offs among the various sources. These trade-offs should be assessed in terms of the previously identified criteria of fitness for purpose, credibility, and consistency. Fitness for purpose should be weighed more heavily to ensure the data reflect the jurisdiction(s) of interest.

Administrative databases represent a potentially valuable source of information for informing estimates of resource use.[101] These databases provide real-world information on resource use specific to jurisdictions.[102-105] When information for individuals is linked across various databases, they may also offer insight into quantifying resource use beyond those resources directly relevant to the intervention(s) or condition of interest. In particular, administrative data may be useful in estimating resource utilization over time for defined patient cohorts. To the extent that administrative databases capture the impact of actual practice within a jurisdiction, they also have the potential to serve as a way of updating the results of model-based evaluations based on real-world observations. Despite strengths, administrative databases are not necessarily designed to address the specific decision problem, and researchers should exercise the usual caution in ensuring that the data are fit for purpose, credible, and consistent with sources used for estimating other parameters in the model.

Where information is collected in a study, researchers should be mindful of protocol-driven resource use.[106] This may be assessed by considering relevant Canadian clinical practice patterns. Where resource use from the clinical studies is driven largely by the protocol and does not appear to reflect Canadian practices,[2,106,107] or does not reflect the target population, researchers should seek alternative sources of information.

Researchers should also be careful to avoid double counting when measuring resources; for example, avoiding situations where resource use is estimated separately as well as being bundled together in a single case-mix estimate (e.g., medical devices). In such situations, researchers must be careful to avoid overestimating the use ascribed to a particular resource.

When incorporating resource use into an economic evaluation, researchers should estimate the mean value as well as the associated uncertainty (i.e., probability distribution) (see Uncertainty section). Where more than one data source is judged to be appropriate, the potential implications of trade-offs among the sources in terms of fitness for purpose, credibility, and consistency should be reflected in the analysis of uncertainty. Furthermore, where there is a lack of data that are sufficiently relevant to the jurisdiction(s), resulting in researchers having to rely on assumptions or methods of adjustment (e.g., simple adjustment procedures based on differences in factors such as practice patterns and settings[108]), these should be clearly explained and justified and assessed in the reference case probabilistic analysis, or using scenario analyses to explore the impact of alternative assumptions.

Resource Valuation

One concept central to health economics is that of opportunity cost. From a theoretical perspective, the proper price for a resource should be determined by its opportunity cost (i.e., the value of the resource's best alternative use).[2,96,109-113] The determination of opportunity cost will depend on the perspective of the analysis as defined by the decision problem.

In theory, the price of a resource reflects the opportunity cost; however, in practice, fees or charges paid directly within the Canadian health care system for services may be used to approximate the opportunity cost when conducting an analysis from the perspective of the Canadian public payer.[96] When referring to fee schedules, researchers should be mindful of situations where the listed fees or charges are not followed in actual practice. For example, in psychology and other non–medical doctor disciplines, as well as in the case of nursing home care, the stated fees may not reflect the actual payments and, consequently, they will not be a good measure of the cost. In such situations, researchers should investigate alternative data sources. Researchers should also consider the possibility of a copayment where resources are valued from the perspective of the public payer. Copayments covered out of pocket by an individual, or by a private insurer should be excluded from the estimated payments of the public payer. A copayment can have implications for the effectiveness of an intervention (e.g., in terms of adherence, compliance to therapy), as such, potential implications on effectiveness should be considered.

When a perspective broader than that of the public payer is adopted, fees or charges other than those paid directly by the public payer should be considered. For example, when conducting an evaluation from the perspective of a private payer, costs beyond those paid for by the public payer must be included. From a societal perspective, incorporating individual out-of-pocket costs, charges, and fees that reflect actual payments by an individual (e.g., copayments) would be applicable. The researcher must consider the perspective(s) (see Perspective section) when selecting and justifying the choice of data sources.

Similar to the selection of data sources for resource use estimates, researchers should select data for costs based on an assessment of the respective trade-offs among the available sources, considering fitness for purpose, credibility, and consistency. Like resource use, concerns regarding the generalizability of costs both between and within jurisdictions require that particular emphasis be placed on the fitness of the data for estimating costs relevant to the jurisdiction defined in the decision problem. Consequently, the use of costs from international sources is not recommended. When Canadian costs are not available, the researcher may consult experts and/or current available fees or charges for similar existing or new technologies to estimate the costs, and look to international sources as a guide (e.g., in terms of estimating the costs of one drug relative to another). Where an economic evaluation is conducted to support a decision at a regional level, region-specific sources should be used. For an analysis from a pan-Canadian perspective, Canada-wide sources should be used. Where this information is not available, a weighted average of costs (costs from jurisdictions weighted by the proportion of use in Canada represented by the jurisdiction) or costs from a representative jurisdiction may be used. Uncertainty in the estimated value should be incorporated into the reference case probabilistic analysis. The sensitivity of the results to alternative sources for the estimates should be tested extensively using a scenario analysis.[96]

When estimating costs from a study or administrative database, mean values should be used and the associated uncertainty reflected in the estimates (i.e., probability distributions). As

with survival data, mean costs estimated in the presence of censoring should be adjusted accordingly.[114] The use of administrative data makes it more tractable to estimate ongoing health care costs, including the cost of death, for any disease and any combination of covariates. Research as well as the continued development of administrative data sources for estimating total health care costs will further facilitate the inclusion of these costs in economic models.

When estimating costs based on resource use and unit costs, researchers should consider the time period for which the unit-cost data are available. Where unit costs are available only for a previous time period, costs should be assessed to ensure they reflect current practice, and they should be updated to the current year. As there are no price indices in Canada for hospital or physician services, a general price index such as the Consumer Price Index can be used.[96]

The appropriate valuation of resources also requires that researchers consider the relationship between the quantity of resources used and the unit-cost estimates, and how they may change over time. That is, whether the relationship between efficiencies obtained through experience or lower doses can be expressed as potential cost savings should be considered. Any assumptions and the approach undertaken should be justified. Where applicable, assumptions and methods used for allocating overhead costs, shared labour costs, and administrative costs should also be clearly described and justified. The impact of possible alternative assumptions and/or approaches should be assessed using scenario analysis.

Patient and Caregiver Time

Should the decision problem reflect a broader societal perspective, it may be relevant to include the impact on patient and/or caregiver time[115] (e.g., the condition may result in the patient or informal caregiver giving up activities that would otherwise be undertaken, involving both costs and time implications)[116] in a non-reference case analysis. The costs of lost time may reflect the individual's inability to participate in paid labour (e.g., absenteeism, presenteeism) or unpaid labour (e.g., informal caregiving) as a result of illness, treatment, disability, or premature death.[115] Researchers should be careful not to double-count the costs as lost time from both paid and unpaid labour.

There has been debate about the appropriate approach for estimating lost time in an economic evaluation. The two primary approaches for valuing lost time from paid work are the human capital approach and the friction cost approach.[2,17,117-123]

The human capital approach is the simplest approach for valuing lost time from paid work. In this approach, the value of lost time is derived by multiplying the amount of time off work by the lost compensation rate (age- and sex-adjusted),[118] plus fringe benefits (e.g., pension benefits; health and life insurance). The human capital approach assumes there is (near) full employment; however, this may overestimate the value of lost production.[124]

In the friction cost approach, it is assumed that when a person is away from work, he or she is eventually replaced with a previously unemployed person, so that the productivity loss to society is limited to the time before the person is replaced (i.e., the friction period).[125-127] 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 or informal caregiver's lost production may be partly restored 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.[2]

As the focus of the societal perspective is on lost productivity, the friction cost approach is the recommended method when estimating productivity losses due to absence from paid work. Any other approaches may be considered in additional non-reference case analyses and the results compared among the analyses.

Productivity may be lost even though the patient or informal caregiver remains at work (i.e., presenteeism).[2] Depending on the nature of the health condition being studied in the evaluation, presenteeism may account for a large proportion of the productivity losses (e.g., mood disorders, migraines).[2,115] Questionnaires have been developed to help estimate lost productivity due to presenteeism, absenteeism, and unpaid work,[119] but there is no consensus as to the best instrument,[119] and the ideal instrument may vary depending on the application. When appraising a work productivity instrument for use in an evaluation, some general criteria to consider relate to the purpose, perspective, population, psychometrics, and practicality of the instrument relative to the requirements of the evaluation.[119]

Two methods have been proposed to place a value on a patient's or caregiver's lost time from unpaid work: the opportunity cost approach, and the replacement cost approach.[96] The former approach is based on forgone wage, while the latter approach assesses the cost of purchasing the resource (worker) to undertake a service (e.g., housekeeping) that was forgone by the informal caregiver or patient.[96] The opportunity cost method is recommended to estimate productivity costs related to unpaid labour, as this approach values time spent on unpaid work based on the value of spending this time in an alternative capacity (e.g., paid work) rather than relying on the value of a market substitute (e.g., hired housekeeper). When determining the opportunity cost, researchers should seek direction from the decision-maker as to their preferred estimate of opportunity cost (e.g., average Canadian wage rate). Researchers should be aware of and transparent about any equity implications, but ultimately the estimated opportunity cost should reflect the decision-maker's equity position.

It is recommended that lost leisure time be excluded as a cost in evaluations.[128-131] Lost leisure time would be partly captured by the QALY, as part of the preference for a state of health. Individuals participating in exercises measuring preferences should be told to value changes to leisure time (and to assume that health care costs and income losses are fully reimbursed and/or compensated).

12. Analysis

This section relates to how data inputs obtained to populate the economic model (e.g., clinical effects, harms, resource use, utilities) should be analyzed in terms of estimating outputs for expected costs, outcomes, and ICERs.

