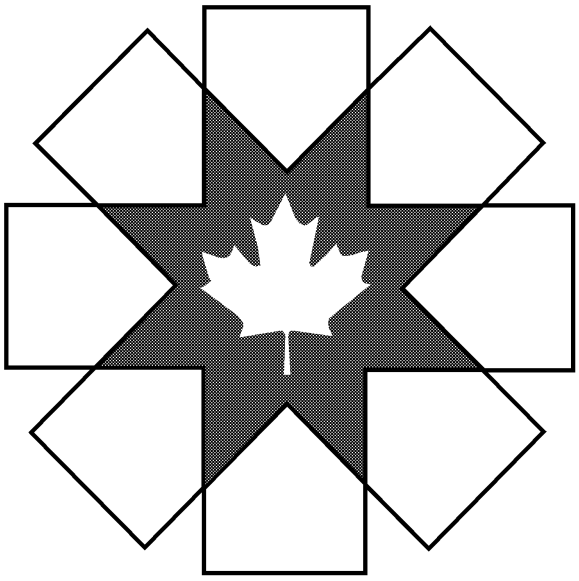
Canadian Coordinating Office for Health Technology Assessment

Office Canadien de Coordination de l'Évaluation des Technologies de la Santé

GUIDELINES FOR ECONOMIC EVALUATION OF PHARMACEUTICALS: CANADA

2nd Edition
November 1997



Cite as: Canadian Coordinating Office for Health Technology Assessment. *Guidelines for economic evaluation of pharmaceuticals: Canada*. 2nd ed. Ottawa: Canadian Coordinating Office for Health Technology Assessment (CCOHTA); 1997.

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

Legal Deposit - 1997
National Library of Canada
ISBN: 1-895561-45-0

CCOHTA is a non-profit organization, funded by the federal, provincial and territorial governments. It was established to encourage the appropriate use of health technology by influencing decision-makers through the scientific evaluation of medical procedures, devices and drugs.

To obtain copies of publications please contact:

CCOHTA Publications
110-955 Green Valley Crescent
Ottawa, Ontario, Canada K2C 3V4
Telephone: (613) 226-2553
Facsimile: (613) 226-5392
Email: pubs@ccohta.ca

Or download full-text from Website:
http://www.ccohta.ca

Aussi disponible en français:

Lignes directrices pour l'évaluation économique des produits pharmaceutiques: Canada.

FOREWORD TO THE SECOND EDITION

Pharmacoeconomics is the application of the methods of economic evaluation of health care programs to interventions involving pharmaceutical products. The techniques and these Guidelines are, on the whole, transferable to the evaluation of any health technology; and the Committee encourages the use of these Guidelines for other health care interventions.

In contrast to the first edition, the focus of the second edition of these Guidelines is to provide assistance to the >doer' of the analysis or study, for the purposes of providing standardized and reliable information for the target audience - the >user'. To this end, changes to the Guidelines have focused on key areas where there have been significant developments in theory and/or methodology. At times, these developments have been considerable and/or controversial, in which case the Committee has tried to address all sides of the issue. Any expansion of a given topic should not be interpreted as implying a preference on the part of the Guidelines towards that methodology. Investigators should be assured that emphasis on evolving methodologies is purely for purposes of explanation and clarification.

These Guidelines attempt to be as prescriptive and helpful as possible, and to provide opportunities for creativity and innovation. There is reference to a partner document, *A Guidance Document for the Costing Process* (CCOHTA, 1996), which provides more details on costing methodology. The second edition should be useful for directing study design, as well as providing a template for the final reports of evaluations. However, investigators should note that, first, each study is unique and considerable thought is required just to implement the Guidelines intelligently. Second, creativity can be exercised to justify particular deviations, or omissions, from the recommended set of methods. Third, there is unlimited scope in adding additional analyses to those suggested in the Guidelines. Finally, creative exploration in methodological research is required to investigate alternative pharmacoeconomic methods for use in updating the Guidelines themselves. To this end, it is suggested that users should be flexible in terms of taking into account deviations from the methodologies suggested herein, provided that these deviations are appropriately justified by the doers.

The second edition of the *Guidelines for Economic Evaluation of Pharmaceuticals: Canada* has undergone extensive revisions both in format and content. Like the first edition, a great many individuals representing academia, government, private industry and the health care professions have provided insight and suggestions for change. There were three major reasons for revising the first edition of the Guidelines. First, one of the commitments made by the Canadian Coordinating Office for Health Technology Assessment (CCOHTA) was to review and update the Guidelines. Second, many individuals and organizations have used the Guidelines, thus providing a wealth of experience and development in this field of study. Third, consensus has developed in some areas which were previously controversial.

The first task of the Guidelines Review Committee was to survey all individuals who had received the first edition of the Guidelines since its publication. From the analysis the Committee identified key concerns regarding methodology, organization and comprehensiveness. Although the Guidelines provide further clarification regarding some issues, they are neither a textbook on pharmacoeconomics nor are they intended to be rigid, since many methodologic issues remain unresolved.

The major changes to this edition are as follows:

- Major rewrites have been done in the following sections: cost-benefit analysis, equity, utilities, treatment comparators, effectiveness, preferences, costing, uncertainty, and portability.

- Several sections have an annotated bibliography which recommends representative readings on the subject area. The sections to which this feature was added were those where the Committee felt major controversies existed. These references are subject to change over time, and future editions of these Guidelines will include additional sections with bibliographies provided.
- The references have been expanded and updated.
- A new reporting structure format has been developed to address the needs of users.

As with any project, there are many individuals who have been instrumental in developing the final product. The Committee worked diligently in the development of these Guidelines. In particular, Dr. George Torrance, who was the principal writer of the first edition of the Guidelines, and Dr. Judith Glennie, who has been editor and contributing author for this second edition, deserve special credit for their efforts. The Committee is also indebted to the following individuals who provided their expertise in the development of these Guidelines: Dr. Bernie O'Brien, Dr. Peter Tugwell, Dr. Lawrence Joseph, Mr. John Hoar, Dr. Jayanti Mukherjee and Mr. David Moher, as well as the many individuals around the globe who responded to our surveys and provided constructive comments and suggestions. This document incorporates most of their comments; however, CCOHTA and the editor take responsibility for its form, content and any errors or omissions.

CCOHTA continues to encourage comments and suggestions regarding this document from individuals and organizations alike. They may be forwarded to:

Director, Pharmaceuticals and Extramural Research
CCOHTA
110-955 Green Valley Crescent
Ottawa, Ontario,
Canada K2C 3V4
Tel: 613-226-2553 Fax: 613-226-5392
Comments may also be sent electronically via the webmaster, annieh@ccohta.ca

Nick Otten
Chairman, Guideline Review Committee

OCTOBER 1997

GUIDELINE REVIEW COMMITTEE MEMBERS

Mr. Jean-François Baladi (Abbott Inc. Canada) Ms. Catherine Berka (Glaxo-Wellcome Inc.)

Dr. Judith Glennie (J.L. Glennie Consulting Inc.) Ms. Eleanor Hubbard (N.S. Dept. of Health)

Dr. Devidas Menon (Institute of Pharmaco-Economics) Dr. Nick Otten (CCOHTA)

Dr. Marc Riviere (Aeterna Labs) Dr. George Torrance (McMaster University)

FOREWORD TO THE FIRST EDITION

Australia was the first country to develop and implement guidelines for the economic evaluation of pharmaceuticals. Draft guidelines were released in 1990, revised in 1992, and are currently going through the process of second revision. In Canada, the process for developing these guidelines began when the Province of Ontario issued draft guidelines for comment in the Fall of 1991. During 1992 it was determined that it would be useful to develop a set of Canadian guidelines, that each Province could adopt, with or without modifications, as they saw fit. That process led to the Canadian Collaborative Workshop on Pharmacoeconomics, held at Le Chantecler, Sainte-Adèle, Quebec, June 21-22, 1993, attended by 73 participants representing the major Canadian stakeholders plus selected international experts (Schubert, 1993). The Proceedings of the workshop were published in November 1993. Under the direction of George Torrance, McMaster University, a steering committee comprised of David Blaker, Health Canada; Eleanor Hubbard, Nova Scotia Department of Health; Allan Detsky, University of Toronto; Theresa Firestone, Ontario Ministry of Health; Wendy Kennedy, Université de Montréal; Richard Konchak, Patented Medicine Prices Review Board; Devidas Menon, CCOHTA; François Schubert, Burroughs Wellcome; and Peter Tugwell, University of Ottawa produced four drafts of these Guidelines culminating with the original publication.

The responsibility for this document now resides with CCOHTA. Due to the emerging science of health economics, its increasing importance in health care resource allocation, and our need to ascertain the utility of the Guidelines under a variety of circumstances, our goal is to ensure this document remains current. To this end a process will be developed to evaluate the Guidelines and provide a forum for stakeholder input. This will be conducted on an annual basis. Should significant changes be approved, new versions will be published.

This document is intended for a diverse audience ranging from academia to industry to government. We encourage your comments and suggestions. They may be forwarded to:

Coordinator, Pharmaceutical Assessment
CCOHTA
110-955 Green Valley Crescent
Ottawa, Ontario
Canada K2C 3V4

NOVEMBER 1994

TABLE OF CONTENTS

FOREWORD TO THE SECOND EDITION i

FOREWORD TO THE FIRST EDITION iii

I. SUMMARY OF GUIDELINE STATEMENTS 1

II. DESIGN AND METHODOLOGY FOR ECONOMIC EVALUATIONS 6

1. INTRODUCTION 6
1.0 Introductory Remarks 6
2.0 Role of Studies 7
2.1 General Considerations 7
2.2 Target Audience 8
Guideline 1. Target Audience 8
3.0 Timing of Studies 9
Guideline 2. Timing of Studies 9
4.0 Management of Studies 10
Guideline 3. Management of Studies 10

2. METHODS 11
5.0 Research Design 11
5.1 Incremental Analysis 11
Guideline 4. Incremental and Total Analysis 11
5.2 Total Costs and Outcomes 11
5.3 Analytic Technique 11
Guideline 5. Analytic Technique 11
5.4 Indications by Target Population 14
Guideline 6. Indications 14
Methodologic Tips on Subgroup Analysis 15
5.5 Treatment Comparator 16
Guideline 7. Treatment Comparator 16
5.6 Perspective 18
Guideline 8. Perspective 18
5.7 Analytic Horizon 18
Guideline 9. Analytic Horizon 18
5.8 Assumptions 19
Guideline 10. Assumptions 19

6.0 Measuring and Reporting Outcomes 20
6.1 Efficacy Versus Effectiveness 20
Guideline 11. Efficacy Versus Effectiveness 20
Finding the Data 20
The Use of Meta-analysis 21
Methodologic Tips on Meta-analysis 22
6.2 Health-Related Quality of Life as an Outcome 23
Guideline 12. Health-Related Quality of Life (HRQOL) 23
General Concepts 23
Specific Measures 24
Generic Measures 24
Preference-based Measures 24
Modelling HRQOL 24
6.3 Outcomes for Cost-Utility Analysis 25
Guideline 13. Outcomes for Cost-Utility Analysis 25
QALYs 26
Other Approaches 27
Generic/Specific HRQOL Measures versus Preferences 28
Preferences 28
Direct Measurement 28
Indirect Measurement 29
Summary 30
6.4 Outcomes for Cost-Benefit Analysis 31
Guideline 14. Outcomes for Cost-Benefit Analysis 31
6.5 Source of Preferences 34
Guideline 15. Source of Preferences 34
6.6 Equity 36
Guideline 16. Equity 36
6.7 Discounting Future Outcomes 38
Guideline 17. Discounting Future Outcomes 38
7.0 Measuring and Reporting Costs 39
7.1 Cost Identification 39
Guideline 18. Cost Identification 39
Unrelated Costs 39
Future Health Care Costs 40
Costs of Lost Time 40
Summary 42
7.2 Cost Measurement 43
Guideline 19. Cost Measurement (Resources Used) 43
7.3 Cost Valuation 44
Guideline 20. Cost Valuation (Unit Prices) 44
7.4 Discounting Future Costs 45
Guideline 21. Discounting Future Costs 45
8.0 Dealing with Uncertainty 46
Guideline 22. Dealing With Uncertainty 46
8.1 Statistical Approaches 47
8.2 Sensitivity Analysis Approaches 47

3. RESULTS AND DISCUSSION 50
9.0 Reporting the Analysis and Results 50
Guideline 23. Reporting Results 50
9.1 General Recommendations 50
9.2 The Portability of Economic Evaluations 51
Guideline 24. Portability of Economic Evaluations 51
Methologic Tips on Portability of Economic Evaluations 53
10.0 Disclosure of Relationships 55
Guideline 25. Disclosure of Relationships 55

III. STANDARDIZED REPORTING STRUCTURE 56
Reporting Structure 56

IV. GLOSSARY OF TERMS 62

22. QUALITY ASSURANCE TIPS 67
Study Documents 67
Data Collection 67
Data Entry and Data Management 67
Data Analysis 67
Quality Control 68

VI. REFERENCES 69

VII. INDEX 84

I. SUMMARY OF GUIDELINE STATEMENTS

Guideline 1. Target Audience

The primary target audience (decision-maker) for the study must be identified. Secondary target audiences (if applicable) should also be listed.

Guideline 2. Timing of Studies

Pharmacoeconomic studies can be undertaken at any point in a product's life cycle. Suggestions are given on the timing of studies and on the types of decision-making situations which call for the presentation of economic evidence.

Guideline 3. Management of Studies

There are no restrictions on who can do studies. All studies should, however, be consistent with these Guidelines.

Guideline 4. Incremental and Total Analysis

Costs and effects must be reported as increments (that is, as differences between two alternatives) and as totals. Increments must be used in the pharmacoeconomic evaluation.

Guideline 5. Analytic Technique

If all consequences are essentially identical between the drug and the relevant comparators, a cost-minimization analysis (CMA) is adequate. In other instances, a cost-consequence analysis (CCA) is required plus one or more of the following: cost-effectiveness analysis (CEA), cost-utility analysis (CUA), and cost-benefit analysis (CBA). Consistent with the desire to permit broad comparisons, CUA or CBA are preferred. Researchers should present the data using a variety of techniques, to maximize the information content and to contribute to the development of these methodologies.

Guideline 6. Indications

The study must clearly specify the target population for the drug. Any investigations of patient subgroups, disease subtypes, severity levels, comorbidity groups, etc., should be clearly identified by an explicit hypothesis in the study protocol. Economic evaluation should be performed overall and, data permitting, for those subgroups that were identified in the protocol for their possible differential effectiveness, costs and/or preferences.

Guideline 7. Treatment Comparator

The drug treatment should be compared with both existing practice and minimum practice. The relevant comparators may be other drugs, other medical care such as surgery, or even no treatment. Existing practice would either be the single most prevalent clinical practice (if there is one that is dominant), or it could be current practice weighted by market share. Minimum practice would

normally be either the lowest cost comparator that is more effective than placebo, or the do-nothing alternative, as appropriate. In addition to these two formal comparators, all other reasonable alternative therapies should be at least discussed in the report.

Guideline 8. Perspective

All studies should report from a comprehensive societal perspective. That perspective should be transparently broken down into those of other relevant viewpoints, including that of the primary decision-maker. A financial impact analysis from the viewpoint of the primary decision-maker may also be undertaken, if requested, but technically is a budgeting exercise and does not constitute part of the economic evaluation.

Guideline 9. Analytic Horizon

Every effort should be made to extend the analytic horizon to capture all relevant outcomes. When modelled data are needed to meet this requirement, the structure and rationale of the model must be presented.

Guideline 10. Assumptions

A comprehensive listing of assumptions and associated rationale must be contained within the explanation of the methodology for the analysis.

Guideline 11. Efficacy Versus Effectiveness

Ideally, pharmacoeconomic studies should report on drug effectiveness rather than efficacy. Because effectiveness data are generally not available, appropriate modelling techniques based on sound pharmacoepidemiology (e.g. using epidemiologic studies to estimate patient compliance with therapy in the real world) are permissible. All assumptions used in such extrapolation techniques must be stated explicitly and thoroughly tested with sensitivity analysis.

Guideline 12. Health-Related Quality of Life (HRQOL)

If HRQOL is being included in a prospective study as an outcome, it is normally advisable to include, where possible and feasible, one instrument from each of the following three types: specific measures, generic profiles, and preference-based measures. Any drug product that demonstrates improved effectiveness over its comparator(s) and impacts on a patient's HRQOL should probably be evaluated for this outcome using these tools.

Guideline 13. Outcomes for Cost-Utility Analysis

Both quantity of life (survival) and HRQOL results should be reported separately, and the method of combining the two described in a transparent manner. The current recommended method for the primary analysis is to combine quantity and quality of life using quality-adjusted life years (QALYs). Alternatives, such as disability-adjusted life years (DALYs), healthy years equivalents (HYEs) or saved young life equivalents (SAVEs) may be useful as secondary analyses in some studies. To be

suitable for use as quality weights for calculating quality-adjusted life years, scores must be based on preferences and measured on an interval scale where dead has a score of 0 and healthy has a score of 1.

Direct preference measurements can be undertaken with various instruments. Analysts should select an instrument that suits the problem, and should justify their selection. Alternatively, preferences can be determined indirectly using one of the major systems available. Here again, analysts should select a system that suits the problem, and should justify their selection.

Guideline 14. Outcomes for Cost-Benefit Analysis

The human capital approach (HCA) to assigning values to outcomes in CBA is incomplete because it focuses primarily on lost work time. If this approach is used, measures taken to overcome these shortcomings should be clearly described.

While realizing the incomplete and experimental nature of this method, theoretically, the preferred approach to the assignment of values to outcomes in CBA is contingent valuation as a means of eliciting an assessment of willingness to pay (WTP). If this approach is used, the investigator(s) should be explicit with regards to the assumptions and methods utilized. In addition, the measures taken to reduce bias and an outline of the scope tests carried out to determine validity should be clearly described.