Probabilistic Analysis

The final results should be based on expected costs and expected outcomes. These should be estimated through a probabilistic analysis, which will provide less biased estimates of costs and outcomes than a deterministic analysis (see Uncertainty section for further details). Probabilistic analysis requires data parameters to be represented by statistical distributions rather than point estimates. This allows for the characterization of the underlying uncertainty (due to uncertainty in the evidence base) regarding the estimated costs and outcomes for the parameters included in the CUA.

The principal argument in favour of adopting probabilistic techniques for the reference case analysis is the potential for deterministic analyses to lead to non-optimal decisions

because of non-linear relationships between the input variables and outputs.[132] Given the characteristics of decision-analytic models in health care, especially Markov models, there is potential for discordance between the results of probabilistic models and deterministic models.[133] Deterministic analysis, in which the researcher considers only the expected values of individual data elements, will often give incorrect estimates of costs and outcomes. Probabilistic analysis, which incorporates the likelihood of each parameter taking different values, provides a less biased estimate of costs and outcomes.

In most cases, probabilistic analysis will take the form of a Monte Carlo simulation. A Monte Carlo simulation is a suitable technique for conducting an analysis to determine the propagation of uncertainty from input parameters to outcomes.[134] Within a Monte Carlo simulation, costs and outcomes for each intervention are obtained by rerunning the model employing random values for each input parameter: each set of input values and the associated costs and outcomes are referred to as a replication. This step is repeated a number of times, and the expected values of costs and outcomes are estimated as the mean of the values obtained from all replications. The number of replications should be sufficiently large that the expected values for costs and outcomes are stable (i.e., are unlikely to change substantially with a greater number of replications) as they are not subject to substantial Monte Carlo error. In most cases, a minimum of 5,000 replications will be required, but researchers should investigate to determine whether larger numbers are necessary for stability.

The values for each input parameter are drawn randomly from a probability distribution specified for each parameter. Thus, to conduct a Monte Carlo simulation, it is necessary to specify a probability distribution for each input parameter for which there is uncertainty; typically, this will relate to parameters that were obtained from sample data or from expert elicitation. Alternatively, where the data analysis and propagation of uncertainty are conducted using a Bayesian approach, parameter uncertainty is sampled from a Bayesian posterior distribution combining a prior distribution and a likelihood function. In that way, the probability distribution can incorporate information from both the sample data (likelihood function) and the expert elicitation (prior distribution).[135] Probability distributions are unlikely to be required for data inputs such as general population mortality rates, which reflect baseline rates and are effectively treated as known quantities. Probability distributions provide the potential range of values for an input parameter and the corresponding likelihood of each value. Given the focus of analysis on parameter uncertainty, the range reflects the potential mean values of the parameter for the population of interest, not the potential values for an individual.[136]

Individual patient simulation models require incorporation of both patient variability and the underlying population uncertainty. Individual patient variability is captured by the standard deviation. However, with such models, it is necessary to conduct a probabilistic analysis, where the uncertainty at the population level is appropriately characterized by standard errors.

The choice of the form of the distribution should reflect the nature of the input parameter and should follow standard statistical methods.[34] It is essential that the choice of distribution recognizes the bounds of values for each parameter; that is, utilities cannot exceed 1, costs cannot be less than 0, and transition probabilities must be bound by 0 and 1. Beta distributions are the natural distribution for transition probabilities; in certain instances, a Dirichlet distribution (the multinomial equivalent to beta) may be required. Gamma or log-normal distributions can be used for data that are skewed right, such as costs and disutilities. Log-normal distributions can be used for relative effects. Distributions such as triangular and uniform distributions that impose artificial bounds on parameter estimates should not be used.[34]

The degree of uncertainty for each input parameter should be based on the mean and standard error from the relevant data source. If data on the degree of uncertainty are unavailable, a conservative approach should be adopted whereby an estimate of the standard error is assumed that allows for plausible parameter values. When data on the degree of uncertainty are unavailable, the researcher must provide details on how uncertainty was handled and provide a justification for the approach adopted, including the plausibility of the parameter values being considered. The lack of sufficient data is not a justification for failing to incorporate the impact of uncertainty with respect to a given parameter. Expert elicitation is an acceptable strategy for dealing with uncertainty in the absence of sufficient data.[137]

Correlation among parameters should be considered, as it can affect both expected values and their degree of uncertainty.[34] When regression analysis is used to define input parameter values, it is necessary to incorporate the correlation among coefficients through the use of the Cholesky decomposition of the covariance matrix.[34] Such methods should also be considered carefully when there are other instances where correlation among input parameters may be high. When evidence synthesis results need to be incorporated into a probabilistic analysis, four alternative approaches that will correctly propagate the uncertainty and correlation structure have been identified.[138]

The assumptions regarding the uncertainty with respect to parameter values, the form of probability distributions, the number of Monte Carlo replications, and the consideration of correlation should be stated and justified.

Estimated Values

The results should be reported as ICERs (i.e., the difference in expected costs between two interventions divided by the difference in expected outcomes). The net monetary benefit measure may be used as an additional (but not alternative) measure to the ICER.[139] It should be noted that when interpreting ICERs, it is not always the case that the intervention with the lowest ICER is the most cost-effective and when evaluating multiple comparators, net monetary benefit may be particularly helpful in terms of aiding interpretation of the cost-effectiveness results. When calculating net monetary benefit, the cost-effectiveness threshold should be stated for each net benefit estimate, and based on that specified by the decision-maker.

When there are more than two interventions being compared, the expected costs and outcomes of the interventions and the relevant incremental ratios should be calculated sequentially.[140] In the reference case, a sequential analysis involves calculating the ICER for a less costly comparator compared with the next most costly comparator, excluding all comparators that are either dominated or subject to extended dominance. This approach requires the identification of interventions that lie on the cost-effectiveness frontier and those that do not (i.e., those that are subject to dominance or extended dominance).

An intervention is dominated when it is more costly and less effective than at least one other intervention. An intervention is subject to extended dominance when it would never be the optimal intervention regardless of the cost-effectiveness threshold. Extended dominance occurs when the ICER for a given intervention compared with a lower-cost alternative is higher than the ICER for the comparison of a higher-cost intervention with the same lower-cost alternative. This is illustrated in Table 2. Intervention B is subject to extended dominance, as there is no threshold value for a QALY where it will be preferred to both intervention A and C. Note that the ICER for intervention B versus intervention A is higher than the ICER for

Table 2: Comparison of Incremental Cost-Effectiveness Ratios by Intervention

	Cost	QALYs	Incremental Cost per QALY Gained (ICER)	
			Versus Intervention A	Sequential ICER
Intervention A	$3,000	4		
Intervention B	$4,500	4.1	$15,000	Subject to extended dominance through interventions A and C
Intervention C	$5,000	5	$2,000	$2,000

ICER = incremental cost-effectiveness ratio; QALY = quality-adjusted life-year.

Table 3: Example of Analysis of Dominance by Intervention

	Cost	QALYs	Incremental Cost per QALY Gained (ICER)	
			Versus Alternative A	Sequential ICER
Intervention A	$3,000	4		
Intervention C	$5,000	5	$2,000	$2,000
Intervention E	$8,000	6	$2,500	$3,000
Intervention F	$12,000	6.05	$4,390	$80,000
Intervention B	$4,500	4.1	$15,000	Subject to extended dominance through interventions A and C
Intervention D	$7,900	4.3	$16,333	Dominated by intervention C. Subject to extended dominance through interventions A and E and interventions A and F
Intervention G	$50,000	6.01	$23,383	Dominated by intervention F

ICER = incremental cost-effectiveness ratio; QALY = quality-adjusted life-year.

intervention C versus intervention A. Where the decision-maker's threshold value is greater than $2,000, intervention C would represent the most cost-effective intervention among the three comparators (all else being equal). In that way, the sequential ICER of $2,000 can be viewed as a critical value where for threshold values greater than $2,000, intervention C would be the most cost-effective intervention, and for values less than $2,000, intervention A would be the most cost-effective. (See Table 2.)

An appropriate approach for the presentation of a sequential analysis is to report each of the interventions on the cost-effectiveness frontier in a table in increasing order of cost. The ICER for each intervention is then calculated as the ratio of the incremental costs to incremental outcomes when compared with the intervention on the frontier, which is placed above it in terms of having lower costs. A frequently adopted approach is to provide the ICER for each therapy only compared with the least costly alternative; this, however, may only be presented in addition to a sequential analysis of the ICERs.

Interventions not on the cost-effectiveness frontier should then be listed in the table with a description of how they are subject to dominance or extended dominance. An example is provided to illustrate this method (see Table 3). In this example, there are three critical values for the cost-effectiveness threshold (i.e., $2,000, $3,000, and $80,000) around which the relative cost-effectiveness of the various interventions changes.

All expected costs and expected outcomes should be reported separately for each subgroup identified within the target population, with sequential analyses conducted for each stratum. If the decision problem requires an overall estimate, researchers can provide an estimate for the entire target population through weighting the results by subgroup.

13. Uncertainty

Economic evaluations are undertaken to inform decision-makers about the expected costs and outcomes of alternative courses of action. It is important that decision-makers be provided with accessible information on any uncertainty regarding the results. As such, researchers should take a systematic and consistent approach to the specification and analysis of uncertainty in economic evaluations. Three categories of uncertainty need to be explicitly addressed: methodological, parameter, and structural.[2,32,141]

Methodological uncertainty is concerned with unresolved questions about methods; for example, the appropriate discount rate to apply to costs and outcomes, and what utility algorithm should be used to construct QALY estimates. *Parameter uncertainty* is concerned with the extent to which the estimated value reflects the "true" value for each parameter in the analysis. *Structural uncertainty* considers whether all the relevant aspects of the health condition and intervention(s) are adequately captured by the chosen model structure.