Guideline 15. Source of Preferences

The appropriate source of preferences depends on the use of the analysis and the viewpoint. For provincial drug plans, which are tax supported, the appropriate viewpoint is societal and the appropriate source of preferences for outcomes is the informed general public. The three major systems for the indirect determination of preferences (QWB, HUI, EQ-5D) are all scored based on preferences from the general public and, so, are suitable.

Analysts who wish to measure preferences directly should ideally do so on the general public, suitably informed. Patients in a study may, however, be a reasonable proxy for the informed general public, especially when they are providing preferences for hypothetical states. Preferences should be based on scientifically sound measurements. Investigators undertaking direct measurements of preferences must justify their source of subjects, and describe the exact population from which the preferences were derived and the precise methods of measurement.

Guideline 16. Equity

All equity assumptions, whether implicit or explicit, must be highlighted in any analysis intended for use in the resource allocation process. Results should be presented using equal weights for all lives, life-years or QALYs, but the presentation should be sufficiently transparent to make it feasible for decision-makers to substitute different weights. Analysts should identify which groups of individuals would be the main beneficiaries if the program were implemented.

Guideline 17. Discounting Future Outcomes

Future outcomes should be discounted at the same rate as costs. The base case discount rate is 5% per year. This rate must be varied in a sensitivity analysis, with a discount rate of 0% (no discounting) at minimum. Analysts should also consider using a 3% rate for comparability with future studies. When it is believed the analysis should differentiate between discount rates for outcomes and costs, these results should be presented as a supplementary analysis and the relevance fully explained.

Guideline 18. Cost Identification

A probability tree of the therapeutic pathway which describes all relevant downstream events should be provided, when appropriate. From the societal viewpoint, cost items that should be included are all direct health care costs, social services costs, spillover costs on other sectors such as education, and costs that fall on the patient and family. Cost items that should be excluded are those not relevant to the therapeutic pathway such as those not related to the treatment being evaluated, costs relevant only to the clinical trial, and transfer payments such as sickness pay, unemployment insurance and welfare payments.

When relevant, lost time should be documented and reported as part of the description of the impact of the intervention. If HRQOL is an outcome measure in the study, some lost time will likely contribute to changes in HRQOL. Depending on the viewpoint, some lost time will represent a real cost in terms of lost resources and should be included as a cost item, but should also be tested with sensitivity analysis.

Guideline 19. Cost Measurement (Resources Used)

Resources used in treatment must first be described in natural (non-dollar) units. All resource utilization data derived from international trials must be validated for Canadian practice.

***Guideline 20. Cost Valuation (Unit Prices*)**

Economic definitions of costs must be used and the concept of opportunity cost recognized. Investigators performing analyses in the Canadian setting should refer to the CCOHTA *Guidance Document for the Costing Process* for further direction regarding costing issues.

Guideline 21. Discounting Future Costs

As with future outcomes, all studies must discount future costs at an initial rate of 5% per year in the base case. This rate must be varied in a sensitivity analysis, with a discount rate of 0% (no discounting) at minimum. Analysts should also consider using a 3% rate for comparability with future studies. If differential discount rates are to be used for outcomes and costs, then the results should be presented as a supplementary analysis and the relevance fully explained.

Guideline 22. Dealing With Uncertainty

All studies must clearly address the issue of uncertainty (whether it arises from sampling error or from assumptions) and justify the methods used. Sampling errors can be dealt with by making use of confidence intervals. In addition, for each important assumption, alternative plausible assumptions must be included. Investigators are encouraged to use approaches such as Monte Carlo simulation, which varies all factors simultaneously.

Guideline 23. Reporting Results

All results must be reported in disaggregated detail first, with aggregations and the use of value judgements (e.g. preference scores) being introduced into the presentation as late as possible. A probability tree of clinical outcomes should be provided for the relevant alternatives. Detailed technical reports, with patient confidentiality protected, should be made available to decision-makers. Reports should either follow the standardized reporting structure or be linked to it.

Guideline 24. Portability of Economic Evaluations

The portability of an economic evaluation is an issue which should be considered during the development of the study, as well as during the interpretation and dissemination of study results. Consideration must be given to two aspects of the applicability of the analysis to the local setting. The first aspect is the distinction between efficacy and effectiveness. The second aspect is the validity of transferring results (i.e. economic, clinical and humanistic) from one country or health care jurisdiction to another. These considerations are especially important when working in the context of multinational, multi-centre trials.

Guideline 25. Disclosure of Relationships

Funding and reporting relationships must be clearly described. The investigators must have independence regarding methodological considerations at all stages of the study, and must have the right of publication in the journal of their choice.

II. DESIGN AND METHODOLOGY FOR ECONOMIC EVALUATIONS

1. INTRODUCTION

1.0 Introductory Remarks

Pharmacoeconomics concerns the application of the methods of economic evaluation of health care programs to interventions involving pharmaceutical products (Feeny *et al.*, 1986; Detsky and Naglie, l990; Freund and Dittus, 1992; Laupacis, *et al.*, 1992; Eisenberg, *et al.*, 1994; Bootman, *et al.*, 1996; Drummond, *et al.*, 1997). The purpose of the methods, and the studies, is to help inform programmatic decision-making regarding the appropriateness and availability of health care interventions including drugs. Results of such programmatic decision-making (e.g. **formulary** listings, clinical guidelines, appropriate prescribing practices) will often impact on decision-making regarding treatments for individual patients. However, it is important to note that the methods of economic evaluation are generally not directly targeted at bedside decision-making.

These Guidelines focus specifically on pharmacoeconomics, and should be useful for directing study design and for providing a template for final reports. As pharmacoeconomics is simply a specific application of economic evaluation, the Guidelines should also be of use to researchers and decision-makers involved with evaluations of other kinds of health care interventions. The principles of technology assessment can be applied to process as well as intervention evaluations, and will be useful in the overall movement towards disease management in health care.

Note that terms defined in the Glossary (Section IV) are **bolded** the first time they appear in the text.

2.0 Role of Studies

2.1 General Considerations

Pharmacoeconomic studies compare the **costs** and **consequences** of pharmaceutical products with relevant alternatives. These studies are pertinent to the decision-making process when trying to balance additional costs associated with one alternative over another, versus their respective differences in clinical outcome. If there were unlimited resources for health care treatments in general, and drugs in particular, there would be no need for pharmacoeconomic studies. One would simply support all therapies that did more good than harm; that is, all treatments and drugs that lead to net improvements or **effectiveness**, where effectiveness combines the good from the treatment and the harm from possible side effects and adverse events.

Resources are limited, however, and are inadequate to support all treatments that demonstrate effectiveness in clinical trials. Thus, supporting one effective treatment or drug means that some other effective therapy cannot be supported. The good that could have been done with this other treatment or drug which is not supported is the **opportunity cost** of the first choice; that is, there is an opportunity cost associated with rejecting effective treatments. The overall technical goal of pharmacoeconomics is to identify treatments and drugs which may be worthy of support, such that the overall good that is done is maximized (or equivalently, the opportunity costs incurred are minimized) within the constrained resources available.

It should be noted that all purchasers, including the provincial Ministries of Health, have always considered costs in addition to effectiveness and safety. Those cost analyses have, however, almost exclusively been limited to comparisons of unit prices of new versus old products. In the l990s, it is clear that the economic arguments must be expanded beyond those of simple unit cost comparisons. The latter leave out offsets that might be achieved by newer and better drugs; and certainly do not provide a measure of **value** in terms of improved clinical outcomes, such as reducing side-effects, better performance, cure, palliation or prevention of disease.

The need to use more complete information in decision-making is reflected in the growing number of pharmacoeconomic guideline documents in the literature (Jacobs, *et al.*, 1995; Drummond and Jefferson, 1996; Ikeda, *et al.*, 1996). The Canadian Guidelines are intended to provide assistance to the "doer" of the analysis or study, for the purposes of providing standardized and reliable information for the target audience - the "user". To this end, these Guidelines provide direction in areas where there is methodologic agreement; and examine issues in key areas wherein there have been significant developments in theory and/or methodology.

Pharmacoeconomic studies in their proper role are used to inform decision-making, not to replace it. The studies are not to be used in a thoughtless, mechanistic fashion. They do not replace hard thinking, careful consideration, good judgement and common sense. When properly used and properly qualified, they provide essential information as input into the decision-making process. They are not the only input, however; other considerations such as justice, **equity**, access, choice and process factors also come into play.

There are a variety of decision-making situations where pharmacoeconomic studies can play a useful role:

1) Research and development decisions by a firm. Using the best estimates available and acknowledging wide bands of uncertainty, pharmacoeconomic studies can be undertaken for drugs under development to identify promising areas for research and development investment. As the drugs move through the development process, the studies can be updated with increasingly more precise estimates to monitor the development of the drug with respect to its projected pharmacoeconomic performance. Such studies could be used for "go/no go" decisions at critical points in the drug's development.

2) Pricing decisions. Both the firm and government regulators (such as the Patented Medicine Prices Review Board [PMPRB] and provincial formulary managers) could use pharmacoeconomic evidence to help to establish an appropriate price for a product. This price may be higher, or lower, than it would have been in the absence of such studies; but the advantage is that the price will be based on a more rational, open and transparent process.

3) Formulary decisions. Provinces, hospitals, insurers, and other payers could use pharmacoeconomic evidence in determining new listings, whether to continue listings, or what portion of the cost of a given drug product they are willing to pay.

4) Clinical guidelines for prescribing decisions. Those preparing clinical guidelines for providers need to consider not only the **efficacy** and effectiveness of clinical alternatives (including drugs), but also their cost-effectiveness. High quality pharmacoeconomic studies can provide important input into the development of such guidelines.

5) Post-marketing surveillance. There is a need to monitor continuously the performance of drugs and to update periodically the pharmacoeconomic studies with accumulating evidence from actual utilization experience, including unanticipated effects, both beneficial and adverse. These updated studies, in turn, can be used to update decisions on pricing, formulary listings and clinical guidelines.

2.2 Target Audience

Guideline 1. Target Audience

The primary target audience (decision-maker) for the study must be identified. Secondary target audiences (if applicable) should also be listed.

The primary users of pharmacoeconomic studies will be the decision-makers; for example, those involved in the decisions listed in Section 2.1. Secondary users, however, include a wide variety of other interested parties who could be consumers of pharmacoeconomic studies. These include health care workers of all types, academics, patients, the press, the general public, consumers' associations, and professional organizations such as medical, nursing and pharmacy groups.

3.0 Timing of Studies

Guideline 2. Timing of Studies

Pharmacoeconomic studies can be undertaken at any point in a product's life cycle. Suggestions are given on the timing of studies and on the types of decision-making situations which call for the presentation of economic evidence.

Pharmacoeconomic studies can be undertaken at any point in a product's life cycle. The timing of studies depends upon the needs of the users of studies, as described in Section 2. Early studies, during the research and development (R&D) phase of the drug, may be undertaken by the company to guide future (i.e. Phase III and IV) R&D decisions and marketing planning. Phase III studies may have a particular role in pricing and formulary decisions early in the product's life cycle. These studies may also play a role in initial clinical and prescribing guidelines.

Phase IV pharmacoeconomic studies (post-marketing) would contribute by updating previous studies on the basis of the new effectiveness data; and provide better evidence regarding utilization and adverse events. These post-marketing review studies could be scheduled on the basis of time (3 to 5 years after the product is marketed), or on the basis of "trigger" events (changes in medical practice, costs, comparator[s], or the emergence of new adverse or beneficial events). Post-marketing studies of drugs may also be initiated by organizations other than the pharmaceutical firm. This would be particularly true for the development of clinical guidelines and possibly for formulary decision-making in particular diseases, or with respect to particular categories of drugs or for particular patient groups.

4.0 Management of Studies

Guideline 3. Management of Studies

There are no restrictions on who can do studies. All studies should, however, be consistent with these Guidelines.

Who should do pharmacoeconomic studies? Because the objective is to achieve high-quality unbiased information, merely assigning the responsibility for studies to one group or another, for example to academics, does not guarantee the achievement of the objective. Moreover, at a practical level, there are not enough well-trained individuals currently in any group to be able to handle all the studies required.

At some time in the future it may be appropriate for pharmacoeconomic studies to follow the same basic approach as clinical studies; that is, for a research protocol to be submitted in advance to a central agency for comment, approval and sign-off. At the present time this is not possible. It would be desirable, however, to be able to provide independent reviews of protocols when requested. The important principle to be followed is that the research protocol should be consistent with the Guidelines. It is not important who designed the protocol; it may be designed within the pharmaceutical industry, within the academic community, by a consulting firm, or through a mixture of these approaches. (See also Section 10 on Disclosure of Relationships.)

Every effort should be made to publish study results in peer-reviewed scientific journals. Publication bias with regard to negative study results could be avoided by designating a central agency, possibly CCOHTA, to collect all pharmacoeconomic studies that are conducted, not just those that are published; and to maintain a central data base. A link should be established between the persons who undertake the protocol review (if done, as outlined above) and those who review the final study. These reviews (pre and post) should be undertaken by a multi-disciplinary panel of experts with no direct ties to the specific project.

2. METHODS

5.0 *Research Design*

Guideline 4. Incremental and Total Analysis

Costs and effects must be reported as increments (that is, as differences between two alternatives) and as totals. Increments must be used in the pharmacoeconomic evaluation.

5.1 Incremental Analysis

All pharmacoeconomic studies must be comparative and express results in incremental terms. The drug treatment under study must be compared to one or more relevant alternative treatments, which may include a "do nothing" alternative (if clinically relevant). Costs and consequences must be measured as increments; that is, as differences between the two alternatives. Cost-effectiveness ratios, cost-utility ratios and cost-benefit differences (i.e. net cost o**net benefit**) must be based on incremental results, not totals or averages. Further details can be found in a variety of references (Detsky and Naglie, 1990; Bootman, *et al.*, 1996; Drummond, *et al.*, 1997). The methods become more complicated with multiple relevant comparisons (Drummond, *et al.*, 1993).

5.2 Total Costs and Outcomes

In addition to the incremental analysis, it is also useful to report total costs and total outcomes for each alternative (i.e. the drug and each comparator treatments). The totals enable "users" to better appreciate the magnitudes involved, which can, of course, be totally masked by a ratio of two increments. The totals also allow future users to make comparisons with new or different comparator treatments, as part of bringing studies up to date or transferring study results across geographic or practice boundaries.

Guideline 5. Analytic Technique

If all consequences are essentially identical between the drug and the relevant comparators, a cost-minimization analysis (CMA) is adequate. In other instances, a cost-consequence analysis (CCA) is required plus one or more of the following: cost-effectiveness analysis (CEA), cost-utility analysis (CUA), and cost-benefit analysis (CBA). Consistent with the desire to permit broad comparisons, CUA or CBA are preferred. Researchers should present the data using a variety of techniques, to maximize the information content and to contribute to the development of these methodologies.

5.3 Analytic Technique

Although there is considerable overlap among the various analytic techniques that can be used, it is useful to identify the following five methods. Not all of these approaches have been widely used, but conceptually they are distinct and the distinctions are useful in helping to clarify the field.

1) Cost-Minimization Analysis (CMA). Cost-minimization analysis is appropriate when the clinical outcomes (i.e. efficacy and safety) for the drug and the comparator(s) are virtually the same. In such a case, the decision simply revolves around the costs.

2) Cost-Consequence Analysis (CCA). This is a disaggregated type of study that makes the least assumptions and puts the greatest burden on the decision-makers. It is a "Consumer Reports" style of study. The costs and consequences of the drug compared to one or more relevant alternatives are simply listed in disaggregated form (e.g. drug costs, hospital costs, other costs, strokes avoided, minor side-effects, major side effects, etc.). Any weighting of the component factors and aggregation is left to the user of the study.

3) Cost-Effectiveness Analysis (CEA). In cost-effectiveness analysis, the **incremental costs** are compared to the incremental outcomes as measured in physical or natural units. Natural units could range from clinical measures, such as millimeters of mercury blood pressure reduction, through disability days averted, to lives saved, or life-years gained.

4) Cost-Utility Analysis (CUA) Cost-utility analysis refers to a particular form of CEA where the outcomes are measured in terms of **quality-adjusted life years (QALY) gained** QALYs combine changes in quantity and **quality of life** (QOL; mortality and morbidity) into one composite measure which is independent of program or disease. The quality-adjustment factors should reflect aggregated **preferences** of individuals for the outcomes. The factors have been measured directly on patients or the general public, taken from published tables or formulae, or estimated by professional judgement. Readers should beware that not everyone makes the distinction between CEA and CUA which has been made in this document. Some researchers refer to CUA studies as CEA or as **cost per QALY (gained)** studies.

5) Cost-Benefit Analysis (CBA). In cost-benefit analysis the incremental outcomes are expressed in dollar terms, usually using the **contingent valuation** approach of estimating benefits to elicit an assessment of **willingness to pay (WTP)**, so that the overall analysis can be conducted entirely in dollars.

Consistent with the desire to permit broad comparisons, the expression of results in cost-utility or cost-benefit terms is preferred. (See measurement issues in Section 6 on Measuring and Reporting Outcomes.) Cost-minimization analysis is, of course, appropriate in those rare cases where clinical outcomes across alternatives are virtually the same. The expression of results in only cost-effectiveness or cost-consequence terms is acceptable, with justification, for example, when there is no important impact on **health-related quality of life** (**HRQOL**).

A process should be established within each disease category to agree upon standard clinical outcomes that could be used for CCA, CMA, and CEA. Moreover, the outcomes could form the basis for the preference elicitations required in both of CUA and CBA.