In common with other good-practice guidance for economic evaluation, these *Guidelines* recommend that sources of methodological uncertainty be addressed by comparing the results of a reference case analysis to those from a non-reference case analysis. While the methods of economic evaluation have developed substantially over the last two decades, there are still areas where there is no consensus regarding the correct technical approach. The reference case specifies a set of methods to be used for all evaluations that promote uniformity and transparency, and enable the comparison of results for different technologies and different decisions. The guideline statements specify the content of the reference case with regard to statements about the decision problem, types of evaluations, target population, comparators, perspective, time horizon, discounting, measurement and valuation of health, resource use and costs, analysis, uncertainty, and equity.

Non-reference case analyses can accompany the reference case and be provided to decision-makers, but the impact of departing from the reference case should be explicitly stated. For example, uncertainty regarding the appropriate discount rate might be addressed through the comparison of a non-reference case analysis that includes a rate that differs from that recommended in the reference case analysis. Any non-reference case analyses should still use a probabilistic analysis to provide decision-makers with unbiased estimates of the costs and outcomes of the technologies being evaluated.

Parameter uncertainty should be addressed using a probabilistic analysis; that is, the uncertainty about the true value of parameters included in the economic evaluation (inputs) should be expressed using probability distributions, and the uncertainty should be propagated through the analysis, in order to quantify the resulting uncertainty in the outputs (expected costs and outcomes) using simulation methods. Whenever possible, the analysis of parameter uncertainty should take account of any correlation between parameters in order to avoid understating or overstating the true uncertainty in the costs and outcomes (see Analysis section). Similar to the approach discussed in the Analysis section, parameter uncertainty can also involve consideration of the critical values that could alter the decision as to whether an intervention is cost-effective. In this case, the critical values would refer to

parameter values that, for a given cost-effectiveness threshold, may change the decision. In addition to identifying critical parameter values, researchers should also give thought to the likelihood of the parameters taking on these values. Expert judgment may be helpful in making such assessments.

As the correct estimation of expected costs and outcomes requires the use of probabilistic analysis, the analysis of parameter uncertainty is undertaken concurrently with the estimation of the expected values. Hence, the term "probabilistic sensitivity analysis" is redundant. Deterministic analyses of parameter uncertainty (i.e., one-way, multi-way, or threshold analyses) are not recommended, as they will be misleading when models are non-linear, and in the presence of correlated parameters. The impact of changes in deterministic parameters (e.g., prices of new drugs) that do not have associated uncertainty, but rather are assumed to be known, should be assessed using scenario analysis. A comparison of the cost-effectiveness results (given values of the cost-effectiveness threshold) for the various scenarios could help to identify critical values for deterministic parameters.

Decision uncertainty (i.e., the risk of making the wrong decision) should be summarized using CEACs. As the assessment of the value of a technology should use the expected value of the incremental costs and outcomes, CEAFs should be included for evaluations with more than two technologies. Scatter plots on the cost-effectiveness plane may be provided but, because difficulty in their interpretation, these are not part of the reference case. Similarly, confidence ellipses on the cost-effectiveness plane are not recommended, as they add no further information compared with that provided by CEACs or CEAFs.

To enable the development of additional research to inform future decisions, decision-makers increasingly consider reimbursement options that combine some degree of adoption of a technology into the health system. There are a wide range of nomenclatures for such schemes, including coverage with evidence development, risk-sharing, and access with evidence development. An important differentiation in this area is between those schemes that make the technology available to all patients (irrespective of engagement with the research process), and those that make the technology available only to patients contributing data to the research.

When the decision problem includes consideration of further research to inform future decisions, a value-of-information analysis should be undertaken as part of the reference case. This should involve consideration of the critical values that could alter the decision as to whether an intervention is cost-effective. To identify these critical values and correctly quantify the impact of a parameter taking a specific value (on both the probability of an intervention being cost-effective and the expected net benefit), recent methodological work suggests that a two-stage expected value of perfect parameter information[142] analysis may be useful.[143] When decision-makers are interested in assessing the value of collecting additional information to reduce uncertainty in specific parameters, consideration may be given to plotting one-way probabilistic analysis graphs for these parameters.[143] Given that one-way probabilistic analysis represents an emerging method, its role should be viewed as supplemental (outside the reference case recommendations).

The expected value of perfect parameter information should be provided for all parameters identified as being critical to the decision in order to support the decision-maker's consideration of the contribution of each parameter or, where appropriate, groups of parameters (e.g., when parameters are correlated) to the total decision uncertainty. The population expected value of perfect parameter information should also be provided, reflecting both the likely size of the population and the lifetime of the intervention. Value-of-sample information and net-benefit-

of-sampling analyses will support decision-makers' assessments of the return on investment of further research when specific parameters or groups of parameters are identified as being responsible for a substantial portion of the total decision uncertainty.

Structural uncertainty is concerned with those components of the decision problem that are not captured by the parameters (e.g., the clinical pathway with and without treatment, model structure, time dependency, and the functional forms chosen for model inputs, such as time-to-event parameters). While some forms of structural uncertainty can be expressed as parameter uncertainty, this is not always possible. In such circumstances, structural uncertainty should be addressed through scenario analyses (i.e., one complete analysis should be provided for each alternative approach); for example, one analysis using a Gompertz curve and a second using a Weibull curve for survival data. In all other respects, these analyses should remain unchanged to allow the decision-maker to assess the impact of the alternative structural approaches on the results. For the purposes of informing decision-makers, researchers should identify a recommended structural approach for the reference case.

14. Equity

A starting place for all economic evaluations should be to acknowledge and respect both horizontal and vertical equity. The former requires that people with like characteristics (of ethical relevance) be treated the same, while the latter allows for people with different characteristics (of ethical relevance) to be treated differently.[144,145] Given the social decision-making viewpoint of these *Guidelines*, respecting both vertical and horizontal equity is essential. For any given vertical equity position adopted by the researcher or decision-maker where, for example, subgroup analyses might be considered, it is essential that horizontal equity be respected.

The use of QALYs is compatible with the aim of maximizing the value of health effects, where the value of health effects is independent of the characteristics of the individual who receives those effects, the condition being treated, and the technology that produces them. However, it is widely recognized that this is unlikely to be the sole goal of either the health care decision-maker or the populations they serve. Addressing health disparities and concerns for equity represent examples of additional policy objectives.

Equity is taken here to refer generally to notions of fairness, and can be considered in terms of health and health care.[144,145] Common to most definitions of health equity is the idea that some health or health care disparities (or inequalities) are unfair. In the context of health, equity is concerned with judgments about the fairness of the distribution of health outcomes and experiences in a population. When thinking about health care, the equity concern relates to the fair allocation of resources (in the form of interventions, technologies, etc.) among individuals or groups.

In economic evaluations in health care, the primary equity focus has traditionally been on health outcomes; the dominant practice being implicitly to assign equal social value to a unit of health improvement, irrespective of who receives the benefit. For example, in the context of a CUA, QALYs gained by socially disadvantaged groups are assumed to have exactly the same social value as those of advantaged groups: a QALY for a woman, an Indigenous person, a poor person, or a sick person counts the same as those of men, non-Indigenous people, wealthy people, and healthy people.

This assumption can be relaxed. Concerns relating to the unfair distribution of health outcomes can, in theory, be addressed by using differential weighting of outcomes, with health gains for the disadvantaged being given a higher value.[146] The consequence of this is, of course, that interventions targeting disadvantaged groups will, all else being equal, look more cost-effective.

Much work has been undertaken internationally to understand social views on what constitutes ethically relevant personal characteristics in the health care context. Candidate equity characteristics include age, gender, socioeconomic status, availability of alternative therapies, and prevalence of the condition. For example, the "Fair Innings" argument, originally presented by Alan Williams,[146] posited "lifetime experience of health" as a personal characteristic by which to differentiate beneficiaries of health care interventions. That is, those who have already experienced a fulsome and health-filled life should not necessarily, at that point, be given further advantage; rather, health gains made early in life, or made by those who have suffered poor health, should be valued more highly.

In the context of equity weights, Wailoo et al.[147] remind us of the founding principle of economic evaluation — namely, opportunity cost, which necessitates consideration of the characteristics of both those who stand to gain and those who stand to lose.[147] If explicit consideration is given only to the equity characteristics of patients receiving the new intervention, then the implicit assumption is that those who bear the opportunity cost have no special characteristics and their health is valued less than the average. Another key point here relates to the identifiability of those who stand to lose. Unless those on whom the opportunity cost falls can be identified, it cannot be determined if the value of the health produced from the expenditure of the health system budget is increased, decreased, or unaffected. Hence, in adopting equity weights for the identified beneficiaries alone, we cannot be confident that we are acting in line with the equity principles that society regards as valuable.

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 incidences of conditions; access to health care; or technology uptake in population groups. When the intervention can be provided selectively to certain subgroups, then cost-effectiveness information can be presented for each subgroup. Any stratified analysis of subgroups motivated by vertical equity considerations must be explained and justified. Further, such subgroup analyses must be consistent with CADTH's Guidelines on Target Population. If costs and outcomes differ among subgroups defined in terms of equity-related characteristics, this should be reported, allowing decision-makers to assess the distributional impacts of the investment in question.

Further, groups that are likely to be disadvantaged by the adoption and implementation of the intervention should also be identified, where possible. 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.

Given that many decision-makers are concerned about equity, economic evaluations should be presented in a manner that supports equity concerns being reflected in decision-making.[148] Hence, although the reference case analysis should weight all outcomes equally (regardless of the characteristics of people receiving the health benefit), analyses should be presented in a disaggregated manner, with full descriptions of the relevant patient populations, to allow for consideration of any subsequent distributional and equity-related policy concerns.