CCA is based on the normative premise that the authorized decision-makers should, and can, make the value judgement trade-offs necessary to integrate a disparate list of pros and cons (costs and consequences) of the various alternatives and reach a final decision. The theoretical issue here is whether these decision-makers, whoever they are, are the right source of values and trade-offs across outcomes; or whether a more appropriate source is patients or the general public (see also Section 6.5: Source of Preferences). The practical issue is whether the decision-makers can cope with the cognitive burden of making all the necessary value judgements and tradeoffs. Most of the psychological research into decision-making would suggest that the task exceeds the cognitive processing capacity of individuals (Miller, 1956).

CBA is based on the theoretical foundations of welfare economics and the normative principle of a potential Pareto improvement (Johannesson and Jonsson, 1991). As such, it has the soundest theoretical basis and is the only technique that, at least in theory, allows for comparisons across health and other sectors (e.g. the environment, education, and defense spending). These advantages come at the expense of difficult measurement issues (see Section 6.4).

CEA is based on the methods of constrained optimization (Winston, 1991), and provides a solution that yields the greatest effectiveness for a given cost; or, alternatively, one that achieves a given effectiveness at minimum cost. The difficulty is that the measure of effectiveness used must be appropriate and common to the treatments being compared.

CUA is generally seen as a special case of CEA, in which the measure of effectiveness is quality-adjusted life years (QALY) gained. This has been called the "extra welfare" foundation of CUA (Culyer, 1989 and 1990). Alternatively, CUA can be linked to welfare economics and its normative principles, by using some additional assumptions that, it is argued, are reasonable (Garber and Phelps, 1997).

It is important to note that these various analytic techniques are in no way mutually exclusive. There is a natural hierarchy which requires a CCA to be undertaken first, and for this data to then flow on into a CEA, and/or CUA, and/or CBA. In many cases the additional effort at the margin required to add an additional analytic technique is small. Researchers are encouraged to present the data using a variety of analytic techniques. In any case, the analysis should be presented in a transparent manner, enabling the reader to follow the path from cost-consequence to the other forms of analysis used.

Annotated Bibliography:

Garber AM, Phelps CE. Economic foundations of cost-effectiveness analysis. Journal of Health Economics 1997;16(1):1-31.

Garber and Phelps provide an overview of many of the controversies relating to the application of CEAs. Among other things, they discuss the problem of indirect time-related costs of treatment or benefits, as well as unrelated future medical costs incurred during the years of life extended by an intervention. In addition they suggest that the validity of CEA has never been rigorously established. The authors go on to demonstrate that, in fact, CEA can offer a valid means of choosing between health interventions, based on the theoretical foundations of the von Neumann-Morgenstern utility framework.

5.4 Indications by Target Population

Guideline 6. Indications

The study must clearly specify the target population for the drug. Any investigations of patient subgroups, disease subtypes, severity levels, comorbidity groups, etc., should be clearly identified by an explicit hypothesis in the study protocol. Economic evaluation should be performed overall and, data permitting, for those subgroups that were identified in the protocol for their possible differential effectiveness, costs and/or preferences.

The following directives and caveats apply primarily to prospective economic evaluations. Application of these principles to retrospective studies may be limited to a large extent as, by definition, these studies involve a modelling process using previously collected data. Any subgroup analysis conducted in the context of a modelling study must be based on clinical or economic evaluations which have established their subpopulations on sound subgroup analysis principles (as outlined below).

A pharmacoeconomic study must clearly specify the target population for the drug. Target populations may be defined using baseline epidemiologic features describing the type of patient (e.g. age, gender, socio-economic status), with a specific disease, of a certain severity, with or without other co-morbidities or risk factors, their geographic distribution, usual compliance rates, typical patterns of treatment, and so on. Target population subgroups which are defined based on effectiveness (from previous research), cost and/or preferences may differ in terms of the cost-effectiveness of an intervention used in those subgroups.

While subgroup differences may be important considerations for decision-makers (Baltussen *et al.*, 1996), the precision of the cost-effectiveness estimate may be compromised by inadequate statistical power due to inadequate sample size. If these competing factors can be balanced in the study development phase, then **subgroup analysis** should be investigated. Because a drug may be cost-effective for some subgroups of patients and not for others, it is important to identify clearly the groups under study *a priori* and, when appropriate, to undertake separate analyses for different groups.

The issue of subgroup analysis is a source of particular contention in economic evaluation. There are questions as to whether these analyses are statistically sound; and concerns that recommendations based on subgroup analysis may be misleading and result in harmful clinical or economic decisions (Oxman, 1996). Therefore, caveats must be noted before subgroup analysis is contemplated.

In their discussion of framing and designing cost-effectiveness studies, Torrance and his colleagues (Torrance, Seigel, *et al.*, 1996) make frequent reference to the evaluation of subgroups. The direction given to investigators is clear: subgroups must be definable based on explicitly outlined parameters prior to the study, and a minimum of subgroup analyses should be carried out within a single study. (It is suggested that the research protocol be made available for confirmation of predefined parameters and the number of subgroup analyses.) Where there is *a priori* evidence of the likelihood of differences in outcome(s) for certain subgroups versus a broader population, it is

suggested that data be collected for both groups. If the subgroups differ in any controversial or discriminatory fashion (e.g. differences in time costs between the retired elderly and employed middle-aged individuals), Torrance *et al.* suggest that **sensitivity analysis** be used to demonstrate the effect of group-specific estimates.

There are situations where subgroups emerge in the process of exploratory data evaluation. Such subgroups may serve a useful purpose in hypothesis generation for future studies. Information derived from subgroups discovered *post hoc* must be analyzed using appropriate statistical methods (e.g. adjusting for multiple comparisons; using regression techniques [Yusuf *et al.*, 1991]); and must be interpreted cautiously (i.e. may introduce unintended bias or uncertainty, akin to random variation, to the analysis). These Guidelines do not discount the value or discourage the analysis of information derived from *post hoc* subgroup evaluations. However, investigators must clearly disclose the source of this information (e.g. how many subgroups were investigated versus the number of subgroups with significant results [see point iv) in box below]); and must convey due caution to readers in interpreting the results of *post hoc* subgroup analyses.

There is guidance available from the world of evidence-based medicine for those evaluating economic evaluations where subgroup analyses are reported. (Given that differences in clinical outcomes often drive subgroup analyses, this linkage of economic and clinical outcomes is both intuitive and appropriate.) Oxman and Guyatt (1992) have suggested a list of questions to be used to determine if supposed subgroup differences are, in fact, real differences (answering "yes" to most or all of these questions increases one's confidence in the subgroup analysis). A summary of these questions is outlined below, and may prove to be a useful guide in the context of pharmacoeconomic evaluations.

Methodologic Tips on Subgroup Analysis

Issues in Subgroup Analysis

1. Is the magnitude of the observed difference in subgroup responses clinically important (so that it would lead to different treatment recommendations for each subgroup?)
2. Was the observed difference in subgroup responses statistically significant?
3. Did a hypothesis that the subgroups ought to differ precede rather than follow the analysis?
4. Was this subgroup analysis one of a small number of subgroup analyses performed in the study?
5. Was the observed difference in subgroup responses suggested by comparisons within a study (as opposed to between two studies)?
6. Was the observed difference in subgroup responses consistent across two or more studies?
7. Is there indirect (e.g. biologic) evidence to support the hypothesized difference in subgroup responsiveness?

5.5 Treatment Comparator

Guideline 7. Treatment Comparator

The drug treatment should be compared with both existing practice and minimum practice. The relevant comparators may be other drugs, other medical care such as surgery, or even no treatment. Existing practice would either be the single most prevalent clinical practice (if there is one that is dominant), or it could be current practice weighted by market share. Minimum practice would normally be either the lowest cost comparator that is more effective than placebo, or the do-nothing alternative, as appropriate. In addition to these two formal comparators, all other reasonable alternative therapies should be at least discussed in the report.

There may be a variety of relevant comparators for a drug, and they may differ across the various subgroups of patients. As previously noted, relevant comparators may include other drugs, other medical care (e.g. surgery or watchful waiting), and no treatment. In theory, all other possible treatments for the same patients are relevant comparators. In practice, studies will have to identify one, or a small number, of primary relevant comparators.

The issue of relevant comparators is complicated because there are two possible questions. Is the new drug cost-effective relative to the existing drugs or treatments that it will in fact replace (local cost-effectiveness)? Or, is the new drug cost-effective relative to optimally cost-effective treatment (global cost-effectiveness)? For example, if current practice is itself unevaluated (which is often the case) and if current practice is in fact not cost-effective, the new drug can appear to be cost-effective (locally cost-effective) when in fact it is not (not globally cost-effective).

As an extreme numerical example to make the point, consider a new treatment program A which compared to existing treatment B costs $1,000 more and produces 1 more QALY. Such a program would have a (local) cost-effectiveness ratio of $1,000 per QALY gained, which is relatively cost-effective by any standards. Now suppose that program B had never itself been evaluated and, in fact, was quite cost-ineffective. Suppose that, compared to no treatment, program B costs $999,000 and gains 1 QALY. Then, compared to no treatment, program A would cost $1,000,000 and gain 2 QALY for a (global) cost-effectiveness ratio of $500,000 per QALY gained, which is relatively cost-ineffective by any standards. For a further discussion and more realistic example of this point see Drummond *et al.* (1993). In summary, to address both questions of local and global cost-effectiveness, the new drug should be ideally compared both to existing practice and to all relevant alternatives, including “do nothing” and the lowest cost therapy that is more effective than placebo.

Pharmacoeconomic researchers are encouraged to investigate both local and global cost-effectiveness of the new drug. At minimum, they are encouraged to compare to both existing practice and minimum practice in their studies. Existing practice would either be the single most prevalent clinical practice (if there is one that is most prevalent), or it could be current practice weighted by market share (e.g. a naturalistic control arm in a clinical trial that consists of normal practice and varies across physicians). Minimum practice would normally be either the lowest cost comparator (i.e. lowest treatment costs, as defined below) that is more effective than placebo, or the do-nothing alternative,

as appropriate.

In the ideal situation, one would compare the current most cost-effective option (as reflected [theoretically] in current practice guidelines or criteria for use) to the new agent. Practically, one often cannot identify such a comparator and, therefore, will use the agent with the lowest treatment costs (i.e. the sum of drug costs, administration costs, and the costs of treating any side effects) for a given course of therapy. This is more appropriate than using the drug with the lowest unit price as the comparator. However, even choosing the lowest cost comparator can be difficult. The selection of an appropriate comparator requires input from the decision-makers, as the choice of comparator relates to the question(s) the target audience wants answered. Thus, analysts are encouraged to obtain input from decision-makers as they develop their research protocols.

In selecting comparators and interpreting incremental comparisons it is crucial to understand the concept of **dominance** (Weinstein, 1990; Seigel, *et al.*, 1996; O'Brien, Heyland, *et al.*, 1997). The assessment of dominance is based on a comparison of the costs and effectiveness of each option. An option that has higher costs and lower effectiveness than another single option is said to be strongly dominated by that option. Weak dominance can occur when a new treatment has the same incremental cost but greater incremental effectiveness; or a lower incremental cost but the same incremental effectiveness. Finally, a more complicated type of dominance arises if an option is not dominated by any other single option, but rather is dominated by a weighted average of two other options. For example, option B might be dominated by a 50/50 weighted average of options A and C. This is known as extended dominance.

The importance of the concept of dominance in cost-effectiveness and cost-utility analysis is that all dominated options, strong weak and extended, are "inefficient". In the analysis, dominated options are all ruled out immediately. The non-dominated options form the **efficient frontier**, and the incremental cost-effectiveness or cost-utility ratios are formed along the efficient frontier. So, the practical implication is that a dominated option is never appropriate even as a comparator. The concepts of dominance are difficult to describe and understand in words, but easy to see graphically. For this reason, a graphical representation of cost-effectiveness and cost-utility results is strongly recommended within the final report of the evaluation (Seigel, *et al.*, 1996).

In addition to these two primary comparisons, there will often be a variety of other alternative therapies. All reasonable alternative therapies should be discussed in the study. The identification of appropriate comparators is likely to be an important point of discussion for any pre-submission review of the research protocol.

5.6 Perspective

Guideline 8. Perspective

All studies should report from a comprehensive societal perspective. That perspective should be transparently broken down into those of other relevant viewpoints, including that of the primary decision-maker. A financial impact analysis from the viewpoint of the primary decision-maker may also be undertaken, if requested, but technically is a budgeting exercise and does not constitute part of the economic evaluation.

Normally, the primary **analytic perspective** or viewpoint for pharmacoeconomic studies should be the comprehensive societal perspective; that is, all costs and benefits should be identified regardless of who incurs the costs or who receives the benefits. However, the comprehensive societal perspective should be transparently disaggregated into multiple viewpoints, including that of the primary decision-maker (the decision-maker, if any, to whom the study is primarily targeted). Relevant subsidiary viewpoints could include the health care system, major third party payers such as ministries of health, and the patient and family viewpoint. No matter what the viewpoint chosen, it should obviously be consistent on both sides of the cost-outcome ratio (i.e. in both the numerator and the denominator).

Generally, more narrow viewpoints such as that of a particular provider institution or provider group would be inappropriately narrow for a pharmacoeconomic study. A financial impact analysis, however, as opposed to an additional pharmacoeconomic viewpoint, should be conducted for major organizations that would be affected by the decision. This latter group certainly includes provincial drug plans, and possibly hospital formularies.

5.7 Analytic Horizon

Guideline 9. Analytic Horizon

Every effort should be made to extend the analytic horizon to capture all relevant outcomes. When modelled data are needed to meet this requirement, the structure and rationale of the model must be presented.

The analytic horizon for pharmacoeconomic studies should extend far enough into the future to capture the major clinical and economic outcomes related to the treatment(s) under study. It must be emphasized that the same time horizon must be applied to<u>both</u> costs and outcomes. In many cases, this would mean that the analysis must follow patients for the duration of their lifetime. Frequently, the appropriate analytic horizon will extend beyond the availability of primary data. In this case, the study will consist of primary data and modelled data. The assumptions of modelling should be explicit, well-justified, and thoroughly tested by sensitivity analysis. In many studies it may be useful to analyze the data using several analytic horizons: a short-term horizon that includes only primary data, and a long-term horizon that also incorporates modelled data.

5.8 Assumptions

Guideline 10. Assumptions

A comprehensive listing of assumptions and associated rationale must be contained within the explanation of the methodology for the analysis.

The issue of defining and identifying assumptions is mentioned frequently in the preceding text. However, it is worthwhile to re-emphasize the importance of communicating these assumptions as they are so vital to the investigator's evaluation. Derived from all of the choices made in designing the economic model (i.e. perspective, analytic technique, target population[s], important clinical assumptions, expert opinions, observed data, etc.), there must be a clear and concise delineation of all of the assumptions which form the base case of the analysis.

A comprehensive listing of assumptions and their associated rationale should be contained within the explanation of the methodology for the analysis. The magnitude and direction (i.e. conservative versus over-estimate) of the bias in the assumption should be indicated, and published evidence supporting the assumption should be provided wherever possible. These assumptions then go on to form part of the basis for the evaluation of uncertainty (i.e. the sensitivity analysis; see Section 8.2) once the data are collected and analyzed.

6.0 Measuring and Reporting Outcomes

6.1 Efficacy Versus Effectiveness

Guideline 11. Efficacy Versus Effectiveness

Ideally, pharmacoeconomic studies should report on drug effectiveness rather than efficacy. Because effectiveness data are generally not available, appropriate modelling techniques based on sound pharmacoepidemiology (e.g. using epidemiologic studies to estimate patient compliance with therapy in the real world) are permissible. All assumptions used in such extrapolation techniques must be stated explicitly and thoroughly tested with sensitivity analysis.

Finding the Data

Efficacy refers to the performance of a drug under highly controlled circumstances -- that is, administered according to a strict written protocol by highly motivated, research-oriented clinicians to consenting, compliant patients who are a carefully selected subgroup of patients meeting restrictive inclusion and exclusion criteria. Effectiveness, on the other hand, refers to the performance of a drug in the real world with a wide variety of providers administering the drug as they see fit to a broad heterogeneous group of patients who are less well-informed, less compliant, and liable to be influenced by a variety of concomitant diseases and/or medications not investigated in the original efficacy trials (Tugwell, *et al.*, 1985; Bailey, 1994; Bloomfield Rubins, 1994; Davis, 1994).

Pharmacoeconomic studies should use effectiveness data as their source of clinical evidence regarding the impact of an intervention. Unfortunately, the only data available prior to the launch of a new product are Phase III efficacy data. Thus, prelaunch pharmacoeconomic studies must extrapolate from trial efficacy to utilization effectiveness using modelling techniques (Johannesson, *et al.*, 1997). Buxton, *et al.* (1997) provide a useful review of the role of modelling in economic evaluations [see annotated bibliography at end of this section]. The assumptions used in this extrapolation (for example, patient compliance rates) must be explicit and must be tested thoroughly with sensitivity analysis. After launch, it is possible to incorporate the above limitations to mount effectiveness trials to gather primary data on the cost-effectiveness of the drug. Such studies are strongly encouraged.

The process of obtaining efficacy or effectiveness data can present its challenges. Some would argue that prospective data reflecting the "real-life" experience of drug use (i.e. effectiveness) in a large number of patients are most desirable, while those derived prospectively from several large randomized controlled trials (RCTs; i.e. efficacy) would be next best. Retrospective data from either effectiveness or efficacy data sources represent viable but not ideal alternative information sources. In practical terms, the preferred source of data is dependent on the complexity of the question being investigated. Analysts must think carefully about the economic question at hand and the most appropriate sources of data for that question. No matter what the origin, analysts must make the presentation of the data transparent and explain the rationale for the source of data used in the study.