15. Reporting

The reporting of the economic evaluation should be clear and detailed, and the analysis and results should be presented in a transparent manner. Researchers should provide enough information to enable the reader or user (e.g., decision-maker) to critically assess the evaluation, including how each element of the economic evaluation, as outlined in these *Guidelines*, has been handled. The format of the economic evaluation should be well structured and easy to follow. To enhance clarity and facilitate the comparison of economic evaluations, researchers can use the structured reporting format in Appendix 1.

An executive summary should be included at the beginning of the evaluation and written in a manner that is easily understood by a reader without a technical background. Wherever possible, researchers should use plain language, and define jargon or technical terms that may be unfamiliar to the reader or user.

Elements of the Economic Evaluation

A decision problem should be clearly stated at the outset of the evaluation. This will detail the interventions being considered, the context in which the interventions are being assessed, the perspective(s), the costs and outcomes being considered, the time horizon, and the target population of interest.

Researchers should report how each of the elements of the economic evaluation have been conducted. Where the methods differ from the recommended reference case analysis, justification should be provided and the results reported as a non-reference case analysis that accompanies the reference case.

All cost and effect items contributing to the economic evaluation should be presented in a disaggregated manner before being aggregated. This should also be done for all scenario and non-references case analyses. To facilitate understanding, researchers are encouraged to present the results of the analysis in graphical (or visual) and tabular forms, as detailed in the Analysis and Uncertainty sections. All tables and graphics should be appropriately discussed and not used to replace a written description or interpretation of the results. The impact of methodological, parameter, and structural uncertainty on the results should also be reported.

Quality Assurance

Researchers should ensure the methodological rigour of the process underlying the economic evaluation, and provide adequate justification for any assumptions. This can be achieved by thoroughly documenting the decisions and choices made throughout the evaluation in a technical appendix, which will ensure transparency and rigour in the process. Model testing and validation processes should also be detailed. A description of the statistical analysis (i.e., data sources, methods, and results) should also be made available.

Documents specific to the quality assurance process should be made available. An operable (i.e., unlocked) electronic copy of the model with an adequate user interface and sufficient accompanying information to facilitate understanding of the model, what it does, and how it works should be made available to decision-makers upon request (under conditions of strict confidentiality and protection of property rights), both for review and to permit an analysis of uncertainty to be undertaken using the decision-maker's data and assumptions.

Disclosure of Relationships

Funding and reporting arrangements should be stated as part of the economic evaluation, or in a letter of authorship accompanying the evaluation. The disclosure should include a list of all key participants in the study and their contributions. It should also list the sponsor of the study and indicate whether the sponsor had any review or editing rights regarding either the analysis or the reporting of the evaluation.

Declarations of any conflicts of interest (financial or otherwise) by the researchers, or a declaration that no conflicts exist, should accompany the evaluation. Guidelines for the declaration of conflicts of interest and a declaration template can be found in the *Guidelines for Authors of CADTH Health Technology Reports*.[149]

Appendix 1: Standard Reporting Format

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

It is suggested that economic evaluations follow this format as much as possible, although, 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 Conclusion sections should be written so they can be understood by a non-technical reader.

Other reporting tools and formats have been published[150,151] and may be used as an alternative.

Preface

- List of authors, affiliations, and a description of contributions
- Acknowledgements
- Disclosure of funding and reporting relationships, study sponsor, contractual arrangements, autonomy of researchers, and publication rights; declaration of conflicts of interest (guidelines and a declaration template can be found in the *Guidelines for Authors of CADTH Health Technology Reports*.[149]

Executive Summary

The Executive Summary should be no more than three pages long and written in non-technical language. It should include the following sections:

- Issue: a statement about the policy or economic issue, or reason for evaluating the technology
- Objectives and Decision Problem
- 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

Objectives

Description of Issue(s) Addressed in the Economic Evaluation

- Set the scene for the reader, and include reasons for the analysis (e.g., policy issues, funding or costs implications, issues of competing technologies).

Statement of Decision Problem

- Define the decision problem, state it in an answerable form, and make it relevant for the target audience.
- Define the target population(s) and comparators.
- State the perspective and any non-reference case perspectives.
- Identify the primary target audience and possible secondary audiences.

Background

General Comments on Condition

- State the condition and population group(s) being studied.
- List the etiology, pathology, diagnosis, risk factors, and prognosis (if relevant).
- Describe the epidemiological (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 for the condition in Canada. Refer to clinical practice guidelines (if relevant). Include a description or comparison of interventions for the indication.

Technology Description

For drugs, state the brand and generic names, dosage form, route of administration, recommended dosage, duration of treatment, therapeutic classification, and 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 the adverse events, contraindications, cautions, and warnings.
- Describe the setting for the technology if relevant (e.g., hospital-based).
- Give the unit costs of the comparators.

Regulatory Status

- List the approved indication(s) in Canada that are the topic of the study, including applicable population and subgroups, and date of approval.
- List any additional approved indication(s) in Canada.
- Include the regulatory status and approved indications in other countries.

Review of Economic Evidence

- Discuss existing economic studies that address the same technology and similar decision problem.
- Where validated economic models have been developed that address similar decision problems, indicate whether the approach can be used to address the current decision problem.

Methods

As outlined in these *Guidelines*, report how each element of the economic evaluation has been handled.

Type of Economic Evaluation
- Describe the cost-utility analysis (CUA).
- Provide justification if another type of evaluation was conducted.

Target Population
- Describe target population(s) and the care setting for the intervention or expected use.
- Describe and justify the basis for the 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, or costs).
- If no subgroup analyses were conducted, provide justification for why they were not required.

Comparators
- Describe and justify selected interventions; relate choice of comparators to the study population, and the local context or practice.

Perspective
- State and justify the perspective(s) used in the analysis (e.g., public payer, societal).
- Where additional non-reference case perspectives are considered, indicate how they address the decision problem.
- Describe how other types of variability (e.g., variation in costs or practice patterns) were analyzed, and provide justification.

Time Horizon
- Indicate the time horizon(s) used in the analysis, and provide justification.

Discount Rate
- Indicate the discount rates used for costs and outcomes, and provide justification.

Modelling
a) Modelling Considerations

- Describe the model structure: description of the scope, structure, and assumptions made (with justification); inclusion of a model diagram 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 the use of different validation methods (e.g., comparison with other models). Results from validation exercises can be attached as appendices.

b) Data Considerations

- List data sources and justify assumptions. This may include details about epidemiological factors, such as prevalence or incidence of the condition.
- Describe any statistical analyses.

Effectiveness

a) Evidence of Efficacy and Effectiveness

- Give details about the evidence on efficacy and how this relates to the estimates of effectiveness used in the analysis (if lengthy, place in an appendix).
- For clinical studies, report on PICOS (participants, intervention, comparator or control, outcomes, and study design).
- Describe adverse events, where important and relevant.
- Indicate sources of information (e.g., trials, a meta-analysis or network meta-analysis, literature).

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 parameter estimates (e.g., short-term to long-term outcomes, surrogate to final outcomes). Describe the strength of the evidence for the relationships and links.

Measurement and Valuation of Outcomes

- Identify, measure, and value all relevant outcomes, including important adverse events, for each intervention.
- Give the sources of information, assumptions, and justification.
- Indicate the health-related quality-of-life measurement used and include justification (a copy of the instrument may be included in an appendix). Describe the methods for eliciting preferences and the population measured.
- Include other outcomes that were considered but rejected (with rationale).

Resource Use and Costs

- Identify, measure, and value all resources included in the analysis.
- Report the costing methods used (e.g., patient level).
- Classify resources into categories relevant to the perspective (e.g., relevant agencies comprising the public payer).
- Report resource quantities and unit costs separately.
- 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, data, and assumptions.

Uncertainty

- Identify sources of uncertainty in the analysis.
- Clearly delineate the reference case analysis from non-reference case analyses.
- Provide sources and justification for the probability distributions used in probabilistic analyses. State the number of Monte Carlo iterations.
- For scenario analyses, state the values and assumptions tested; provide sources and justification for each.

Equity

- State equity assumptions (e.g., all quality-adjusted life-years [QALYs] are equal).
- Identify the equity-relevant characteristics of the main subgroups that may benefit, or be adversely affected by, the technology, and describe how they were analyzed.

Results

Study Parameters

- Report and justify the sources of information used for input parameters.
- Report the probability distributions for all parameters.
- Provide the input values for study parameters with reference to the sources of information in a table.
- List all assumptions.

Analysis and Results

- Present all analyses in a step-by-step fashion so the calculations can be replicated, if desired. This includes outcomes, costs, and quality of life by comparator.
- Present the analysis first in a disaggregated fashion, showing all components 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.
- Report aggregate costs and outcomes over the time horizon by perspective.
- Show the components of the incremental cost-effectiveness ratio (ICER) numerator (mean costs of each intervention), and denominator (mean outcomes of each intervention).
- For outcomes, express in natural units first, then translate into alternative units such as QALYs or monetary benefits.
- Provide tables of results in appendices; a visual display of results is encouraged.
- Results should be reported in this manner for all relevant subgroups.

Results of Scenario Analyses

- Report the results for scenarios analyzed.
- Describe the interpretation of the results in relation to the reference case.
- Indicate the results of analyses for types of variability (e.g., variation in costs or practice patterns).

Model Validation

- Provide details on the process for validating the model.
- Where details of the exercise are relevant for inclusion, consider including this as an appendix to the economic evaluation.
- Where other economic studies have been reviewed, compare the methods and results of these studies with the present study.