The Use of Meta-analysis

It is often the case that studies reporting efficacy or effectiveness data are either insufficient or are conflicting, yet there is still the need for information to support valid retrospective model development. **Meta-analysis** a process of combining study results in such a way as to be able to draw conclusions about therapeutic effectiveness (L'Abbe,*et al.*, 1987). As such, it is a tool for increasing the precision of estimated differences between the proposed drug and appropriate comparators which can then be used in a pharmacoeconomic model. It can also highlight advantages and disadvantages of the proposed drug and its comparators which are too small to be detected accurately in individual trials.

If the investigator determines that a meta analytic approach is required, the method used must be rigorous, should be thoroughly rationalized in comparison to the use of other data sources, and should be submitted as part of the economic evaluation. An outline for systematic assessment based on the approach by the Cochrane Collaboration is presented in the*Methodological Tips* box below. L'Abbé, *et al.* (1987) present a related checklist for establishing the quality of published meta analyses. Readers are referred to several sources of information regarding classical as well as Bayesian approaches to meta-analysis: Hunter and Schmidt, 1990; Rosenthal, 1991; Velanovich, 1991; Carlin, 1992; Eddy,*et al.*, 1992; US General Accounting Office, 1992; Fleiss, 1993; Gibaldi, 1993; Oxman, 1993; Cooper and Hedges, 1994; Sorofman and Milavetz, 1994; Petitti, 1994; Speigelhalter, *et al.* 1994; Chalmers and Altman, 1995; McIntosh, 1996; Tweedie,*et al.*, 1996; Oxman, 1996; Su and Po, 1996; Cook,*et al.*, 1997.

While there are currently no standardized formats, the reporting of meta analyses is an area which is evolving. Similar to the efforts towards a standardized reporting structure for RCTs proposed in the CONSORT (Consolidated Standards of Reporting Trials) statement (Begg,*et al.*, 1996; Huston and Hoey, 1996), efforts are currently underway for a similar approach to reporting meta analyses. The goal of the QUORUM working group on systematic reviews is to create a reporting format which avoids systematic bias in treatment estimates (personal communication, D. Moher, February 26, 1997).

Methodologic Tips on Meta-analysis

An Outline for the Meta-analytic Approach (adapted from Cochrane Collaboration)

a) Use a reproducible, exhaustive and well-described search strategy (computerized, hand and citation searches) to identify trials. List inclusion and exclusion criteria, as well as a list of excluded studies and reasons for their exclusion (to avoid selection bias).

b) Summarize the general characteristics of each trial included in the review (i.e. gender, age range, severity of disease, dose range of intervention, year and language of publication, methodologic quality [with evidence it was reproducibly applied], sources of funding). The comparative trial information should be provided in a tabular form so that readers can make an informed decision about the distribution pattern of characteristics in the trials and the appropriateness of the synthesis.

c) Tabulate the results (point estimates and 95% confidence intervals) of the individual trials. Economic evaluation often requires the construction of a decision tree that requires estimates of the number of patients who demonstrated a response ("responders") rather than mean improvement. If so, the responder analysis should be the source of the primary estimates listed in the tree. However, for the purposes of the final economic assessment, data from all patients (i.e. intent to treat analysis) should be used.

d) Plot the results of the individual trials (point estimates and 95% CIs) and examine for heterogeneity.

e) The methodology for primary and subgroup statistical analysis must be specified prior to evaluation. Once a statistical assessment of heterogeneity is completed, if the visual presentation and/or statistical test indicate that the results are heterogeneous, check out clinical and methodologic reasons that explain this heterogeneity.

f) Statistically combine the results for the absolute and relative risk reductions using both the fixed effects and random effects models (Armitage and Berry, 1994) - giving 4 combinations in all. The approach used in the statistical combination of the results (e.g. pooled hazard ratios) should be justified and explained in a short technical document or attachment to the submission.

g) Show sensitivity analyses to test the robustness of results based on features of the primary studies (e.g. scientific merit) and key assumptions and decisions made in the selection, analysis and presentation of studies and their findings, as selected on an a priori basis.

h) Select one estimate from the 4 options in (f) above for use in the economic evaluation. Justify the selection.

Annotated Bibliography:

Buxton MJ, Drummond MF, Van Hout BA, Prince RL, Sheldon TA, Szucs T, et al. Modelling in economic evaluation: an unavoidable fact of life. Health Economics 1997;6(3):217-227.

Buxton, et al. review the role of decision analytic and statistical models in health economics as a means of communicating the complexity of interventions, as well as simulating these interventions or evaluating alternative scenarios. There is also a thorough presentation of: concerns regarding the use of models in economic evaluations; the use of "pragmatic" trials as an alternative to modelling in evaluating the cost-effectiveness of interventions in "real world" conditions; and suggestions for creating a "best practices" approach to modelling when no other options are available.

Johannesson M, Jonsson B, Karlsson G. Outcome measurement in economic evaluation. Health Economics 1996;5(4):279-296.

The authors provide an excellent review of the different types of outcomes that can be measured for not only cost-effectiveness studies, but also cost-utility and cost-benefit analyses. The pros and cons of each category of outcome within each analytic method (e.g. surrogate endpoints versus endpoints versus survival versus health status/quality of life for CEAs) are presented. They make recommendations for the outcomes which should be measured in each instance, recommendations which are being discussed by the "Harmonization by Consensus of the Methodology for Economic Evaluation of Health Technologies in the European Union" project members.

6.2 Health-Related Quality of Life as an Outcome

Guideline 12. Health-Related Quality of Life (HRQOL)

If HRQOL is being included in a prospective study as an outcome, it is normally advisable to include, where possible and feasible, one instrument from each of the following three types: specific measures, generic profiles, and preference-based measures. Any drug product that demonstrates improved effectiveness over its comparator(s) and impacts on a patient's HRQOL should probably be evaluated for this outcome using these tools.

General Concepts

Quality of life is a broad concept that includes many aspects of living in addition to health, for example; wealth, freedom, political system, and cleanliness of the environment all contribute to the overall QOL (Patrick and Erickson, 1993). Health-related quality of life refers to those aspects of QOL that are related to health. The overall goal of the health system is to improve both survival (life expectancy) and HRQOL. Accordingly, many tools have been developed to measure HRQOL. Any drug product that demonstrates improved effectiveness over its comparator(s) and impacts on a patient's HRQOL should probably be evaluated for this outcome using these tools. A useful taxonomy for these instruments is that developed by Guyatt and colleagues in which the methods are partitioned into three major sets: specific instruments, generic profiles, and **utility** (preference-based) measures (Guyatt, *et al.*, 1989; Guyatt, *et al.*, 1993; Feeny, Torrance and Labelle, 1996).

No single measure of HRQOL has yet been accepted as the gold standard. In fact, given the multi-dimensional nature of HRQOL, it is unlikely that a single measure will ever become the gold standard. Therefore, if HRQOL is being measured in a prospective study, it is normally advisable to include one reasonably precise, reliable and valid scale from each of the following three types: generic, specific, and a preference measure. Other scales being developed can, of course, be included. The choice of instrument(s) is based on many factors, including: the content of the tool(s) being considered, the basis of scoring, and the question and/or disease state being investigated. The investigator should also consider the burden on the patient (or proxy) of completing multiple instruments and surveys.

Specific Measures

Specific instruments include those that are targeted at specific diseases, such as the Functional Living Index - Cancer (Clinch, 1996) or the Western Ontario-McMaster Osteoarthritis Index (Bellamy, *et al.*, 1988); specific populations, such as the Care and Resource Evaluation Tool for the Elderly (Fretwell, 1996); and specific functions, such as visual function measured by the Activities of Daily Vision Scale (Mangione, *et al.*, 1992; Mangione, *et al.*, 1996). Specific instruments are designed for a particular type of application and are generally not relevant for other purposes. For example, a cancer index may not detect a change brought about by an arthritis intervention. The advantage of specific instruments is that they would be expected to have higher responsiveness to change (Guyatt, *et al.*, 1996).

Generic Measures

Generic health profiles are applicable to a wide range of patients and diseases and, thus, are more generalizable but probably less responsive than specific instruments (Guyatt, *et al.*, 1996). They provide scores on a number of dimensions (a profile of scores) and typically are not aggregated into an overall summary score. Three well known instruments in this category are the Short Form 36 (SF-36; Ware, 1996), the Sickness Impact Profile (SIP; Damiano, 1996), and the Nottingham Health Profile (NHP; McEwen and McKenna, 1996).

Preference-based Measures

Preference-based measures provide a single summary score which reflects numerically the HRQOL. Like generic profiles, these approaches are applicable to a wide variety of diseases and individuals and, thus, are highly generalizable. However, they are likely to be the least responsive of the three approaches (Guyatt, *et al.*, 1996). The advantage of preference-based measures is that they are the only approach that provides a score reflective of HRQOL that is suitable for use in CEA and CUA. (Issues surrounding the use of preference-based measures are discussed further in Sections 6.3 and 6.5.)

Modelling HRQOL

To this point, the discussion has focused on HRQOL measurement in prospective evaluations. In retrospective modelling studies, the analyst does not have the luxury of specifying the HRQOL instruments that will be used to gather the data. Typically, the analyst must work with results from clinical trials that did not incorporate such instruments. In this case, the analyst can undertake a CEA using the primary clinical effectiveness measure from the trials. If the analyst wishes also to undertake a CUA, the effectiveness outcomes from the trial must be somehow mapped onto utility scores.

There are fundamentally three means by which effectiveness can be mapped onto utility scores. One method is to develop written scenarios that describe the relevant health states from the trial, and to measure the utility of these states on a sample of the general public using an technique like the **standard gamble (SG)**. An alternative method is to map the health states from the trial onto a multi-attribute system like the health utilities index (HUI). A third possibility would be to find patients

currently in the health states relevant to the trial and to measure their utility for these states. The analyst must support the method used, describe it fully, and subject it to rigorous testing for robustness (i.e. using sensitivity analysis or statistical methods).

Annotated Bibliography:

Spilker B, editor. Quality of life and pharmacoeconomics in clinical trials, 2nd ed. Philadelphia, PA: Lippincott-Raven; 1996.

Spilker has put together what is likely the most comprehensive text on quality of life measurement in the context of clinical trials. Starting with a basic introduction, experts in the field go on to discuss fully 127 different topics related to: specific tools; the process of choosing and administering QOL instruments; analysis, interpretation and presentation of results; the perspectives through which various stakeholders interpret QOL issues; cross-cultural and cross-national issues; the interaction between QOL and health policy; special subpopulations and disease states for QOL assessment; and the relationship between QOL studies and pharmacoeconomics.

McDowell I, Newell C. Measuring health: A guide to rating scales and questionnaires, 2nd ed. New York: Oxford University Press; 1996.

McDowell and Newell provide a comprehensive review of the theoretical underpinnings of health measurement tools, emphasizing the means by which the theory affects their technical application. The majority of this textbook is devoted to presenting and critically evaluating the validity and reliability of a wide variety of general and disease-specific health status and QOL measurement tools.

6.3 Outcomes for Cost-Utility Analysis

Guideline 13. Outcomes for Cost-Utility Analysis

Both quantity of life (survival) and HRQOLresults should be reported separately, and the method of combining the two described in a transparent manner. The current recommended method for the primary analysis is to combine quantity and quality of life using quality-adjusted life years (QALYs). Alternatives, such as disability-adjusted life years (DALYs), healthy years equivalents (HYEs) or saved young life equivalents (SAVEs) may be useful as secondary analyses in some studies. To be suitable for use as quality weights for calculating quality-adjusted life years, scores must be based on preferences and measured on an interval scale where dead has a score of 0 and healthy has a score of 1.

Direct preference measurements can be undertaken with various instruments. Analysts should select an instrument that suits the problem, and should justify their selection. Alternatively, preferences can be determined indirectly using one of the major systems available. Here again, analysts should select a system that suits the problem, and should justify their selection.

In CUA the quantity of life improvement (survival) and the HRQOL improvement (morbidity) are combined into a single metric (e.g. quality-adjusted life years gained). This combination is essential in order to achieve the broad comparability across programs which is the hallmark of the CUA. It is also important, however, to report the quantity and quality results separately and to be entirely transparent in how the two are combined.

QALYs

The current generally accepted method of combining quantity and quality is through the use of QALYs (Torrance and Feeny, 1989; Weinstein, 1990; Gold, Patrick, *et al.*, 1996; Torrance, 1996; Drummond, *et al.*, 1997). A QALY is calculated by multiplying the number of life years added via a program by a standardized weight (between 0.0 and 1.0) that reflects the health-related quality of life during that time (where 0.0 is the weight given to immediate death and 1.0 is the weight given to perfect health for a defined period of time). In a slight variation on this theme, some approaches to QALY weights provide for the possibility of negative weights for states considered worse than death (Torrance, 1984; Patrick, *et al.*, 1994; Ditto, *et al.*, 1996).

QALY weights are measured by asking relevant individuals which consequences they prefer and by how much, thereby reflecting the value people place on different health outcomes. The weights must be preference-based, cardinal, and measured on the 0 to 1, dead to healthy scale. Preference can be measured either directly or indirectly (see below for further discussion). It should be noted that weights cannot be obtained from the other two types of health-related quality of life instruments (i.e. specific instruments and generic profiles). Their scores do not have the necessary measurement properties for QALY weights identified above.

The QALY approach is useful in policy analysis and program decision-making, in part, because it is completely general. It can be applied to any population, any disease, any intervention, and can be used to compare across quite diverse programs. However, for the comparisons to be valid, the QALY studies must use the same methodology; for example, the same QALY weights, the same perspective, the same **discount rate**, etc. Unfortunately, because of the lack of standardization in the field to date, such comparisons must be undertaken with extreme caution (Drummond, *et al.*, 1993; Mason, *et al.*, 1993; Drummond, *et al.*, 1995).

The QALY approach contains a number of assumptions and limitations:

1) It assumes that all QALYs are equal. For example, it assumes that it is equally desirable to provide a one QALY gain to a teenager or to a senior citizen, to a woman or a man, to a labourer or to a professional, and so on.
2) It also assumes that it is equally desirable to provide a small gain to many people or a large gain to a few, as long as the QALY totals are the same. For example, a gain of 0.1 QALY to each of 1000 people would be considered equal to a gain of 25 QALYs each to four individuals.
3) As usually practiced, the QALY approach assumes that the relative weights for health states are independent of the duration of the health states. However, it is possible to circumvent this assumption by measuring the weights specifically for the durations that are relevant (Drummond, *et al.*, 1997).
4) The QALY approach also assumes that the preferences that individuals have for paths of changing health states can be reasonably estimated by adding up the time-weighted preferences that the individual has for the components of that path (Kuppermann, *et al.*, 1997).

Despite these assumptions and limitations, the QALY approach remains the most common approach for combining quantity and quality of life, and using cost-utility analysis. The frequency of use of this approach is probably the result of its clarity, simplicity, ease of application, face validity, and, when the weights are based on von Neumann-Morgenstern utilities, its theoretical foundation.

Nevertheless, a number of alternatives to QALYs have been suggested, as described below. The alternatives are based on relaxing or modifying specific assumptions of those listed above. For example, DALYs change assumption 1) by weighting on the basis of age; **healthy years equivalent** (HYEs) change assumptions 3) and 4) by directly measuring preferences over paths; and SAVEs change assumption 2) by directly asking person trade-off (PTO) questions.

Other Approaches

Disability-adjusted life years apply unequal weights to the QALYs that are gained based on the age of the recipient (Murray, 1994; Murray and Lopez, 1996). For example, a life year gained when a person is 25 years of age is assigned a higher value than when the person is younger or older. These weights were selected to capture social roles at different ages. However, both the age weights and the weights for the health states are not empirically founded, but are simply based on expert opinion.

The Healthy Years Equivalent method has been suggested as an alternative to the QALY model that does not require assumptions 3) and 4) above (Mehrez and Gafni, 1989; Mehrez and Gafni, 1991; Mehrez and Gafni, 1992). The basic approach of the HYE method is to measure preferences over a complete path of health states, rather than each state singly; and to use a two-stage standard gamble technique to perform preference measurements, rather than a single standard gamble.

There has been extensive discussion and debate on all aspects of HYEs. The following references are examples of the most recent contributions to, or summaries of, the issues: Buckingham, 1993; Culyer and Wagstaff, 1993; Fryback, 1993; Gafni, *et al.*, 1993; Johannesson, *et al.*, 1993; Mehrez and Gafni, 1993; Loomes, 1995; Williams, 1995; Gafni, 1996; Gold, Patrick, *et al.*, 1996; Wakker, 1996; Weinstein and Pliskin, 1996; Drummond, *et al.*, 1997. While the debate regarding HYEs is complex, it can be summarized by stating that one aspect of the approach has merit while another aspect does not. Measuring preferences over a path of health states has merit, and indeed is theoretically superior to the traditional QALY approach. However, it is difficult to implement in practice, and likely to be impractical in many projects. The two-stage standard gamble technique of HYE does not have merit. The two-stage procedure is, in fact, theoretically equivalent to a much simpler technique, a one-stage **time trade-off (TTO)** procedure (Rittenhouse, 1997).

An approach using person trade-offs to determine saved young life equivalents (SAVEs) has been suggested as an alternative to the QALY (Nord, *et al.*, 1993; Nord, 1995; Nord, 1996). Proponents of this method suggest that the weights for QALYs reflect an individualistic perspective not a societal perspective and, thus, the conventional QALY does not measure social value. Research shows that when members of the general public are asked specific PTO questions (like how many patients of type A should be cured to be equivalent in social value to curing 10 patients of type B), the results do not match conventional QALYs. With this approach, the SAVE is the common metric which allows comparisons between disparate programs. All programs are converted through PTO measurements to their SAVE value, and programs are compared on the basis of costs and SAVEs. It is argued that

this method is more fair from a societal point of view.

The approach is new and the measurement techniques are still under development (Ubel, Loewenstein, *et al.*, 1996). In some studies, for some decision-makers, analysts may wish to explore the impact of using the SAVE methodology. If the approach gives dramatically different answers to the resource allocation question versus QALY methods, then a discussion of the reasons could be quite enlightening for the decision-makers.

Generic/Specific HRQOL Measures versus Preferences

In using the QALY model, the question arises as to the best source of quality-adjustment weights (rating scale, standard gamble, time trade-off, Health Utilities Index, Quality of Well Being, EQ-5D), and whether they can be approximated by experts. Some studies use expert judgement with extensive sensitivity analysis as the source of quality-adjustment weights (see Section 6.5 for further discussion on source of preferences). This approach is not favoured, although clearly there will be some studies where the results can be shown to be quite insensitive to the quality weights, and approximate estimates will suffice.

Occasionally, studies use the global summary score of a generic HRQOL profile (for example, Sickness Impact Profile) as the quality weight for QALYs. This is completely inappropriate (Gold, Patrick, *et al.*, 1996). First, such scores for some generic HRQOL measures are not based on preferences and, thus, score changes do not reflect gradations in the preferability of outcomes. Second, the scores are not measured relative to death so they cannot be used to combine morbidity and mortality; that is, the health state for mortality (dead) is not on the scale (see discussion regarding preference measurement below).

Preferences

Typically, to be suitable for calculating QALYs, preference-based scores are measured on a cardinal (i.e. interval) scale on which death is 0 and perfect health is 1. Note that states worse than death are allowed on this scale, and would take on values less than 0. The preference score (i.e. the weight for a CUA) can be measured directly or indirectly, and is the best source of quality-adjustment weights.

Direct Measurement

Direct measurement requires the analyst to conduct complex and costly measurement tasks using one of the three more widely used instruments: the standard gamble or time trade-off for **revealed preferences**; or the visual analog scale for stated preferences (Torrance, 1986; Furlong, *et al.*, 1990; Bennett and Torrance, 1996; Feeny, Torrance and Furlong, 1996). (Ratio scaling and paired comparisons methods have been used occasionally). Among the direct measurement tools, the standard gamble is favoured by many economists and decision scientists because of its strong normative foundation in von Neumann-Morgenstern utility theory (von Neumann and Morgenstern, 1944; Bennett and Torrance, 1996). On the other hand, there are those who argue against the superiority of the standard gamble (Richardson, 1994).

There have been many head-to-head studies comparing the three main instruments, and the general

finding is that standard gamble scores exceed time trade-off scores, which in turn exceed visual analog scores; and that the relationships are curvilinear (Drummond *et al.*, 1997). Thus, it is possible to take the measurements using one instrument, for example the visual analog scale, and convert to what would have been obtained with another instrument, for example the standard gamble (Torrance, Feeny, *et al.*, 1996).

In addition to these empirical differences, the various approaches have differing theoretical foundations and characteristics (Bennett and Torrance, 1996; Gold, Patrick *et al.*, 1996). Analysts who wish to undertake direct preference measurements should select a measurement technique that has theoretical and empirical properties that suit the problem being addressed, and should clearly explain their selection.

Indirect Measurement

Alternatively, the weights can be obtained indirectly through the use of "off the shelf" preference-weighted health status systems. Three well-known instruments in this category are the Quality of Well Being (QWB; Kaplan and Anderson, 1988; Kaplan and Anderson, 1996), the Health Utilities Index (HUI; Feeny, *et al.*, 1993; Furlong, *et al.*, 1994; Feeny, *et al.*, 1995; Torrance, *et al.*, 1995; Feeny, Torrance and Furlong, 1996), and the EuroQol, now renamed the EQ-5D (EuroQol Group, 1990; Essink-Bot, *et al.*, 1993; Dolan, *et al.*, 1995; Brooks, 1996; Kind, 1996). All of these systems have the same overall structure, whereby health status is described by multiple attributes and levels of function within each attribute.[1] To use these instruments the analyst only has to classify the health status of the patient into the classification system provided and compute the preference score from the formula. The score represents an estimate of the mean preference score that would be given to that health state by a random sample of the general public.

In each indirect preference measurement system, every unique combination of levels across attributes defines a unique health state. All of the systems have a scoring algorithm that provides a single summary score for each individual health state. The systems, however, differ in terms of their construct and coverage of health status (attributes and levels), the number of levels available in each attribute, the number of unique health states available in the system, the preference measurement technique that underlies the scoring algorithm (category scaling, time tradeoff, standard gamble), and the structure of the scoring algorithm (additive, additive with adjustment terms, and multiplicative).

For comparability across studies, it would be desirable for all studies to use the same system. No head-to-head comparisons of these systems are, however, available to guide users. At the moment, the best advice is for users to study the alternative systems, to select in advance the one that best suits the study objectives, to justify the selection in the study protocol, and to stick with it. It is not

[1] For example, the attributes for the QWB are mobility, physical activity, social activity and symptom-problem complex; attributes for the HUI2 (i.e. version 2 of the HUI) are sensation, mobility, self-care, emotion, cognition, and pain/discomfort; and the attributes for the HUI3 (i.e. version 3 of the HUI) are seeing, hearing, speaking, ambulation, dexterity, emotion, cognition, and pain/discomfort.

appropriate to try a variety of approaches and simply pick the one that puts the product in the best light.

Summary

Cost-utility analyses are being performed and published on an increasingly frequent basis. It is important that this methodology be considered as one of the means by which an economic evaluation will be carried out, when and where it is applicable.

Annotated Bibliography:

Drummond MF, O'Brien BJ, Stoddart GL, Torrance GW. Methods for the economic evaluation of health care programmes, 2nd ed. Oxford: Oxford University Press; 1997.

Chapter 6 of this textbook (Cost-Utility Analysis) provides a comprehensive and up-to-date treatment of all of the material mentioned in this section of the Guidelines, including utilities, multi-attribute utilities, measuring preferences, preference-based health status systems, quality-adjusted life years, and alternatives to QALYs.

Gold MR, Patrick DL, Torrance GW, Fryback DG, Hadorn DC, Kamlet MS, Daniels N, Weinstein MC. Identifying and valuing outcomes. In Gold MR, Russell LB, Seigel JE, Weinstein MC, editors. Cost-effectiveness in health and medicine. New York, NY: Oxford University Press; 1996. p. 82-134.

This chapter of the Washington Panel text provides a thorough discussion of the concept of HRQOL and the various means by which it is combined into a summary measure with quantitative measures of the duration of life. Discussion also ensues regarding health status measures, the arguments regarding sources of preference measurement, and the techniques for taking such measurements.

Nord E. The person-trade-off approach to valuing health care programs. Medical Decision Making 1995;15(3):201-208.

Nord reviews the rationale underlying the person-trade-off method for estimating the value of health care interventions via CUA. Because of its grounding in the societal perspective, Nord suggests that this method is preferred for resource allocation decision-making. At the same time, the disadvantages of the method are presented (e.g. require large groups to minimize random error, process of scenario presentation is complex in order to avoid introduction of bias). Nord discusses the evidence for the reliability and validity of the method by reviewing previous studies using the PTO technique.

Wakker P. A criticism of healthy-years equivalents. Medical Decision Making 1996;16(3):207-214.

Wakker provides a comprehensive outline of the debate regarding outcomes evaluations via decision analysis using QALYs versus HYEs. Background is provided by way of a review of expected utility theory; and the means by which QALYs and HYEs address each of the assumptions of the theory are compared and contrasted. Readers should review the papers which accompany this article (in the same issue) to put the debate into a more complete perspective: a preamble by Weinstein and Pliskin (pp. 205-206), further input from Gafni (pp. 216-217), and a final rebuttal by Wakker (p. 217).

6.4 Outcomes for Cost-Benefit Analysis

Guideline 14. Outcomes for Cost-Benefit Analysis

The human capital approach (HCA) to assigning values to outcomes in CBA is incomplete because it focuses primarily on lost work time. If this approach is used, measures taken to overcome these shortcomings should be clearly described.

While realizing the incomplete and experimental nature of this method, theoretically, the preferred approach to the assignment of values to outcomes in CBA is contingent valuation as a means of eliciting an assessment of WTP. If this approach is used, the investigator(s) should be explicit with regards to the assumptions and methods utilized. In addition, the measures taken to reduce bias and an outline of the scope tests carried out to determine validity should be clearly described.

This section attempts to review the significant research activity which has taken place in this area, as well as to delineate the issues of controversy which have arisen therein. The intent of the emphasis placed on CBA within these Guidelines is not to promote CBA as the ideal methodology for health related economic evaluations, but rather to inform "doers" regarding new research directions which may affect the discipline in the future.

In broad terms, when a CBA is carried out, the value (i.e. utility) of a particular program is measured by what the consumer is willing to pay for it. A large part of that "value" may be reflected in the health improvement achieved by a given program / intervention, which is then translated into an equivalent dollar value for comparison with the incremental costs. There have been two predominant approaches to converting these outcomes into dollar values: the **human capital approach (HCA)** and contingent valuation methods to elicit an expressed preference through assessment of willingness to pay (WTP). (Readers should also see Section 7.1 for a brief discussion of the **friction cost method** as an alternative to the HCA.)

In early studies, outcomes were valued using the human capital approach, which, in addition to measuring change in **direct medical and direct non-medical costs**, also assesses the dollar value of lost productivity and the value of missed formal (paid) work in the absence of the program/intervention (or in the reverse, the increased productivity as measured by increased earnings due to the presence of the program/intervention). Exclusive reliance on the traditional HCA to CBA is falling out of favour, in part because it is incomplete and because it is not consistent with the welfare economic theory that underlies CBA (Mishan, 1971; O'Brien and Gafni, 1996).

The HCA focuses primarily on lost work time and, as such, restricts its scope to persons in the labour force. It does not deal well with other kinds of time losses, for example, time lost by those not in the employed work force (e.g. homemakers, unemployed, retired, children), or patients with certain disease states (e.g. Alzheimer's disease). Although, adjustments can be made to attribute wages for some of these groups, the adjustments do not cover everyone. In addition, the approach cannot attach a value to health improvements that have no impact on working time, even though these improvements may be highly valued by the recipients.

There are, however, particular study perspectives which might be more amenable to the human capital methodology, and thus the technique should not be abandoned entirely. For instance, studies from the employer's perspective might focus on decreased illness-related absenteeism (and, presumably, increased productivity), an important outcome which could be accurately measured via the HCA. (In fact, it could be argued that employers/third party insurers should consider the overall health and HRQOL of employees/beneficiaries, as these also affect productivity.) In any event, if the HCA is used, the measures taken to overcome the aforementioned shortcomings should be clearly described.

Theoretically, the preferred approach to assessing benefits in CBA is contingent valuation. By definition, contingent valuation is a survey methodology which uses hypothetical "questions to elicit people's preferences for public goods by finding out what they would be willing to pay for specified improvements in them" (Mitchell and Carson, 1989; Portney, 1994). Thus, a specific monetary term is elicited to describe their preferences as reflected by their willingness to pay. There is significant debate regarding the validity and, therefore, use of contingent valuation/WTP as a means of placing value on a given outcome (Diamond and Hausman, 1994; Hanemann, 1994; Johannesson, 1996). This debate can be framed in terms of the relative advantages and disadvantages of the HCA versus WTP approaches.

Compared to the HCA, the WTP approach has two major advantages. First, it measures the outcome of primary interest; that is, the value of improved health per se, not just the impact on productivity. (Here "value" is defined as the most money the respondent would sacrifice to have the program/intervention in place.) Thus, the value of improved health can be determined even if the individual does not re-enter the labour force. Second, the question can be asked such that it captures externalities, spillover benefits and harms to other individuals (Gafni, 1991; Labelle and Hurley, 1992).

There are also difficulties with the WTP approach. The diverse nature of the possible outcomes in health care often make contingent valuation approaches difficult to implement (O'Brien and Gafni, 1996). It is more often the case that an intervention impacts on a subjective endpoint involving HRQOL, as opposed to a potentially more quantifiable endpoint such as avoided mortality. In addition, particular attention needs to be focused on creating believable payment scenarios in WTP questions, using realistic payment vehicles which make sense in the context of how the health care system is funded (for example, "willingness to be taxed" for the availability of an intervention, in the context of the Canadian public system).

Despite these methodologic quandaries, WTP studies have increasingly appeared in both the pharmaceutical and general health technology assessment literature (Berwick and Weinstein, 1985; Thompson, 1986; Appel, *et al.*, 1990; Donaldson, 1990; Estaugh, 1991; Johannesson and Jonsson, 1991; Johannesson, Jonsson and Borgquist, 1991; Johannesson, 1992; Johannesson, 1992; Neumann and Johannesson, 1994). None of these papers represents the ideal, and significant conceptual and practical issues have yet to be resolved before these methods can be used routinely.

Early contingent valuation research involved asking the respondent to provide a direct dollar amount as to her/his WTP using open-ended questions. These methods were too difficult for respondents. Other early research used an open-ended "bidding game techniques", which suffered from contextual

biases, particularly starting point bias. A dichotomous question (yes/no) for a single price has been advocated to eliminate these biases, but at the expense of requiring very large sample sizes (Johannesson, l992b). In terms of the context of the question, the breadth and benefits of the program must be clearly defined and presented to the respondents (for instance, do benefits include only personal health benefits, or do they also include the health effects on productivity, income or public expenditure consequences?).

Who should be asked WTP questions? Some analysts suggest that patients in need of a specific treatment should be asked their WTP for the actual health improvement itself (i.e. the use-based approach) (Thompson, 1986; Johannesson 1992; Johannesson, 1992; Johannesson, 1996). Others argue that the theoretically correct question is the WTP for an increased insurance premium that would make the treatment available free of charge if, and when, needed (i.e. the insurance-based approach) (Gafni, 1991). These are fundamentally different approaches and produce different dollar valuations (Neumann and Johannesson, 1994). Johannesson (1996) has produced a theoretical framework for connecting the use-based and insurance-based approaches, based on the degree of risk aversion.

From a methodologic perspective, a stumbling block to the regular use of WTP is its reliability and validity (Johannesson, Johansson, and Jonsson, 1992). For instance, O'Brien and Viramontes (1993) were able to demonstrate the reliability of WTP responses; but the large variation in the replies compromised the discriminant validity of this tool. As a means of assessing validity, it has been suggested that WTP studies incorporate "scope tests", wherein the congruence of the direction of the WTP response with the direction of benefit is assessed (Neumann and Johannesson, 1994; O'Brien, Goeree, *et al.*, 1997). The ideal gold standard for validating WTP survey data would, however, involve actual observed WTP.

In summary, CBA techniques and, specifically, WTP methodologies are very much at an experimental and incomplete stage of development in terms of their use as tools in the economic evaluation of health related interventions. More research on the feasibility, reliability, validity and responsiveness of WTP is required. Analysts are encouraged, however, to incorporate and experiment with contingent valuation approaches in studies, both for the potential usefulness of the information and for the methodological benefits of gaining more experience.

Annotated Bibliography:

Johannesson M. The contingent valuation controversy in environmental economics and its relevance to health services research. Journal of Health Services Research and Policy 1996;1(2):116-117.

Johannesson provides a useful overview of the controversies regarding the use of contingent evaluation as a means of deriving expressed preferences for alternative programs. While the sources cited refer to economic evaluations in the environmental field, Johannesson points out that these same issues are applicable to the assessment of willingness to pay in health care. The author also proposes means by which health-oriented studies can be designed to overcome some of the criticisms aimed at the contingent valuation method.

O'Brien B, Gafni A. When do the "dollars" make sense? Toward a conceptual framework for contingent valuation studies in health care. Medical Decision Making 1996;16(3):288-299.

O'Brien and Gafni provide an overview of the general approaches to valuation of outcomes for CBAs. There is discussion of the underlying theories and relative merits of the HCA versus WTP approaches. The focus of the article is on developing a conceptual framework for designing and interpreting such studies for health care program evaluation.

6.5 Source of Preferences

Guideline 15. Source of Preferences

The appropriate source of preferences depends on the use of the analysis and the viewpoint. For provincial drug plans, which are tax supported, the appropriate viewpoint is societal and the appropriate source of preferences for outcomes is the informed general public. The three major systems for the indirect determination of preferences (QWB, HUI, EQ-5D) are all scored based on preferences from the general public and, so, are suitable.

Analysts who wish to measure preferences directly should ideally do so on the general public, suitably informed. Patients in a study may, however, be a reasonable proxy for the informed general public, especially when they are providing preferences for hypothetical states. Preferences should be based on scientifically sound measurements. Investigators undertaking direct measurements of preferences must justify their source of subjects, and describe the exact population from which the preferences were derived and the precise methods of measurement.

In both CUA and CBA, preferences for health outcomes are required and are expressed as utilities or through WTP questions. Whose preferences are the right preferences? The two main candidate groups are the general public and selected patients. The latter group can be subdivided into patients who currently have the particular condition of interest; patients who have previously experienced the condition of interest; and patients who have no direct personal experience with the condition of interest but, because of their experience with similar or related conditions, may be expected to have a better understanding of the condition interest than members of the general public.

The limited evidence to date suggests that there is little difference between the scores for hypothetical health states obtained from the general public and the scores for hypothetical health states obtained from patients or others with greater disease-specific knowledge relevant to the states (Feeny *et al.*, 1991). On the other hand, there is equivocal evidence that patients currently experiencing a state may rate it differently than the general public (Sackett and Torrance, 1978; Llewellyn-Thomas *et al.*, 1993). Normally the direction of this difference is such that those in a particular health state see it as not as bad as those for whom the state is hypothetical. This may be caused by the change in perspective and the natural adaptability of people to circumstances in which they find themselves.

It can be argued that in a publicly funded system, such as that in Canada, the preferences of the general public are the relevant ones. Not only is the general public the ultimate payer of the system, but also members of the general public are at risk for the various diseases and health states under study. Certainly it would seem reasonable to use the general public's preferences when evaluating preventive programs (e.g. screening), because these are designed to prevent the general public from entering undesirable health states in the future. On the other hand an argument can be made that, for

treatment decisions for persons with already impaired health, the patients' preferences are more relevant. There is, however, a practical issue here: that each patient can only provide a preference score for the one health state in which he/she is in, and any other preference scores provided by the patient are hypothetical. Thus, any analysis that attempts to gather preferences on health states only from individuals currently experiencing that health state has a very large logistical problem of patient recruitment.