Discussion

Summary of Results

- Critically appraise and interpret the main findings of the analysis in the context of all reasonable interventions.
- Address the intervention's place in practice, based on the evidence.
- Discuss the uncertainty of the results and the key drivers of results.
- Discuss the trade-off between outcomes and costs.

Study Limitations

- Discuss key limitations and issues concerning the analysis, including methodological limitations and issues, validity of assumptions, strength of the data, and relationships or links used in the model. Describe whether the data and methods used may bias the analysis in favour of any intervention.

Generalizability

- Comment on the generalizability or relevance of results, and the validity of the data and model for 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.

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 due to geographic location or 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 and for whom there is no alternative treatment (e.g., "rule of rescue")?

Future Research

- Identify knowledge gaps and areas for further research that are relevant to Canada.

Conclusions

- Address the decision problem(s).
- Summarize the main findings of the study: aggregate impact, uncertainty about the results, appropriate uses for the intervention (e.g., population subgroups), and any caveats.

References

Appendices

- Depending on practical considerations and amount of material, include the following in the appendices: a table of data sources; data collection forms, questionnaires, and instruments; a diagram of the model structure; step-by-step details of analyses, including intermediate results; tables of results; and visual presentations of results (e.g., figures, graphs).

Appendix 2: Reference Case

Table 4 presents the recommended guidance for the reference case analysis. There may be situations where the analysis will differ from that presented below. Researchers should identify any such discrepancies and provide justification based on the decision problem for any additional non-reference case analyses.

Table 4: Recommended Guidance for Reference Case Analysis

Section	Guidance
Decision Problem	Specify the interventions, setting, perspective, costs, outcomes, time horizon and target population for the evaluation.
Types of Evaluations	Conduct a cost-utility analysis (CUA) capturing health outcomes in terms of quality-adjusted life-years (QALYs).
Target Population	Identify the population(s) for which the interventions are to be used.
	Conduct stratified analysis where distinct subgroups are identified.
Comparators	Compare all relevant interventions, including "current care" (i.e., the intervention(s) presently used in a Canadian context).
Perspective	Adopt a publicly funded health care payer perspective.
Time Horizon	Select a time horizon that is long enough to capture all relevant differences in the future costs and outcomes associated with the interventions being compared.
Discounting	Discount costs and outcomes at a rate of 1.5% per year.
Measurement and Valuation of Health	Identify, measure, and value all relevant health outcomes based on the perspective of the publicly funded health care payer.
	Use health preferences that reflect the general Canadian population.
	Obtain health preferences from an indirect method of measurement that is based on a generic classification system.
Resource Use and Costs	Identify, measure, and value all relevant resources and costs based on the perspective of the publicly funded health care payer.
	Estimate Canadian resources and costs using data that reflect the jurisdiction(s) of interest.
Analysis	Derive expected values of costs and outcomes for each intervention through probabilistic analysis, incorporating potential correlation among parameters.
	Where distinct subgroups are identified within the target population, conduct a stratified analysis and present results for each subgroup.
	Calculate incremental cost-effectiveness ratios (ICERs).
	For evaluations with more than 2 interventions, calculate ICERs sequentially.
Uncertainty	Address methodological uncertainty by comparing the reference case results to those from a non-reference case analysis.
	Summarize decision uncertainty, using such items as cost-effectiveness acceptability curves (CEACs) and cost-effectiveness acceptability frontiers (CEAFs).
	Use scenario analysis to address structural uncertainty.
	When a value-of-information analysis is undertaken, summarize the value of additional information using the expected value of perfect parameter information and the population expected value of perfect parameter information.
Equity	Weight all outcomes equally regardless of the characteristics of people receiving, or affected by the intervention in question.

References

1. Claxton K, Paulden M, Gravelle H, Brouwer W, Culyer AJ. Discounting and decision making in the economic evaluation of health-care technologies. Health Econ. 2011 Jan;20(1):2-15.

2. Drummond MF, Sculpher MJ, Claxton K, Stoddart GL, Torrance GW. Methods for the economic evaluation of health care programmes. 4th ed. Oxford: Oxford University Press; 2015.

3. Paulden M, Galvanni V, Chakraborty S, Kudinga B, McCabe C. Discounting and the evaluation of health care programs [Internet]. Ottawa: CADTH; 2016 Mar. [cited 2016 May 20]. Available from: https://www.cadth.ca/sites/default/files/pdf/CP0008_Economic_Evaluation_Guidelines_Discount_Rate_Report.pdf

4. Baron J. A theory of social decisions. Philadelphia: University of Pennsylvania; 2011 Mar 25.

5. Palmer S, Torgerson DJ. Economic notes: definitions of efficiency. BMJ [Internet]. 1999 Apr 24 [cited 2017 Feb 1];318(7191):1136. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1115526

6. Claxton KP, Sculpher MJ, Culyer AJ. Mark versus Luke? Appropriate methods for the evaluation of public health interventions [Internet]. York (GB): University of York, Centre for Health Economics; 2007 Nov. [cited 2016 May 20]. (CHE research paper; no. 31). Available from: http://www.york.ac.uk/media/che/documents/papers/researchpapers/rp31_evaluation_of_public_health_interventions.pdf

7. Whitehurst DG, Norman R, Brazier JE, Viney R. Comparison of contemporaneous EQ-5D and SF-6D responses using scoring algorithms derived from similar valuation exercises. Value Health [Internet]. 2014 Jul [cited 2016 Jun 16];17(5):570-7. Available from: http://www.sciencedirect.com/science/article/pii/S1098301514017811

8. Cookson R. Willingness to pay methods in health care: a sceptical view. Health Econ. 2003 Nov;12(11):891-4.

9. Sculpher M. Subgroups and heterogeneity in cost-effectiveness analysis. Pharmacoeconomics. 2008;26(9):799-806.

10. Phelps CE. Good technologies gone bad: how and why the cost-effectiveness of a medical intervention changes for different populations. Med Decis Making. 1997 Jan;17(1):107-17.

11. Coyle D, Buxton MJ, O'Brien BJ. Stratified cost-effectiveness analysis: a framework for establishing efficient limited use criteria. Health Econ. 2003 May;12(5):421-7.

12. Espinoza MA, Manca A, Claxton K, Sculpher MJ. The value of heterogeneity for cost-effectiveness subgroup analysis: conceptual framework and application. Med Decis Making [Internet]. 2014 Nov [cited 2017 Mar 2];34(8):951-64. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4232328

13. Mauskopf J, Samuel M, McBride D, Mallya UG, Feldman SR. Treatment sequencing after failure of the first biologic in cost-effectiveness models of psoriasis: a systematic review of published models and clinical practice guidelines. Pharmacoeconomics [Internet]. 2014 Apr [cited 2017 Mar 2];32(4):395-409. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3964298/

14. Goodrich K, Kaambwa B, Al-Janabi H. The inclusion of informal care in applied economic evaluation: a review. Value Health. 2012 Sep;15(6):975-81.

15. Hoefman RJ, van Exel J, Brouwer W. How to include informal care in economic evaluations. Pharmacoeconomics. 2013 Dec;31(12):1105-19.

16. Fox-Rushby JA, Cairns J, editors. Economic evaluation. Maidenhead (NY): Open University Press; 2005 Nov.

17. Gold MR, Siegel JE, Russell LB, Weinstein MC, editors. Cost-effectiveness in health and medicine. New York: Oxford University Press; 1996.

18. Zhuang J, Liang Z, Lin T, de Guzman F. Theory and practice in the choice of social discount rate for cost-benefit analysis: a survey [Internet]. Mandaluyong City (PH): Asian Development Bank; 2007. [cited 2016 Oct 24]. (ERD working paper series; no. 94). Available from: https://www.adb.org/sites/default/files/publication/28360/wp094.pdf

19. National health expenditure trends, 1975 to 2015 [Internet]. Ottawa: Canadian Institute for Health Information; 2015 Oct. [cited 2016 May 20]. Available from: https://www.cihi.ca/sites/default/files/document/nhex_trends_narrative_report_2015_en.pdf

20. History of health and social transfers [Internet]. Ottawa: Department of Finance Canada; 2014 Dec 15. [cited 2016 May 25]. Available from: http://www.fin.gc.ca/fedprov/his-eng.asp

21. Galvani V, Behnamian A. A comparative analysis of the returns on provincial and federal Canadian bonds [Internet]. Edmonton (AB): University of Alberta; 2009 Jan. [cited 2016 May 20]. (Working paper no. 2009-07). Available from: http://www.ualberta.ca/~econwps/2009/wp2009-07.pdf

22. Research guides [Internet]. New York: Columbia University. Bloomberg help guide; 2016 [cited 2016 May 24]. Available from: http://library.columbia.edu/subject-guides/business/bloomberg.html

23. Inflation [Internet]. Ottawa: Bank of Canada; 2016. [cited 2016 May 24]. Available from: http://www.bankofcanada.ca/core-functions/monetary-policy/inflation/

24. Renewal of the inflation-control target: background information -- October 2016 [Internet]. Ottawa: Bank of Canada; 2016 Oct. [cited 2017 Feb 1]. Available from: http://www.bankofcanada.ca/wp-content/uploads/2016/10/background_nov11.pdf

25. Annual demographic estimates: Canada, provinces and territories [Internet]. Ottawa: Statistics Canada; 2015. [cited 2016 May 24]. (Catalogue no. 91-215-X). Available from: http://www.statcan.gc.ca/pub/91-215-x/91-215-x2015000-eng.htm

26. Damodaran A. What is the riskfree rate? A search for the basic building block [Internet]. New York: New York University, Stern School of Business; 2008 Dec. [cited 2016 May 20]. Available from: http://117.4iranian.com/uploads/Estimating%20Riskfree%20Rates_881.pdf

27. Caro JJ, Briggs AH, Siebert U, Kuntz KM, ISPOR-SMDM Modeling Good Research Practices Task Force. Modeling good research practices--overview: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force-1. Med Decis Making. 2012 Sep;32(5):667-77.