In some studies health professionals, such as nurses and physicians, have been used as the source of preferences for outcomes. They are both convenient and knowledgeable. They are, however, not a representative sample of the population at risk for incurring the outcomes. Before they can be recommended as a valid source of preferences, there is a need to demonstrate that their preferences are good proxies for those of the general public and/or patients.

The appropriate source of preferences depends on the use of the analysis and the viewpoint. For provincial drug plans, which are tax supported, the appropriate viewpoint is societal and the appropriate source of preferences for outcomes is the informed general public. An informed member of the general public is one who fully understands what it would be like to have the outcome under consideration. One way to provide this information would be to measure the preferences of patients with the outcome, and provide this information to the general public. Although the research has not been done, perhaps many in the general public would "go along with" the preferences of the patients. Thus, a circular argument can be made that patient preferences may be a good proxy for the preferences of an informed general public. Also, as discussed above, patients provide similar preferences to the general public when both are rating hypothetical states.

In summary, for provincial formulary decisions the appropriate source of preferences is a well informed general public, but patients may sometimes be a suitable proxy. Investigators must rationalize their source of subjects for preferences, and must describe the exact population from which the preferences were derived as well as the precise methods of measurement (including a description of health state scenarios).

Annotated Bibliography:

Gold MR, Patrick DL, Torrance GW, Fryback DG, Hadorn DC, Kamlet MS, et al. Identifying and valuing outcomes. In: Gold MR, Russell LB, Seigel JE, Weinstein MC, editors. Cost-effectiveness in health and medicine. New York: Oxford University Press; 1996. p. 82-134.

This chapter of the Washington Panel report explores the issue of whose preferences should be used in CEA. It examines issues such as whether to obtain preferences from the general public or from patients, whether there is such a thing as an average preference, whether community preferences discriminate against persons who are ill or have disabilities, what to do about populations with different preferences, and how to deal with practical considerations.

6.6 Equity

Guideline 16. Equity

All equity assumptions, whether implicit or explicit, must be highlighted in any analysis intended for use in the resource allocation process. Results should be presented using equal weights for all lives, life-years or QALYs, but the presentation should be sufficiently transparent to make it feasible for decision-makers to substitute different weights. Analysts should identify which groups of individuals would be the main beneficiaries if the program were implemented.

An all-encompassing discussion of the concept of equity is beyond the scope of this document. It is an important discussion, however, as equity considerations are a major factor in any economic evaluation that is being used for resource allocation purposes (Culyer, 1993; Ubel, DeKay*et al.*, 1996; Williams, 1997). By assuming equity for the base case of an economic evaluation, it is assumed that every patient involved in a clinical trial and/or economic analysis has a fair opportunity to participate and achieve the outcome(s) of interest.

In practical terms this means that the value of the lives or life-years or QALYs impacted by a given intervention are assumed to be equal no matter what the age, gender, socioeconomic status, etc. of the individual or group of patients who make up the target population of the intervention. It is important for the investigator to clarify the study's stance on equity, especially for those situations where an economic evaluation is carried out alongside a clinical trial. (By definition, a clinical trial, with all its inclusion and exclusion criteria, compromises the equity principle to a certain degree.)

In more practical terms, each of the various types of economic analyses and analytic techniques already contain equity assumptions. It is important to understand them, and to highlight them in the analysis. For example, a CEA using cost per life saved (or life-years gained) contains the equity assumption that all lives saved (gained life-years) are considered equal regardless of the age, co-morbidity, or other circumstances of the individual. It should be noted, however, that the assumptions underlying CEA may not adequately reflect the importance placed on equity by "real" individuals (Ubel, DeKay,*et al.*, 1996).

A CUA contains the equity assumption that a quality-adjusted life-year gained is equivalent regardless of who gains it. That is, it is equally preferable to add a quality-adjusted life year by extending the life of an 80 year old individual or a 30 year old individual. Note that, as with the other means of economic evaluation, all the utility assessment methods are technically neutral with regard to their equity assumptions and are able to accommodate weights for different individuals if society were able to agree on such weights. The limited research on this issue of weights is inconclusive (Williams, 1988), and the best advice at the moment is to use equal weights for all lives, life-years, or QALYs. Analysts should highlight the fact that they have used equal weights in the analysis. In addition, to facilitate the substitution of different weights by "users", analysts should ensure transparent and disaggregated presentation of study methods and the values used. For further discussion and critique of the equity assumptions and practices of cost-utility studies see Loomes and McKenzie (1989),

Gafni and Birch (1991), Broome (1991 and 1993), Sen (1991), Williams (1996), and Bleichrodt

(1997).

Cost-benefit analysis using the WTP technique contains an equity assumption that the existing distribution of income is appropriate, as mechanisms such as taxation exist to change income distributions. As with any market-based mechanism, WTP is limited by ability-to-pay and, thereby, programs aimed at conditions of interest to the wealthy can score better on WTP measures than programs aimed at conditions of interest to the poor.

Some consider this a serious limitation of CBA (Richardson, 1990; Richardson, 1994; Gold, Russell, *et al.*, 1996), while others argue that this is simply a matter of accepting the current distribution of income and maximizing welfare given that distribution (Johannesson and Jonsson, 1991). The argument in favour of the WTP method is that one can redistribute income, if one wants, in a variety of other ways. As an alternative, some researchers have asked WTP questions in terms of percent of income and have averaged these percentages in an attempt to reintroduce equity and to overcome this problem (Thompson, 1986); but this method violates the welfare theory behind WTP.

It is important for investigators to provide explicit information in defining the equity assumptions made in the analysis . For instance, the groups towards which the intervention is directed need to be defined (e.g. young versus old, rural versus urban). The specific changes observed in these groups (and the magnitude of these changes) also require full documentation in the body of the report. Related to this is the need to identify which groups of individuals are expected to be the prime beneficiaries, based on the perspective taken; that is, which population receives the benefit of the intervention, and what is the magnitude of that benefit (e.g. a large number of individuals receiving a small benefit versus a small number of individuals receiving a large benefit). This "big picture" or summary aspect of the evaluation should be reported as part of the aggregate or overall impact analysis section of the report (see part 5 of Section III - Reporting Structure).

In addition to identifying the prime beneficiaries, it may also be possible to identify losses that might be associated with an intervention. These losses may present as decreases in workload (and, therefore, jobs) created by the elimination of health care resource utilization due to the impact of a new intervention. In terms of health outcome losses, one example is the impact of an adverse drug reaction which seriously affects a minority of patients who receive the drug, or has a minor impact on a majority of patients involved in the intervention. It is ultimately up to authorized decision-makers, <u>not</u> the analyst, to provide the necessary weights or judgements to determine the redistribution of resources within various sectors of society given the implementation of the new intervention (Baltussen, *et al.*, 1996).

To reiterate, these Guidelines do not propose any particular equity criterion, but do recommend that analysts be aware of and state clearly the equity criterion they are using. Investigators are encouraged to refer to the papers cited herein for further discussion on the role of equity in the context of economic evaluations and allocative decision-making.

6.7 Discounting Future Outcomes

Guideline 17. Discounting Future Outcomes

Future outcomes should be discounted at the same rate as costs. The base case discount rate is 5% per year. This rate must be varied in a sensitivity analysis, with a discount rate of 0% (no discounting) at minimum. Analysts should also consider using a 3% rate for comparability with future studies. When it is believed the analysis should differentiate between discount rates for outcomes and costs, these results should be presented as a supplementary analysis and the relevance fully explained.

There is some controversy regarding whether or not future outcomes should be discounted, and at what rate (Olsen, 1993). Should future life years, life-years gained, quality-adjusted life years gained, and WTP amounts for future events averted be discounted? Most people are comfortable with the latter suggestion, that is, discounting future amounts of WTP; but many have some discomfort with the concept of discounting health states in term of future lives, life years or quality-adjusted life years.

Strong arguments can be made that these future outcomes must be discounted, and at the same rate as costs, to avoid paradoxical results (Weinstein and Stason, 1977; Keeler and Cretin, 1983; Cairns, 1992; Katz and Welch, 1993; Drummond and Jefferson, 1996). Others have rejected these arguments and stated that the most appropriate course is to discount future outcomes at a rate of or close to zero (Parsonage and Neuburger, 1992; Hillman and Kim, 1995). Experts also call for increased research on individuals' time preferences in order to come to a resolution regarding the differential discounting of costs and benefits. Recent work in this area by MacKeigan *et al.* (1993) and Gafni (1995) demonstrates the difficulties which lie ahead in resolving the time preference issue.

The developers of these Guidelines are more convinced by the arguments in favour of discounting than those opposed, and recommend that future outcomes be discounted and at the same rate as costs, i.e., at 5% per year (see Section 7.5). The discount rate should, however, be varied in a sensitivity analysis. At minimum, a sensitivity analysis involving a discount rate of 0% should be carried out in order to assess the impact of the above argument. In addition, it is suggested that sensitivity analysis based on a 3% rate should be considered, in order to allow comparisons with studies which will be using the 3% rate required by the Washington Panel reference case (Lipscomb, *et al.*, 1996).

7.0 Measuring and Reporting Costs

7.1 Cost Identification

Guideline 18. Cost Identification

A probability tree of the therapeutic pathway which describes all relevant downstream events should be provided, when appropriate. From the societal viewpoint, cost items that should be included are all direct health care costs, social services costs, spillover costs on other sectors, and costs that fall on the patient and family. Cost items that should be excluded are those not relevant to the therapeutic pathway such as those not related to the treatment being evaluated, costs relevant only to the clinical trial, and transfer payments such as sickness pay, unemployment insurance and welfare payments.

When relevant, lost time should be documented and reported as part of the description of the impact of the intervention. If HRQOL is an outcome measure in the study, some lost time will likely contribute to changes in HRQOL. Depending on the viewpoint, some lost time will represent a real cost in terms of lost resources and should be included as a cost item, but should also be tested with sensitivity analysis.

Cost identification involves identifying all the relevant resource items for subsequent measurement and valuation. A useful first step is to develop a probability or decision tree of the therapeutic pathway which describes all relevant downstream events. Then viewpoints for the analysis are selected, and resource items that are applicable to each viewpoint are identified. In the comprehensive societal viewpoint, all costs related to the therapeutic pathway should be included; however, **transfer payments** (e.g. sickness pay, unemployment insurance, welfare payments) should not. If subsidiary viewpoints are presented in the analysis, they should contain the subset of cost items relevant to that viewpoint but which were excluded from the primary societal analysis. This means that subsidiary analyses may include transfer payments if they represent a cost or savings from the viewpoint in question.

From a societal perspective, resource items that should be included are all direct health care costs, social services costs, spillover costs on other sectors (e.g. additional educational costs related to the proportion of children who "graduate" from neonatal intensive care units with learning disabilities), and costs that fall on the patient and family (i.e. direct patient costs, time costs and productivity costs [sometimes borne by the employer]). Costs that should be excluded from the analysis are all those **unrelated costs** that are not specifically attributable to the therapeutic pathway and its consequences.

Unrelated Costs

A cost item may be deemed not relevant because it is caused by an event that is unrelated to the treatment being evaluated (e.g. the costs of a broken leg would normally not be counted in evaluating an acne drug). One option for determining which clinical events are related (or not) to the intervention is via an adjudication committee (blinded to treatment assignment). This allows the investigator to remove unrelated events in an unbiased manner. (Note that the "noise" created by

unrelated events should not be a problem in large RCT-based studies, but will increasingly be a problem for smaller studies. The resolution to this problem is to "strip out" the noise in an unbiased manner [i.e. do one analysis with the unrelated events excluded, and then a sensitivity analysis with them included]). Finally, protocol-driven costs should typically be excluded if they would not occur as part of the intervention on a regular basis (e.g. extra monitoring and tests that patients receive just because they are in a trial).

Future Health Care Costs

One of the more contentious issues in the economic literature is that of dealing with **future health care costs**; that is, the costs associated with patients who live longer and consume health care resources as a result of a given intervention.

On the one hand, some authors argue that the decision to include future health care costs should be based on two predominant considerations. First, future costs should be judged by their relationship to the intervention; for example, any additional care required during "added years of life" as a direct consequence of the program in question. For instance, future costs of care for patients who survive septic shock via a new intervention should include the cost of treating the underlying condition which is now an issue as a direct consequence of giving the new therapy. Alternatively, the impact of a new drug for high cholesterol produces added years of life which occur far into the future. One would not be expected to include the treatment costs of clinically unrelated diseases (e.g. cancers) during the added life years, because these treatment costs are not a necessary and direct consequence of the specific intervention. The second factor which will influence the inclusion of future health care costs is the availability of data.

Others argue for inclusion of all future medical and non-medical expenditures based on grounds of correct economic theory. They argue that these costs should be taken into account in order to be consistent with a lifetime utility maximization model (Meltzer, 1997; Weinstein and Manning, 1997).

CCOHTA's *A Guidance Document for the Costing Process* (CCOHTA, 1996) concludes that future costs of unrelated diseases should be excluded from evaluations at this time. However, investigators are encouraged to consider exploring and evaluating the issue of future health care costs when planning their research protocols.

Cost of Lost Time

Another area of controversy is the **cost of lost time** or **indirect non-medical costs** (i.e. productivity costs, formerly referred to as **indirect medical costs**) as they relate to economic evaluations taking the societal perspective. At issue is where to incorporate this cost category, and how to place a value on it (with the latter being the more difficult of the two considerations).

Patients and/or family members can lose time from work and other activities as part of illness and treatment. For family members, time may be lost in taking patients for treatment, visiting patients in hospital, or caring for patients at home. The amount of lost time, by whom, and lost from what (work, other major activity, leisure) should be recorded. At the most basic level, these data should be reported as consequences of the intervention. For instance, where short-term absences from work

are included in an economic evaluation, justification for their inclusion should be given, and the data from the economic evaluation should demonstrate that productivity decreased as a consequence of treatment or illness.

Valuing time outside the workforce (e.g. volunteer time) is even more problematic, as the individual who loses time is not in the workforce. While volunteer time is a net resource cost, attaching a numerical value to this contribution is difficult. Some argue for its omission, stating that there is no net impact of this time loss on conventional gross economic indicators. Normally, however, caregiver time costs are included in the analysis; and there are two general approaches which have been used for valuation purposes. One option is to use a replacement cost estimate based on the market value of the services being delivered to the patient at home. Alternatively, one could use the opportunity cost method where the value of time in the home (or used consuming leisure) is at least equal to what could be earned in the labour force.

Placing a value on lost time has been the focus of papers by Koopmanschap and his colleagues (Koopmanschap and Rutten, 1993). They have proposed the friction cost method as an alternative to the HCA for incorporating work absence and productivity losses into economic evaluations (Koopmanschap and van Ineveld 1992; Koopmanschap*et al*, 1995; Koopmanschap and Rutten, 1996). This may be an alternative means of accounting for lost time, although it omits the value of the patient's time in the analysis (which is contrary to welfare economic theory) and is most correctly used in a non-full employment scenario.

Controversy also emerges in determining >where to value' the lost time. Should it be valued on the cost side (i.e. in the numerator) as a change in resources; or on the benefit side (i.e. in the denominator) as a change in HRQOL; or on both? The literature reveals definite differences of opinion in this regard. The Washington Panel (Luce,*et al.*, 1996) describes two distinct effects of a given intervention on patient time: the lost/gained time (i.e. productivity change) due to treatment, versus the health effects (e.g. pain, discomfort and treatment) resulting from the time invested in the intervention. The Panel's overall recommendation is that the numerator should include "only the opportunity cost of time in treatment" (i.e. the value of the patient's time due to seeking treatment), while adjustment for any unpleasant/painful treatment interventions should be accounted for by the QALY measurement (or other preference weight) in the denominator.

The issue is not as simple as this summary statement might suggest, and there have been competing recommendations put forth in the literature (Brouwer,*et al.*, 1997). This group disagrees with the measurement of productivity costs via changes in quality of life described by the Washington Panel. They recommend the friction cost methods as a more accurate means of calculating lost productivity from a societal perspective. Clearly the debate will continue on this topic. Readers are encouraged to compare the arguments put forth by both camps when considering the issue of lost time.

Summary

Given the current controversy about the proper treatment of lost time and given the reasoning above, these Guidelines support the following position. All lost time, whether work time or leisure time, should be documented and reported as part of the description of the impact of the intervention. If HRQOL is an outcome measure in the study, some lost time will likely contribute to changes in HRQOL (as measured in the denominator). Depending on the viewpoint, some lost time will also represent a real cost in terms of lost resources (as reflected in the numerator). This should be identified and explained within the context of the study.

Given the potential controversy surrounding the issue of lost time, if the amount is significant it would be wise to undertake sensitivity analyses to display the impact of alternative assumptions. Following these recommendations will mean that some lost time contributes both on the benefit side (to changes in HRQOL) and on the cost side (to changes in cost). This does not represent inappropriate "double counting". It simply represents the appropriate counting of different impacts of the treatment.

In summary, analyses should exclude unrelated costs, exclude future health care costs that are not directly related to the intervention, and exclude those costs which are identical to the alternatives under consideration. Steps must be taken to ensure that costs derived from varying sources are calculated using the same base year (i.e. net of inflation). Finally, it should be kept in mind that the controversy surrounding the issues of time lost (for caregivers or for patients), future health care costs, unrelated costs, etc., is theoretically based and requires much discussion and research in the future.

Annotated Bibliography:

Koopmanschap MA, Turren FFH, van Ineveld BM, van Roijen L. The friction cost method for measuring indirect costs of disease. Journal of Health Economics 1995;14:171-189.

The authors review the debate surrounding the inclusion of indirect costs within an economic analysis. They propose the friction cost method as an alternative to the human capital approach as a means of estimating indirect costs. This method explicitly considers the economic impact of disease (via absenteeism, disability and mortality) on wealth lost to society via reduced productivity; and takes into account factors such as the local labour market and social insurance premiums, amongst others.

7.2 Cost Measurement

Guideline 19. Cost Measurement (Resources Used)

Resources used in treatment must first be described in natural (non-dollar) units. All resource utilization data derived from international trials must be validated for Canadian practice.

Costs are the product of a vector of the quantities of resources (Q) and the unit prices of resources (P). **Cost measurement** consists of determining the quantities, Q, of resources (i.e. health care resources, non-health care resources, informal caregiver time, patient time for treatment) used as part of a given intervention. Cost valuation (see Section 7.3) consists of determining the unit costs/prices, P, of these individual resources. It is important to separate these two concepts, in part, because of the potential to use standard costs for valuation.

Where should one go to determine the resource consumption associated with a particular product? In considering drugs that go through multiple trials during their development, the later trials would more nearly match the actual therapeutic pathway of final use and would be the appropriate source for the resource quantities. In considering international trials, it should be noted that resource quantities cannot be directly imported into the Canadian system, because of the major differences in the way that health care is delivered in many countries (see Section 9.2 re: portability issues). As a minimum, resource quantities must be re-validated for Canadian practice. Some may, in fact, be transportable into Canada, but an explanation and justification is required. The default assumption is that resource quantities are not directly transportable.

Note should also be made regarding the methods by which one analyzes the uncertainty inherent in resource utilization versus unit price data. The former should be subjected to inferential statistical analysis, while with the latter uncertainty should be evaluated via sensitivity analysis.

7.3 Cost Valuation

Guideline 20. Cost Valuation (Unit Prices)

Economic definitions of costs must be used and the concept of opportunity cost recognized. Investigators performing analyses in the Canadian setting should refer to the CCOHTA *Guidance Document for the Costing Process* for further direction regarding costing issues.

Before beginning the cost measurement and valuation process, investigators should refer to the CCOHTA *Guidance Document for the Costing Process* (CCOHTA, 1996) for direction on relevant approaches, especially as they relate to direct (medical and non-medical) costs. The *Guidance Document* is useful in that it highlights the fact that multiple means exist for estimating costs depending on the level of precision needed for the analysis. This document is reviewed regularly and will be updated as new information and techniques become available.

The economic definition of costs should be used in cost valuation, not the accounting definition. Thus, for example, buildings which have been written off in the accounts and no longer incur a depreciation cost would still have a cost from the economic point of view. Also, all land, buildings and equipment would have an opportunity cost of the capital tied up in investment -- a cost which is normally not acknowledged in accounting records.

To the extent possible, standard cost values should be used in costing out the utilization of resources. This will improve the generalizability of the results and the comparability across studies. Situations arise in carrying out economic analyses where actual unit prices are collected as part of the study process (i.e. a standard cost list is not used). In reporting these values, the investigator must clearly define whether mean or median values have been incorporated, along with a justification for using the chosen form of central tendency.

Attaching values to lost work time continues to be a contentious issue. Since the average wage rate may be affected by outlier wages within the study sample, it is suggested that the overall average industrial wage rate be used when determining costs due to lost productivity. (This is consistent with the equity discussion from Section 6.6.) It may also be useful to test the effect of this parameter on the results of the economic evaluation by using a justifiable range of wage values (e.g. including gender differences in wages) in the sensitivity analysis.

7.4 Discounting Future Costs

Guideline 21. Discounting Future Costs

As with future outcomes, all studies must discount future costs at an initial rate of 5% per year in the base case. This rate must be varied in a sensitivity analysis, with a discount rate of 0% (no discounting) at minimum. Analysts should also consider using a 3% rate for comparability with future studies. If differential discount rates are to be used for outcomes and costs, then the results should be presented as a supplementary analysis and the relevance fully explained.

Future costs, and indeed future effects (see Section 6.7), are discounted to reflect the fact that, in general, individuals and society have a positive rate of time preference (Gafni and Torrance, 1984; Lipscomb, 1989; Krahn and Gafni, 1993). In general, people prefer desirable consequences (like benefits) to occur earlier and undesirable consequences (like costs) to occur later. Thus, future benefits are discounted to reflect the fact that they are worth less simply because they occur in the future rather than the present. Similarly, future costs are discounted to reflect the fact that people prefer them to occur in the future rather than the present. (Note that, because of discounting, the impact of these costs may be diminished and may not amount to much for certain disease states.)

There is, however, no reason that the rate of time preference should be a constant percent per year. Indeed, empirical studies suggest that individuals do not exhibit a constant annual rate of time preference (Fuchs, 1982). An alternative model, relative value discounting, has been proposed which uses a varying discount rate that is higher in near years and lower in far years (Harvey, 1994). This model is still experimental, but may hold promise for the future.

The proper rate of discount is controversial (Jenkins, 1981; Krahn and Gafni, 1993), and no precise gold standard exists. For comparability across studies, however, it is important that all studies use a common discount rate. The standard discount rate for the base case analysis is set at 5% per year. Variations about this rate should be undertaken using sensitivity analyses. At minimum, a rate of 0% should be investigated to clearly display the impact of discounting. In addition, it is suggested that sensitivity analysis based on a 3% rate should be considered, in order to be comparable with studies which will be using the 3% rate required by the Washington panel (Lipscomb *et al.*, 1996).

8.0 Dealing with Uncertainty

Guideline 22. Dealing With Uncertainty

All studies must clearly address the issue of uncertainty (whether it arises from sampling error or from assumptions) and justify the methods used. Sampling errors can be dealt with by making use of confidence intervals. In addition, for each important assumption, alternative plausible assumptions must be included. Investigators are encouraged to use approaches such as Monte Carlo simulation, which varies all factors simultaneously.

Economic evaluations of treatment options require the analyst to combine information on the course of disease, the clinical effectiveness of each treatment regimen, costs of the interventions and their sequelae, all of which come from one or more of the following sources:

1) primary data gathered prospectively from clinical trials and other study designs;
2) estimated data based on a meta-analysis of articles (published and unpublished) found in a literature search;
3) unpublished data (e.g. collected by a manufacturer, cost data);
4) parameters estimated by a panel of experts (e.g. estimating patient compliance rates for effectiveness calculations);
5) methodological assumptions (e.g. selecting a discount rate, or an approach to costing lost time); and,
6) modelling assumptions (e.g. projecting data into the future using an epidemiological model of the disease and its treatment).

The cost-effectiveness ratio (as well as CUA or CBA ratios), therefore, incorporates parameter uncertainties and modelling uncertainties, the latter of which include model structure and model process. (Little work has gone into evaluating the uncertainty introduced via the means through which analysts develop models, and thus will not be explored in this document).

Traditionally, such uncertainties have been examined using sensitivity analyses based on either traditional/deterministic (one way, two-way, three-way or threshold analyses) or probabilistic methods. However, Manning *et al.*, (1996) have reported an increased interest in developing statistical measures of uncertainty in the estimated cost-effectiveness ratio. There is also an increasing body of literature regarding the role of statistical analysis as it applies to prospectively collected, patient-specific data in stochastic cost-effectiveness studies (Willan and O'Brien, 1996). No matter what the methods used, it is important that the exact means by which uncertainty has been evaluated (including reference to pertinent software programs) is communicated clearly in the final report.

Before reviewing the details of the aforementioned approaches, it is of value to alert "doers" to the importance of clear presentation of results (especially those reflecting the uncertainty associated with the analysis) from the perspective of the "users" of economic evaluations. Investigators need to recognize that there are issues involved in making drug policy decisions above and beyond the clinical and economic results of their report. In order for decision-makers to put the results of an analysis into perspective, a report's concluding statements should incorporate confidence bounds on the results (see below) to communicate a sense of the investigator's certainty in his/her work. This issue

of determining "policy important differences" as they relate to the contribution of pharmacoeconomic studies is a rapidly evolving area of research. In the meantime, it is important that studies address the issue of uncertainty and rigorously justify the methods used to help provide decision-makers with a sense of the reliability of the analysis.

8.1 Statistical Approaches

A well-known approach to dealing with parameter uncertainty (or data sampling error) is the confidence interval, from the frequentist (or classical) statistical literature. The confidence interval is a numerical range calculated to have a particular probability (typically 95%) of containing the true value of the parameter of interest. In a pharmacoeconomic study, however, statistical issues of dealing with sampling error are often more complicated than in randomized trials. The cost-effectiveness ratio, for instance, is made up of two parameters at once (i.e. costs and effectiveness). This requires a generalization to a confidence region (Sacristan, *et al.*, 1995; O'Brien and Drummond, 1994). There are various ways of determining the uncertainty in the estimate of cost-effectiveness ratios, as outlined below.

Determining a confidence interval around cost-effectiveness ratios is a complicated process, and can involve approximations performed in computer based simulations (i.e. probabilistic Monte Carlo simulations). Other methods include: the Fieller's theorem method; the Taylor series or delta method; non-parametric bootstrapping; and jackknife estimation techniques. This is a developing area in economic evaluation, and investigators are encouraged to explore these methods when applicable to their analyses. Readers are referred to several articles and texts for further discussion of these methods and their application (Selvin, 1991; O'Brien, Drummond, *et al.*, 1994; Mullahy and Manning, 1995; Sacristán *et al.*, 1995; Chaudhary and Stearns, 1996; Mullahy, 1996; Laska, *et al.*, 1997; Polsky, *et al.*, 1997).

The determination of confidence intervals as an approach to characterizing uncertainty is useful for at least three purposes: for testing hypotheses about the direction and magnitude of costs, effectiveness and the cost-effectiveness ratio; to give decision-makers a sense of how much confidence to place in the results of the analysis; and as a means of assessing whether or not the study sample size was sufficient to provide clinically and economically meaningful power to choose among treatments. It should be noted that sample size calculation is still an enormous challenge in economic evaluations, as there are ethical considerations in increasing sample size beyond that needed for clinical purposes. Sufficient statistical power for economic analysis may not always be possible for this reason, and research on this topic is underway. In addition, problems in defining an economically meaningful difference (similar to defining a clinically meaningful difference) have still not been resolved.

8.2 Sensitivity Analysis Approaches

In most pharmacoeconomic studies the uncertainties created by sampling errors are minuscule compared to the potential errors that can be created by using unverified assumptions. Weinstein and Stason (1977) argue that sensitivity analyses (SAs) are fundamental to cost-effectiveness analysis, as it is the nature of pharmacoeconomic studies that unverified assumptions are almost always used in at least some parts of the analysis. The uncertainty inherent in these data is evaluated by performing

a SA which, ultimately, provides an assessment of the certainty of the conclusions of the analysis. For each important assumption, alternative plausible assumptions should be investigated (Briggs, 1995; Briggs, *et al.*, 1994; Krahn, *et al.*, 1997).

A one-way sensitivity analysis involves varying critical components of the analysis across a meaningful range of possible alternative values, and then recalculating the cost-effectiveness ratio. The range used for the analysis should be made transparent to the reader; and there should be a clear description of the means by which the boundaries of each estimate were derived and used in the SA. At minimum, these boundaries should include the upper and lower values of the range of the data points used, and an analysis of the clinically plausible extremes (i.e. best case and worst case scenarios) is also recommended. The difference between the base case ratio and the cost-effectiveness ratio arrived at through SA provides some indication of how sensitive the results might be to changes in the parameter. This allows for a systematic way of examining the robustness of the conclusions.

These analyses provide useful information in terms of identifying the variables which play an important role in the results of the analysis, as well as the degree of robustness of the analysis, which is related to the definitiveness of the results. One-way SAs also have their drawbacks. First, much of the process is arbitrary, in that the analyst has to choose which variables to change and which to treat as fixed; the amount of variation around the base case value of the parameter that is considered clinically meaningful; and how much of a change in the base case result constitutes a robust finding. Important variations can be missed, since the incremental cost and effectiveness depend on multiple parameters and it is often the synergistic or interactive effects of two or more variables that can cause variations in cost-effectiveness ratios. It is important to use methods that allow for a more realistic approach to the interaction of factors, and two- and three-way SAs have been developed to overcome some of these limitations.

Ideally, one would like to vary all parameters simultaneously. Goldsmith, *et al.* have suggested a "factorial" approach to SA as one means of incorporating the wealth of information contained in data from clinical studies based on a factorial design (1987). Monte Carlo methods (see below) are also a means of varying more than a few components of the data at one time. Finally, one-way SA does not allow the calculation of 95% confidence intervals for the main outcome measures. These contribute to confident decision-making in that they provide some perspective of the upper and lower bounds for a given point estimate.

Traditional SA is generally considered a qualitative approach to evaluating uncertainty, and confidence intervals cannot be generated. In situations where there is the possibility of important variation in the cost-effectiveness ratios or suspicion of the interdependence of variables, a probabilistic sensitivity analysis, usually using Monte Carlo simulation, should be performed. These analyses vary all parameters simultaneously, and lead to 95% confidence intervals on the cost-effectiveness ratios (Doubilet, *et al.*, 1985). Probabilistic sensitivity analysis also has the advantage of providing information regarding the relative likelihood that the scenarios under consideration will occur (i.e. the likelihood that the intervention is cost effective).

Annotated Bibliography:

Doubilet P, Begg CB, Weinstein MC, Braun P, McNeil BJ. Probabilistic sensitivity analysis using Monte Carlo

simulation: a practical approach. Medical Decision Making 1985;5(2):157-177 .

An excellent review of the theory underpinning Monte Carlo simulations. Using a practical clinical example, the authors demonstrate the application of probabilistic techniques to a decision tree model of patient management options. While it is reassuring to know that computer programs will do all the work for us, it is important that the methods underlying these simulations are understood.

Manning WG, Fryback DG, Weinstein MC. Reflecting uncertainty in cost-effectiveness analysis. In: Gold MR, Seigel JE, Russell LB, Weinstein MC, editors. Cost-effectiveness in health and medicine. New York: Oxford University Press; 1996. p. 247-275.

This chapter of the Washinngton Panel report reviews the qualitative and statistical approaches to analyzing uncertainty in both the models and the parameters used in models for CEA. A practical clinical example is provided; and, as with every section in this text, useful general recommendations for dealing with this issue serve to summarize the chapter. (Note that other portions of this text also contain helpful discussion regarding sensitivity analysis.)

3. RESULTS AND DISCUSSION

9.0 Reporting the Analysis and Results

Guideline 23. Reporting Results

All results must be reported in disaggregated detail first, with aggregations and the use of value judgements (e.g. preference scores) being introduced into the presentation as late as possible. A probability tree of clinical outcomes should be provided for the relevant alternatives. Detailed technical reports, with patient confidentiality protected, should be made available to decision-makers. Reports should either follow the standardized reporting structure or be linked to it.

9.1 General Recommendations

Pharmacoeconomic reports must be detailed, clear and transparent. It is crucial that interested readers be able to follow thoroughly and exactly what was done and why. All results should first be reported in the form of the detailed steps of the analysis, with aggregations and the use of value judgements (e.g. preference scores) being introduced into the presentation of information as late as possible. An incremental or stepped approach for presenting results is also a useful means of presenting the various perspectives which may have been used in the analysis; for example, presenting the results in terms of a CCA, then a CEA, and finally a CUA and/or a CBA (as applicable).

If the study uses primary data reflecting short-term results obtained directly from clinical trials and modelled data for long-term results, a useful approach is to present the results separately for the short-term and for the long-term. An important aspect of clarity of reporting is to provide a probability tree of clinical outcomes (or other visual representation) for the relevant alternatives.

In terms of communicating the results of any sensitivity analysis carried out in the evaluation, the report should, at minimum, describe and clearly rationalize the methods chosen to assess uncertainty. It is important to clearly identify which components of the model and/or which parameters have the greatest impact on the outputs of the model (Mandelblatt,*et al.*, 1996). As noted in Section 8.1, confidence intervals for the CE ratio may also be included. A graphical representation (net cost versus net effectiveness) of the incremental cost-effectiveness results for each intervention may be a useful addition to the report (Seigel,*et al.*, 1996).

Each report should contain a distinct section within the Discussion which deals with the limitations of the evaluation, given the methods and assumptions which form the basis of the analysis. An objective review of the study's limitations fosters transparency in the communication of study results. Some of the limitations of the assessment may relate to the portability of results, which should be addressed as outlined in Section 9.2 below.

In addition to communicating the limitations of the evaluation, it is also important that the investigator convey a sense of his/her confidence in the results of the analysis. The end result of an evaluation (i.e. "the bottom line") is a highly composite number, and sensitivity analysis on this figure may not convey sufficiently the issues which represent the major sources of variation within the study. There is a tremendous need for further research in testing the uncertainty associated with utility measures, and statistically evaluating the cost and effectiveness measures within the primary data set. In the meantime, it is important that analysts be very transparent in their presentation of the data used in the analysis. They should also describe any relationships amongst the parameters evaluated, so as to provide decision-makers with a sense of the main factors which drive the results of the evaluation.

To enhance clarity and comparability across studies, especially for decision-makers who must review multiple studies, a structured format for reports has been developed (see Reporting Structure in Section III). This format includes a list of questions that must be addressed in compiling the report. Authors may also refer to guidelines for economic submissions developed by the *British Medical Journal* for additional suggestions for improving report content (Drummond and Jefferson 1996).

The journal article is not sufficient as a report format. Detailed technical reports, with patient confidentiality protected, should be available to decision-makers and to others who wish to examine them. Electronic submission of the report, including the data set, may be considered. Journal articles that summarize the study should make explicit reference to the availability of the detailed technical report.