28. Roberts M, Russell LB, Paltiel AD, Chambers M, McEwan P, Krahn M, et al. Conceptualizing a model: a report of the ISPOR-SMDM modeling good research practices task force-2. Med Decis Making. 2012 Sep;32(5):678-89.

29. Siebert U, Alagoz O, Bayoumi AM, Jahn B, Owens DK, Cohen DJ, et al. State-transition modeling: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force-3. Med Decis Making. 2012 Sep;32(5):690-700.

30. Karnon J, Stahl J, Brennan A, Caro JJ, Mar J, Moller J. Modeling using discrete event simulation: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force-4. Med Decis Making. 2012 Sep;32(5):701-11.

31. Pitman R, Fisman D, Zaric GS, Postma M, Kretzschmar M, Edmunds J, et al. Dynamic transmission modeling: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force--5. Value Health [Internet]. 2012 Sep [cited 2016 Jun 16];15(6):828-34. Available from: http://www.sciencedirect.com/science/article/pii/S1098301512016518

32. Briggs AH, Weinstein MC, Fenwick EA, Karnon J, Sculpher MJ, Paltiel AD, et al. Model parameter estimation and uncertainty: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force--6. Value Health [Internet]. 2012 Sep [cited 2016 Jun 16];15(6):835-42. Available from: http://www.sciencedirect.com/science/article/pii/S1098301512016592

33. Eddy DM, Hollingworth W, Caro JJ, Tsevat J, McDonald KM, Wong JB, et al. Model transparency and validation: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force-7. Med Decis Making. 2012 Sep;32(5):733-43.

34. Briggs A, Claxton K, Sculpher M. Decision modelling for health economic evaluation. Oxford (GB): Oxford University Press; 2006.

35. Stahl JE. Modelling methods for pharmacoeconomics and health technology assessment: an overview and guide. Pharmacoeconomics. 2008;26(2):131-48.

36. Marshall DA, Burgos-Liz L, Ijzerman MJ, Osgood ND, Padula WV, Higashi MK, et al. Applying dynamic simulation modeling methods in health care delivery research-the SIMULATE checklist: report of the ISPOR simulation modeling emerging good practices task force. Value Health [Internet]. 2015 Jan [cited 2016 Jun 16];18(1):5-16. Available from: http://www.sciencedirect.com/science/article/pii/S1098301514047640

37. Brennan A, Chick SE, Davies R. A taxonomy of model structures for economic evaluation of health technologies. Health Econ. 2006 Dec;15(12):1295-310.

38. Barton P, Bryan S, Robinson S. Modelling in the economic evaluation of health care: selecting the appropriate approach. J Health Serv Res Policy. 2004 Apr;9(2):110-8.

39. Griffin S, Claxton K, Hawkins N, Sculpher M. Probabilistic analysis and computationally expensive models: necessary and required? Value Health [Internet]. 2006 Jul [cited 2016 Jun 16];9(4):244-52. Available from: http://www.sciencedirect.com/science/article/pii/S1098301510602768

40. Ethgen O, Standaert B. Population- versus cohort-based modelling approaches. Pharmacoeconomics. 2012 Mar;30(3):171-81.

41. Dias S, Welton NJ, Sutton AJ, Ades AE. Evidence synthesis for decision making 5: the baseline natural history model. Med Decis Making [Internet]. 2013 Jul [cited 2016 Oct 24];33(5):657-70. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3704201

42. Cooper NJ, Sutton AJ, Achana F, Welton NJ. Use of network meta-analysis to inform clinical parameters in economic evaluations [Internet]. Ottawa: CADTH; 2015 Jun. [cited 2016 Apr 5]. Available from: https://www.cadth.ca/sites/default/files/pdf/RFP%20Topic-%20Use%20of%20Network%20Meta-analysis%20to%20Inform%20Clinical%20Parameters%20in%20Economic%20Evaluations.pdf

43. Ara R, Wailoo A. NICE DSU technical support document 12: the use of health state utility values in decision models [Internet]. Sheffield (United Kingdom): University of Sheffield; 2011. [cited 2016 Apr 29]. Available from: http://www.nicedsu.org.uk/TSD12%20Utilities%20in%20modelling%20FINAL.pdf

44. Morton A, Adler AI, Bell D, Briggs A, Brouwer W, Claxton K, et al. Unrelated future costs and unrelated future benefits: reflections on NICE Guide to the Methods of Technology Appraisal. Health Econ. 2016 Aug;25(8):933-8.

45. Bojke L, Claxton K, Bravo-Vergel Y, Sculpher M, Palmer S, Abrams K. Eliciting distributions to populate decision analytic models. Value Health [Internet]. 2010 Aug [cited 2016 Jun 16];13(5):557-64. Available from: http://www.sciencedirect.com/science/article/pii/S1098301510600964

46. Bojke L, Soares M. Decision analysis: eliciting experts' beliefs to characterize uncertainties. In: Culyer AJ, editor. Encyclopedia of health economics. Amsterdam: Elsevier; 2014.

47. Karnon J, Vanni T. Calibrating models in economic evaluation: a comparison of alternative measures of goodness of fit, parameter search strategies and convergence criteria. Pharmacoeconomics. 2011 Jan;29(1):51-62.

48. Stout NK, Knudsen AB, Kong CY, McMahon PM, Gazelle GS. Calibration methods used in cancer simulation models and suggested reporting guidelines. Pharmacoeconomics [Internet]. 2009 [cited 2016 Jun 16];27(7):533-45. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2787446

49. McHaney R. Computer simulation: a practical perspective. San Diego (CA): Academic Press; 1991.

50. Thesen A, Travis LE. Simulation for decision making. St. Paul (MN): West Publishing; 1992.

51. Law AM, Kelton DW. Simulation modeling and analysis. 4th ed. Boston: McGraw-Hill; 2007.

52. Kopec JA, Fines P, Manuel DG, Buckeridge DL, Flanagan WM, Oderkirk J, et al. Validation of population-based disease simulation models: a review of concepts and methods. BMC Public Health [Internet]. 2010 [cited 2016 May 3];10:710. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3001435

53. Drummond MF, Barbieri M, Wong JB. Analytic choices in economic models of treatments for rheumatoid arthritis: what makes a difference? Med Decis Making. 2005 Sep;25(5):520-33.

54. Berry DA, Cronin KA, Plevritis SK, Fryback DG, Clarke L, Zelen M, et al. Effect of screening and adjuvant therapy on mortality from breast cancer. N Engl J Med. 2005 Oct 27;353(17):1784-92.

55. Mount Hood 4 Modeling Group. Computer modeling of diabetes and its complications: a report on the Fourth Mount Hood Challenge Meeting. Diabetes Care. 2007 Jun;30(6):1638-46.

56. Zauber AG, Lansdorp-Vogelaar I, Knudsen AB, Wilschut J, van Ballegooijen M, Kuntz KM. Evaluating test strategies for colorectal cancer screening: a decision analysis for the U.S. Preventive Services Task Force. Ann Intern Med [Internet]. 2008 Nov 4 [cited 2016 Aug 10];149(9):659-69. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2731975

57. Ip S, Paulus JK, Balk EM, Dahabreh IJ, Avendano EE, Lau J, et al. Role of single group studies in Agency for Healthcare Research and Quality comparative effectiveness reviews [Internet]. Rockville (MD): Agency for Healthcare Research and Quality (AHRQ); 2013 Jan. [cited 2016 May 24]. (AHRQ publication; no. 13-EHC007-EF). Available from: http://www.ncbi.nlm.nih.gov/books/NBK121314/pdf/Bookshelf_NBK121314.pdf

58. O'Mahony JF, Newall AT, van Rosmalen J. Dealing with time in health economic evaluation: methodological issues and recommendations for practice. Pharmacoeconomics [Internet]. 2015 Dec [cited 2017 Feb 1];33(12):1255-68. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4661216

59. Ballou DP, Pater HL. Modeling completeness versus consistency tradeoffs in information decision contexts. IEEE Trans Knowl Data Eng. 2003;15(1):241-44.

60. Prevost TC, Abrams KR, Jones DR. Hierarchical models in generalized synthesis of evidence: an example based on studies of breast cancer screening. Stat Med. 2000 Dec 30;19(24):3359-76.

61. McCarron CE, Pullenayegum EM, Thabane L, Goeree R, Tarride JE. The importance of adjusting for potential confounders in Bayesian hierarchical models synthesising evidence from randomised and non-randomised studies: an application comparing treatments for abdominal aortic aneurysms. BMC Med Res Methodol [Internet]. 2010 [cited 2016 Jun 16];10:64. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2916004

62. Turner RM, Spiegelhalter DJ, Smith GC, Thompson SG. Bias modelling in evidence synthesis. J R Stat Soc Ser A Stat Soc [Internet]. 2009 Jan [cited 2017 Mar 2];172(1):21-47. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2667303

63. Schmitz S, Adams R, Walsh C. Incorporating data from various trial designs into a mixed treatment comparison model. Stat Med. 2013 Jul 30;32(17):2935-49.

64. Verde PE, Ohmann C. Combining randomized and non-randomized evidence in clinical research: a review of methods and applications. Res Synth Methods. 2015 Mar;6(1):45-62.

65. Robb MA, McInnes PM, Califf RM. Biomarkers and surrogate endpoints: developing common terminology and definitions. JAMA. 2016 Mar 15;315(11):1107-8.

66. Fleming TR, DeMets DL. Surrogate end points in clinical trials: are we being misled? Ann Intern Med. 1996 Oct 1;125(7):605-13.

67. Collett D. Modelling survival data in medical research. 2nd ed. Boca Raton (FL): Chapman & Hall/CRC Press; 2003.

68. Latimer NR. Survival analysis for economic evaluations alongside clinical trials--extrapolation with patient-level data: inconsistencies, limitations, and a practical guide. Med Decis Making. 2013 Aug;33(6):743-54.

69. Guyot P, Ades AE, Ouwens MJ, Welton NJ. Enhanced secondary analysis of survival data: reconstructing the data from published Kaplan-Meier survival curves. BMC Med Res Methodol [Internet]. 2012 [cited 2016 Aug 10];12:9. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3313891

70. Latimer N. NICE DSU technical support document 14: survival analysis for economic evaluations alongside clinical trials-extrapolation with patient-level data [Internet]. Sheffield (United Kingdom): University of Sheffield; 2011. [cited 2016 Aug 10]. Available from: http://www.nicedsu.org.uk/NICE%20DSU%20TSD%20Survival%20analysis.updated%20March%202013.v2.pdf

71. Davies C, Briggs A, Lorgelly P, Garellick G, Malchau H. The "hazards" of extrapolating survival curves. Med Decis Making. 2013 Apr;33(3):369-80.

72. Committee for Medicinal Products for Human Use (CHMP). Reflection paper on methodological issues in confirmatory clinical trials planned with an adaptive design [Internet]. London: European Medicines Agency; 2007 Oct 18. [cited 2016 Jul 7]. Available from: http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2009/09/WC500003616.pdf

73. Schipper J, Clinch JJ, Olweny CL. Quality of life studies: definitions and conceptual issues. In: Spilker B, editor. Quality of life and pharmacoeconomics in clinical trials. 2nd ed. Philadelphia (PA): Lippincott-Raven; 1996. p. 11-24.

74. von Neumann J, Morgenstern O. Theory games and economic behavior. Princeton (NJ): Princeton University Press; 1994.

75. Brazier J. Measuring and valuing health benefits for economic evaluation. Oxford (United Kingdom): Oxford University Press; 2007.

76. Blumenschein K, Johannesson M. An experimental test of question framing in health state utility assessment. Health Policy. 1998 Sep;45(3):187-93.

77. Dolan P, Gudex C, Kind P, Williams A. Valuing health states: a comparison of methods. J Health Econ. 1996 Apr;15(2):209-31.

78. About EQ-5D [Internet]. Rotterdam: EuroQol; 2016. [cited 2016 Apr 29]. Available from: http://www.euroqol.org/about-eq-5d.html

79. Health Utilities Inc (HUInc) [Internet]. Dundas (ON): HUInc. 2015 [cited 2016 Apr 29]. Available from: http://www.healthutilities.com/

80. SF-6D [Internet]. Sheffield (United Kingdom): University of Sheffield; 2016. [cited 2016 Apr 29]. Available from: https://www.shef.ac.uk/scharr/sections/heds/mvh/sf-6d

81. Bansback N, Tsuchiya A, Brazier J, Anis A. Canadian valuation of EQ-5D health states: preliminary value set and considerations for future valuation studies. PLoS One [Internet]. 2012 [cited 2016 Oct 24];7(2):e31115. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3273479/

82. Xie F, Pullenayegum E, Gaebel K, Bansback N, Bryan S, Ohinmaa A, et al. A time trade-off-derived value set of the EQ-5D-5L for Canada. Med Care [Internet]. 2016 Jan [cited 2016 Apr 28];54(1):98-105. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4674140

83. Brazier J, Longworth L. NICE DSU technical support document 8: an introduction to the measurement and valuation of health for NICE submissions [Internet]. Sheffield (United Kingdom): University of Sheffield; 2011. [cited 2016 Oct 24]. Available from: http://www.nicedsu.org.uk/TSD8%20Introduction%20to%20MVH_final.pdf

84. Peasgood T, Ward SE, Brazier J. Health-state utility values in breast cancer. Expert Rev Pharmacoecon Outcomes Res. 2010 Oct;10(5):553-66.

85. Goodwin E, Green C. A systematic review of the literature on the development of condition-specific preference-based measures of health. Appl Health Econ Health Policy. 2016 Apr;14(2):161-83.

86. Brazier J, Rowen D. NICE DSU technical support document 11: alternatives to EQ-5D for generating health state utility values [Internet]. Sheffield (United Kingdom): University of Sheffield, ScHARR, Decision Support Unit; 2011 Mar. [cited 2016 Oct 12]. Available from: http://www.nicedsu.org.uk/TSD11%20Alternatives%20to%20EQ-5D_final.pdf

87. Longworth L, Rowen D. NICE DSU technical support document 10: the use of mapping methods to estimate health state utility values [Internet]. Sheffield (United Kingdom): University of Sheffield; 2011. [cited 2016 Oct 24]. Available from: http://www.nicedsu.org.uk/TSD%2010%20mapping%20FINAL.pdf

88. Longworth L, Rowen D. Mapping to obtain EQ-5D utility values for use in NICE health technology assessments. Value Health [Internet]. 2013 Jan [cited 2016 Jun 16];16(1):202-10. Available from: http://www.sciencedirect.com/science/article/pii/S1098301512041617

89. Pennington B, Davis S. Mapping from the Health Assessment Questionnaire to the EQ-5D: the impact of different algorithms on cost-effectiveness results. Value Health [Internet]. 2014 Dec [cited 2016 Jun 16];17(8):762-71. Available from: http://www.sciencedirect.com/science/article/pii/S1098301514047573

90. Ara R, Wailoo AJ. Estimating health state utility values for joint health conditions: a conceptual review and critique of the current evidence. Med Decis Making. 2013 Feb;33(2):139-53.

91. Hu B, Fu AZ. Predicting utility for joint health states: a general framework and a new nonparametric estimator. Med Decis Making. 2010 Sep;30(5):E29-E39.

92. Basu A, Dale W, Elstein A, Meltzer D. A linear index for predicting joint health-states utilities from single health-states utilities. Health Econ. 2009 Apr;18(4):403-19.

93. Nord E, Daniels N, Kamlet M. QALYs: some challenges. Value Health [Internet]. 2009 Mar [cited 2016 Jun 16];12 Suppl 1:S10-S15. Available from: http://www.sciencedirect.com/science/article/pii/S1098301510600472

94. Fowler FJ Jr, Cleary PD, Massagli MP, Weissman J, Epstein A. The role of reluctance to give up life in the measurement of the values of health states. Med Decis Making. 1995 Jul;15(3):195-200.

95. Whitehead SJ, Ali S. Health outcomes in economic evaluation: the QALY and utilities. Br Med Bull. 2010;96:5-21.

96. Jacobs P, Budden A, Lee KM. Guidance document for the costing of health care resources in the Canadian setting [Internet]. 2nd ed. Ottawa: CADTH; 2016 Mar. [cited 2016 May 19]. Available from: https://www.cadth.ca/sites/default/files/pdf/CP0009_CADTHCostingGuidance.pdf

97. Walter E, Zehetmayr S. Guidelines of health economic evaluation. Consensus paper [Internet]. Vienna: Institute for Pharmaeconomic Research; 2006 Apr. [cited 2016 Oct 24]. Available from: http://www.ispor.org/peguidelines/source/Guidelines_Austria.pdf

98. Briggs AH, Drummond MF. Reporting guidelines for health economic evaluations: BMJ guidelines for author and peer reviewers of economic submissions. In: Moher D, Altman D, Schulz K, Simera I, Wager E, editors. Guidelines for reporting health research: a user's manual. Chichester (United Kingdom): John Wiley & Sons; 2014. Chapter 28.

99. Mogyorosy A, Smith P. The main methodological issues in costing health care services: a literature review [Internet]. York (United Kingdom): University of York, Alcuin College, Centre for Health Economics; 2005. [cited 2016 Oct 24]. (CHE research paper; no. 7). Available from: https://www.york.ac.uk/che/pdf/rp7.pdf

100. Redelmeier DA, Tan SH, Booth GL. The treatment of unrelated disorders in patients with chronic medical diseases. N Engl J Med [Internet]. 1998 May 21 [cited 2016 Mar 29];338(21):1516-20. Available from: http://www.nejm.org/doi/full/10.1056/NEJM199805213382106

101. Wodchis WP, Bushmeneva K, Nikitovic M, McKillop I. Guidelines on person-level costing using administrative databases in Ontario [Internet]. Toronto: Health System Performance Research Network; 2013 May. [cited 2017 Mar 10]. (Working paper series; vol. 1). Available from: http://www.hsprn.ca/uploads/files/Guidelines_on_PersonLevel_Costing_May_2013.pdf

102. de Oliveira C, Bremner KE, Pataky R, Gunraj N, Chan K, Peacock S, et al. Understanding the costs of cancer care before and after diagnosis for the 21 most common cancers in Ontario: a population-based descriptive study. CMAJ Open [Internet]. 2013 Jan [cited 2016 Jun 17];1(1):E1-E8. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3985946

103. Krajden M, Kuo M, Zagorski B, Alvarez M, Yu A, Krahn M. Health care costs associated with hepatitis C: a longitudinal cohort study. Can J Gastroenterol [Internet]. 2010 Dec [cited 2016 Mar 29];24(12):717-26. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3004444

104. Wijeysundera HC, Machado M, Wang X, Van Der Velde G, Sikich N, Witteman W, et al. Cost-effectiveness of specialized multidisciplinary heart failure clinics in Ontario, Canada. Value Health. 2010 Dec;13(8):915-21.

105. Nosyk B, Lima V, Colley G, Yip B, Hogg RS, Montaner JS. Costs of health resource utilization among HIV-positive individuals in British Columbia, Canada: results from a population-level study. Pharmacoeconomics [Internet]. 2015 Mar [cited 2017 Mar 10];33(3):243-53. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4677778

106. Coyle D, Lee KM. The problem of protocol driven costs in pharmacoeconomic analysis. Pharmacoeconomics. 1998 Oct;14(4):357-63.

107. Rittenhouse BE. Exorcising protocol-induced spirits: making the clinical trial relevant for economics. Med Decis Making. 1997 Jul;17(3):331-9.

108. Drummond M, Barbieri M, Cook J, Glick HA, Lis J, Malik F, et al. Transferability of economic evaluations across jurisdictions: ISPOR Good Research Practices Task Force report. Value Health. 2009 Jun;12(4):409-18.

109. Ramsey S, Willke R, Briggs A, Brown R, Buxton M, Chawla A, et al. Good research practices for cost-effectiveness analysis alongside clinical trials: the ISPOR RCT-CEA Task Force report. Value Health [Internet]. 2005 Sep [cited 2016 Jun 17];8(5):521-33. Available from: http://www.sciencedirect.com/science/article/pii/S1098301510604123

110. Morris S, Appleby J, Parkin D, Spencer A. Economic analysis in health care. Chichester (United Kingdom): Wiley; 2012.

111. Culyer AJ. The dictionary of health economics. 3rd ed. Cheltenham (United Kingdom): Edward Elgar Publishing; 2014.

112. Lipscomb J, Yabroff KR, Brown ML, Lawrence W, Barnett PG. Health care costing: data, methods, current applications. Med Care. 2009 Jul;47(7 Suppl 1):S1-S6.

113. Johnston K, Buxton MJ, Jones DR, Fitzpatrick R. Assessing the costs of healthcare technologies in clinical trials. Health Technol Assess [Internet]. 1999 [cited 2016 Oct 24];3(6):1-76. Available from: https://njl-admin.nihr.ac.uk/document/download/2003801

114. Wijeysundera HC, Wang X, Tomlinson G, Ko DT, Krahn MD. Techniques for estimating health care costs with censored data: an overview for the health services researcher. Clinicoecon Outcomes Res [Internet]. 2012 [cited 2016 Aug 11];4:145-55. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3377439

115. Krol M, Brouwer W, Rutten F. Productivity costs in economic evaluations: past, present, future. Pharmacoeconomics. 2013 Jul;31(7):537-49.

116. Brouwer WB, Koopmanschap MA, Rutten FF. Patient and informal caregiver time in cost-effectiveness analysis. a response to the recommendations of the Washington Panel. Int J Technol Assess Health Care. 1998;14(3):505-13.

117. Lensberg BR, Drummond MF, Danchenko N, Despiégel N, Francois C. Challenges in measuring and valuing productivity costs, and their relevance in mood disorders. Clinicoecon Outcomes Res [Internet]. 2013 [cited 2016 Jan 8];5:565-73. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3836685

118. Krol M, Brouwer W. How to estimate productivity costs in economic evaluations. Pharmacoeconomics. 2014 Apr;32(4):335-44.

119. Tang K. Estimating productivity costs in health economic evaluations: a review of instruments and psychometric evidence. Pharmacoeconomics. 2015 Jan;33(1):31-48.

120. Brouwer WB, Koopmanschap MA, Rutten FF. Productivity costs in cost-effectiveness analysis: numerator or denominator: a further discussion. Health Econ. 1997 Sep;6(5):511-4.

121. Krol M, Brouwer WB, Severens JL, Kaper J, Evers SM. Productivity cost calculations in health economic evaluations: correcting for compensation mechanisms and multiplier effects. Soc Sci Med. 2012 Dec;75(11):1981-8.

122. Nicholson S, Pauly MV, Polsky D, Sharda C, Szrek H, Berger ML. Measuring the effects of work loss on productivity with team production. Health Econ. 2006 Feb;15(2):111-23.

123. Robinson R. Cost-benefit analysis. BMJ [Internet]. 1993 Oct 9 [cited 2016 Jun 17];307(6909):924-6. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1679054

124. van den Hout WB. The value of productivity: human-capital versus friction-cost method. Ann Rheum Dis. 2010 Jan;69 Suppl 1:i89-i91.

125. Koopmanschap MA, Rutten FF, van Ineveld BM, van Roijen L. The friction cost method for measuring indirect costs of disease. J Health Econ. 1995 Jun;14(2):171-89.

126. Tranmer JE, Guerriere DN, Ungar WJ, Coyte PC. Valuing patient and caregiver time: a review of the literature. Pharmacoeconomics. 2005;23(5):449-59.

127. Hanly P, Timmons A, Walsh PM, Sharp L. Breast and prostate cancer productivity costs: a comparison of the human capital approach and the friction cost approach. Value Health [Internet]. 2012 May [cited 2016 Jun 17];15(3):429-36. Available from: http://www.sciencedirect.com/science/article/pii/S1098301512000125

128. 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.

129. Johannesson M. Avoiding double-counting in pharmacoeconomic studies. In: Mallarkey G, editor. Economic evaluation in healthcare. Auckland (NZ): Adis International; 1999. p. 155-7.

130. Sendi P, Brouwer W. Leisure time in economic evaluation: theoretical and practical considerations. Expert Rev Pharmacoecon Outcomes Res. 2004 Feb;4(1):1-3.

131. Oostenbrink JB, Koopmanschap MA, Rutten FF. Standardisation of costs: the Dutch Manual for Costing in economic evaluations. Pharmacoeconomics. 2002;20(7):443-54.

132. Thompson KM, Graham JD. Going beyond the single number: using probabilistic risk assessment to improve risk management. Hum Ecol Risk Assess. 1996;2(4):1008-34.

133. Claxton K, Sculpher M, McCabe C, Briggs A, Akehurst R, Buxton M, et al. Probabilistic sensitivity analysis for NICE technology assessment: not an optional extra. Health Econ. 2005 Apr;14(4):339-47.

134. Briggs AH. Handling uncertainty in cost-effectiveness models. Pharmacoeconomics. 2000 May;17(5):479-500.

135. McCarron CE, Pullenayegum EM, Marshall DA, Goeree R, Tarride JE. Handling uncertainty in economic evaluations of patient level data: a review of the use of Bayesian methods to inform health technology assessments. Int J Technol Assess Health Care. 2009;25(4):546-54.

136. Briggs AH, Goeree R, Blackhouse G, O'Brien BJ. Probabilistic analysis of cost-effectiveness models: choosing between treatment strategies for gastroesophageal reflux disease. Med Decis Making. 2002 Jul;22(4):290-308.

137. O'Hagan A. Uncertain judgements: eliciting experts' probabilities. Hoboken (NJ): John Wiley & Sons; 2006.

138. Dias S, Sutton AJ, Welton NJ, Ades AE. Evidence synthesis for decision making 6: embedding evidence synthesis in probabilistic cost-effectiveness analysis. Med Decis Making [Internet]. 2013 Jul [cited 2016 Oct 24];33(5):671-8. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3704202

139. Stinnett AA, Mullahy J. Net health benefits: a new framework for the analysis of uncertainty in cost-effectiveness analysis. Med Decis Making. 1998 Apr;18(2 Suppl):S68-S80.

140. Karlsson G, Johannesson M. The decision rules of cost-effectiveness analysis. Pharmacoeconomics. 1996 Feb;9(2):113-20.

141. Bojke L, Claxton K, Sculpher M, Palmer S. Characterizing structural uncertainty in decision analytic models: a review and application of methods. Value Health. 2009 Jul;12(5):739-49.

142. Coyle D, Oakley J. Estimating the expected value of partial perfect information: a review of methods. Eur J Health Econ. 2008 Aug;9(3):251-9.

143. McCabe C, Awotwe I, Paulden M, Hall P. One-way sensitivity analysis for stochastic cost effectiveness analysis: conditional expected incremental net benefit [Internet]. Edmonton: PACEOMICS; 2017. [cited 2017 Mar 2]. (PACEOMICS working paper; PWP 2017_01). Available from: http://paceomics.org/wp-content/uploads/2017/03/PACEOMICS-working-paper-2017_01.pdf

144. Donaldson C, Gerard K. Economics of health care financing : the visible hand. 2nd ed. Basingstoke (United Kingdom): Palgrave Macmillan; 2005.

145. Culyer AJ. Equity of what in healthcare? Why the traditional answers don't help policy--and what to do in the future. Healthc Pap. 2007;8 Spec No:12-26.

146. Williams A. Intergenerational equity: an exploration of the 'fair innings' argument. Health Econ. 1997 Mar;6(2):117-32.

147. Wailoo A, Tsuchiya A, McCabe C. Weighting must wait: incorporating equity concerns into cost-effectiveness analysis may take longer than expected. Pharmacoeconomics. 2009;27(12):983-9.

148. Baltussen R, Leidl R, Ament A. The impact of age on cost-effectiveness ratios and its control in decision making. Health Econ. 1996 May;5(3):227-39.

149. Guidelines for authors of CADTH health technology assessment reports [Internet]. Ottawa: CADTH; 2003 May. [cited 2017 Mar 2]. Available from: https://www.cadth.ca/guidelines-authors-cadth-health-technology-assessment-reports-0

150. Husereau D, Drummond M, Petrou S, Carswell C, Moher D, Greenberg D, et al. Consolidated Health Economic Evaluation Reporting Standards (CHEERS) statement. Cost Eff Resour Alloc [Internet]. 2013 [cited 2016 Aug 11];11(1):6. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3607888

151. Drummond MF, Jefferson TO. Guidelines for authors and peer reviewers of economic submissions to the BMJ. The BMJ Economic Evaluation Working Party. BMJ [Internet]. 1996 Aug 3 [cited 2016 Aug 11];313(7052):275-83. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2351717