Finally, investigators should make every effort to assure users of economic evaluations of the quality of the process underlying the study (see ***"Quality Assurance Tips"*** in Section V). Such assurance can be achieved by clear and thorough delineation of the conduct of the study, including how it was documented to ensure consistency and quality in the process; and by making documents specific to the quality assurance process available to users.

9.2 The Portability of Economic Evaluations

Guideline 24. Portability of Economic Evaluations

The portability of an economic evaluation is an issue which should be considered during the development of the study, as well as during the interpretation and dissemination of study results. Consideration must be given to two specific aspects of the applicability of the analysis to the local setting. The first aspect is the distinction between efficacy and effectiveness. The second aspect is the validity of transferring results (i.e. economic, clinical and humanistic) from one country or health care jurisdiction to another. These considerations are especially important when working in the context of multi-national, multi-centre trials.

(Note that there is no consistency in terminology related to this topic. Readers may find these same concepts referred to as "generalizability" or "measures of external validity". The authors of these Guidelines hope that the descriptions outlined below will provide clarity to this discussion.)

Users of economic evaluations are particularly interested in incorporating the results of a study into their decision-making processes. The extent to which users can generalize study conclusions to

populations and settings of interest which are external to the original evaluation is the focus of the portability issue (Mason, 1997).

There are two levels of applicability that must be addressed in any discussion of portability: that of the extent to which the evidence of the treatment effect obtained from the study population is *applicable* to the local setting; and the *validity of transferring the results* of an evaluation from one country, health care jurisdiction or setting to another (O'Brien, Heyland *et al.*, 1997). Both RCTs and **decision analysis** model-based evaluations face these problems (Oliver and Smith, 1990). Each issue must be addressed by the investigator when planning a study, and when interpreting and discussing the results of an evaluation. The issues of applicability and validity must also be carefully weighed by decision-makers before the results of the analysis are used in making resource allocation decisions.

The question of the applicability of the evaluation relates primarily to the matter of clinical efficacy versus effectiveness. Are the RCT-based clinical efficacy data which form the basis of the economic analysis (usually high internal validity) comparable to and predictive of the effectiveness which might be achieved in the setting for which the decision is being made? In order to address this issue, one must determine whether the conditions of the RCT (or the decision analysis model) compare favourably or are reasonably consistent with usual clinical practice in one's own jurisdiction.

The transferability problem of importing economic evaluations from one health care jurisdiction to another is founded in a combination of economic, clinical and humanistic concerns. One of the chief economic factors affecting the use of evaluations from other jurisdictions is variation in the relative unit price of resources among countries or regions (O'Brien, 1997). Relative prices are, in turn, impacted by clinical factors (see below) which ultimately determine overall resource utilization patterns (and, thereby, overall costs). The CCOHTA *Guidance Document for the Costing Process* (CCOHTA, 1996) notes that it is not sufficient to "Canadianize" international studies by using Canadian price weights; and that resource utilization patterns should be verified for Canadian practice (at minimum via costing studies or **Delphi panel methods**) before translating quantities to the local setting.

Clinical variation contributes significantly to the overall variability introduced by transferring study results. Clinical factors which come into play under the "transferability" umbrella include the following: the comparability of patient demographics and the epidemiology of disease in the study versus the local situation; clinical practice and usual care; as well as incentives or regulations for health care providers. A final component which must be considered in the portability discussion is the humanistic issue (e.g. patient responses to generic and disease-specific instruments). Neither investigators nor users can assume that humanistic outcomes evaluated within the context of an economic study will be consistent across countries or cultures (O'Brien, 1997). Cross-cultural validation of QOL measurement tools is an active area in outcomes research.

How does one take all these issues into consideration when transferring study results from one setting to another? O'Brien, Heyland, *et al.* (1997) have outlined an approach for interpreting economic analyses, in which they discuss the application of study findings in the context of one's clinical practice situation. Implicit in their outline is a spectrum which represents the certainty with which one can transfer information from one setting to another. For instance, efficacy data (does it work?) is probably more portable than effectiveness data (how well does it work?), which in turn is more portable than economic information. The implication is that, as one goes along the spectrum from efficacy to economic data, one needs increasingly more stringent proof of equivalence before study results can be transferred to the local setting.

Some would argue that clinical applicability/comparability should be the first criterion addressed in determining study portability. If one is unable to verify that comparable clinical efficacy/effectiveness exists, then there may be little benefit (and a lot of work) involved in proceeding with the verification of other clinical and cost data to complete the transfer. Whatever the approach, each of the major components of an evaluation (i.e. clinical, economic, epidemiologic, treatment patterns, treatment comparators) must be verified versus local conditions before a study's results can be considered transferable.

The discussion outlined above focuses on dealing with the portability issue after the study has been completed. What about addressing portability concerns in the study planning stages? As a means of giving direction to investigators, Drummond, *et al.*, (1992) have identified five issues which need to be addressed to facilitate the conduct and interpretation of economic evaluations on an international basis (see below).

Methodologic Tips on the Portability of Economic Evaluations
(see Drummond, *et al.*, [1992])

a) Develop a common economic methodology and agree on the type of evaluation that is feasible given data availability.

b) Address problems of data availability, which can be mitigated to a certain extent by separately reporting quantities of resources used and costs.

c) Address the influence of practice variations, which must be transparently communicated by including information regarding how diseases are defined and treated (e.g. dosage regimens, timing of surgical interventions, hospitalization rates, LOS) so that jurisdiction-specific information can be identified and used in the evaluation.

d) Facilitate the interpretation of results in different settings by presenting the results in a disaggregated manner for all appropriate viewpoints.

e) Address additional factors by incorporating local relative costs differences and patterns of care (remember to focus on the sensitivity of results to key parameters when interpreting for local application).

There is no precise process which is recommended at the present time to adapt studies from one jurisdiction to another. It is the responsibility of the investigator to think carefully about the issues discussed above in the planning, interpretation and communication of study results; and it is the responsibility of the "user" to think diligently when using study results in the context of decision-making.

It should be noted that the issue of international transportability is quite separate from that of international harmonization of guidelines for pharmacoeconomic studies. While it is possible that the future may bring a single set of international guidelines, it is also likely that such a document would emphasize the issues pertaining to the transfer of results across international boundaries or between different health care settings (Genduso and Kotsanos, 1996).

Guideline 25. Disclosure of Relationships

Funding and reporting relationships must be clearly described. The investigators must have independence regarding methodological considerations at all stages of the study, and must have the right of publication in the journal of their choice.

10.0 Disclosure of Relationships

Both the report and any journal article written from the report should state clearly the funding and leadership arrangements for the study. This should include a listing of all key participants in the study and their roles. It should include a description of all administrative arrangements regarding the study such as steering committees, management committees, adjudication committees, and the like, including membership and affiliation of members of the committees and voting arrangements for the committee (presence/absence of any veto power). The important principle is that the investigators should have independence regarding methodological considerations at all stages of the study.

As part of the research plan, the investigators should have the right of publication when and where they like. The pharmaceutical firm or other organization funding the study should be given thirty days to comment on any manuscript prior to its submission; however, all final decisions regarding submission are the responsibility of the investigators.

III. STANDARDIZED REPORTING STRUCTURE

Section II provided guidance on the design and performance of pharmacoeconomic studies. This section provides guidance on the planning and reporting of studies for those responsible for carrying out this type of research. In addition, the Standardized Reporting Structure is intended to provide a framework for decision-makers, to facilitate the incorporation of results of economic evaluations into the decision-making process.

The Reporting Structure provides a format to ensure that studies are reported adequately and in a consistent manner to facilitate their review and comparison. It is expected that all topics listed in this Section will be addressed in the report. If a topic is omitted, a brief reason should be given. Preferably, the report should be organized according to the topics provided. If the report is organized differently, a cover note or markings made directly on the report should be used to indicate where in the report each topic can be found.

In addition to overall structure, each topic in this section also provides a list of questions that will be in the reviewer's mind as reports are being analyzed. These questions are intended to serve as a reminder to authors of those aspects that they should clearly address in their reports. In addition, these questions should alert users to the important issues which must be considered when reviewing the analysis. The questions have been adapted from the Ontario guidelines for economic analysis of pharmaceutical products (Detsky, 1993; Ontario Ministry of Health, 1994).

Reporting Structure

1. EXECUTIVE SUMMARY

- Summary of study and bottom line result(s) in the form of a structured abstract which follows the reporting structure outline
- Interpretation in the context of all reasonable alternative therapies, as well as in the context of the limitations of the analysis.

Q1.	What is the "bottom line" result of the analysis in quantitative terms? The answer to this question should be presented in the form of statements such as the following: a) The cost per QALY gained for using this product compared to the alternative is $X or ranges from $Y to $Z; b) The use of this product compared to the stated alternative will result in expected incremental expenditure of per patient treated with a net reduction of Y major adverse clinical events (e.g. cardiac deaths) and Z minor clinical events (e.g. side-effects).

- Recommendation, if appropriate

2. INTRODUCTION

2.1 *General comments on the disease or condition*

- Pathology/condition
- Epidemiology
- Current clinical practice
- Economic impact

2.2 *Product description*

- Therapeutic classification, brand and generic names, dosage form, route
- Approved indication(s)
- Indication(s) for which economic evaluation is being carried out, including a description of the pathology, epidemiology, current clinical practice relating to this/these indication(s), and specific patient subgroups to which the analysis is being applied

2.3 *Objective of the study*

2.4 *Disclosure of relationships*

- Funding and reporting relationships, contractual arrangements
- Investigators' autonomy and publication rights

Q2.	Who performed the analysis? Have the authors of the report signed a letter indicating their agreement with the entire document presented? Does the report indicate that the authors had independent control over the methods and the right to publish the analysis regardless of its results?

3. METHODS

3.1 *Type of analyses*

- Prospective, retrospective, modelling, or mixture of methods
- Analytic techniques used, reasons and key assumptions which underpin the analysis (e.g. CMA, CCA, CEA, CUA, CBA)
- Study design and procedures used, statistical analyses and validation methods

Q3.	What is the question being asked in the report? Is the economic question relevant? What type of economic analysis is being performed to answer the question (i.e. cost comparison/minimization, cost consequence analysis, cost-effectiveness analysis, cost-utility analysis, cost-benefit analysis).
Q4.	Is the analysis an incremental analysis?

3.2 *Target audience*

- Target audience (decision-makers) for the study (e.g. formulary decision-makers, patient purchasers, prescribers)

3.3 *Viewpoint*

- Viewpoints selected and reasons (e.g. societal, Ministry of Health)

Q5.	Has the viewpoint or perspective for the analysis been stated clearly, along with the reasons for this choice? Is it a societal perspective, third-party payer perspective, or patient perspective? Is the analysis presented in a disaggregated fashion showing these perspectives separately?

3.4 *Treatment comparator*

- Comparators selected and reasons (e.g. no treatment/watchful waiting, drug,

surgery)

- Methods for analysis of clinical efficacy of comparator(s) (e.g. head to head in RCTs versus meta-analysis of separate trials)

Q6.	Has the study included a comparison of alternate treatments for patients with the same clinical condition? Are those alternatives explicitly stated? Are the alternatives chosen valid and reasonable?

3.5 *Time horizon*

3.6 *Related studies/background*

- Systematic review of previous economic analyses that concern the similar problems or the same or similar treatments

3.7 *Outcome measurement*

- Outcome variables
- Clinical efficacy results (method used, source of info, assumptions made)
- Socio-demographic data collected, clinical data (where applicable)
- Clinical outcomes included and how measured (e.g. adverse events, morbidity, mortality)
- Health-related quality of life instruments included (e.g. disease-specific instrument, generic HRQOL profile, preference-based measure)
- Other outcomes considered but rejected (with rationale)

Q7.	Has the evidence of the product's efficacy been established through randomized trials? Has this evidence of efficacy been supplemented by evidence of effectiveness applicable to the patient population or subgroups considered in the study? Has the latter evidence been derived from studies documenting routine use in clinical practice? Have all relevant and significant variations in effectiveness for different subgroups been identified and reported?
Q8.	Are the methods and analysis displayed in a clear and transparent manner? Are the components of the numerator (cost of each alternative) and denominator (clinical outcomes of each alternative) displayed? Are clinical outcomes expressed first in natural units and then translated into alternate units, such as benefits or utility?
Q9.	Are all important and relevant costs and consequences (outcomes), including adverse effects for each alternative, identified?
Q10.	How is HRQOL measured?
Q11.	Is HRQOL an important component of an economic analysis for this question? Based on the sensitivity analysis, how sensitive is the estimate of cost-utility to variations in HRQOL?

3.8 *Cost measurement and valuation*

- Resource items included and how measured (e.g. direct costs, costs of lost time, spillover costs on others sectors, spillover costs on other individuals)
- How were prices determined?
- Source of information or resource utilization data collection

- Assumptions made
- Discounting for costs and outcomes

Q12.	Are costs and consequences modeled (as in a decision tree) with information derived from a variety of sources or estimated directly from (a) specific patient population(s)?
Q13.	Are capital costs and overhead costs included as well as operating costs? How are they measured?
Q14.	How have indirect costs (i.e. productivity costs, cost of lost time) been identified and estimated?

3.9 ***Uncertainty***

- Outline uncertainty due to sampling error versus that derived from the range of plausible assumptions
- Outline those aspects of uncertainty addressed via statistical methods (e.g. confidence intervals) versus sensitivity analysis

3.10 ***Sub-group analyses***

- Are there *a priori* identifiable subgroups for which differential results might be the case (e.g. effectiveness subgroups, preference subgroups, cost subgroups, cost-effectiveness subgroups)

3.11 ***List of assumptions***

- Identify major assumptions and limitations (both economic and clinical) contained within the analysis and how they might affect the results

4. RESULTS

4.1 ***Analysis and results***

- Presentation of all analyses in a clear, step by step fashion so readers can replicate the calculations if interested
- Display any models used, and the assumptions
- Presentation of results in detail first (e.g. study population, socio-demographic profiles, clinical data, resource utilization, etc.), with aggregations and the use of value judgements (e.g. preference scores) introduced into the presentation of the analysis as late as possible
- Interpretation of results in the context of all reasonable alternative therapies
- Statistical and sensitivity analyses
- Outline of important limitations of analysis in a transparent manner (i.e. what issues limit the results and their application/portability to other jurisdictions)
- Comment on how the health care setting (e.g. practice patterns) for the evaluation influences the results

Q15. For variables which are difficult to measure, what method is used to handle this difficulty? Does this method slant the analysis all in favour of one intervention in order to bias the analysis against the expected result?

Q16. How applicable are these results for the intended audience? Comment on the external validity of results and the validity of your data as they apply to the decision-maker at hand. Comment on any international or regional differences in the following parameters: epidemiology of the disease, treatment patterns, effectiveness, costs, and others of relevance. If differences exist, how likely is it that each factor would impact the results? In which direction? Of what magnitude? Taken together, how would they impact the results and of what magnitude? Given these potential differences, how would the conclusions likely change in the current setting? Are you able to quantify this in any manner?

4.2 *Sensitivity analysis results*

4.3 *Subgroup analysis results*

4.4 *Equity*

- Equity assumptions (e.g. a QALY is equal for all).
- Distributional considerations (e.g. primary beneficiaries)

Q17. What equity assumptions have been made in the analysis? For example, are QALYs gained by any individual considered equal?

5. AGGREGATE ANALYSIS

- Aggregate impact analysis

Q18. Is the incremental cost-effectiveness ratio estimated for a specific clinical indication that represents the majority or all of its expected use by those covered under the programs operated by the decision-makers to whom the report is addressed? Are there other indications which have not been considered which involve a large amount of utilization for which the ratio may be very different?

Q19. Is there an estimate of the aggregate incremental expenditure required for the provinces (or other decision-makers to whom the study is addressed) to provide this product to patients covered by their programs? What is the estimate of aggregate incremental costs? Does this estimate cover all of the major indications for use of the product?

6. DISCUSSION

- Should include discussion of limitations of study, robustness of results, bottom line caveats, methodologic issues, transportability issues

7. CONCLUSION

- Should summarize bottom line result, equity assumptions, aggregate impact, confidence in results

8. REFERENCES

9. APPENDICES

- Detailed tables of data
- Step by step details of analyses
- Intermediate results
- Copies (hard or soft) of data collection forms, questionnaires, instruments, etc.

IV. GLOSSARY OF TERMS*

Words which appear in the text in **bold** lettering are defined in the tables below.

Term	Definition/Description
Acquisition cost	The purchase cost of a drug to an institution, agency or person.
Analytic perspective	The viewpoint chosen for the analysis (e.g. societal, government, health care system, payer).
Average cost	Total costs of a treatment or programme divided by total quantity of treatment units provided (see also **marginal cost**).
Consequence(s)	The outcome(s) associated with a disease and/or intervention (e.g. stroke, death, side effects, avoided morbidity).
Contingent valuation	A method for evaluation of benefit or value to individuals of therapy that uses survey methods to establish willingness to pay.
Cost	A product of the quantities of resources (Q) and the unit prices of resources (P).
Cost measurement	The process of determining the quantity of resources (Q) used as part of an intervention.
Cost of lost time	See indirect cost.
Cost per QALY gained	A measure used in CUA to assist in comparisons among programmes; expressed as monetary cost per unit of outcome.
Cost-benefit analysis (CBA)	Type of analysis that measures costs and benefits in pecuniary units and computes a net monetary gain/loss (i.e. as net cost or net benefit) or a cost/benefit ratio.
Cost-consequence analysis (CCA)	Type of analysis that makes no attempt to aggregate across different kinds of consequences (e.g. strokes, deaths, side effects). Any weighting and aggregation is left to the user of the study.
Cost-effectiveness analysis (CEA)	Type of analysis that compares drugs or programmes having a common health outcome (e.g. reduction of blood pressure; life-years saved).
Cost-minimization analysis (CMA)	Type of analysis that finds the least costly programme among those shown or assumed to be of equal benefit.
Cost-utility analysis (CUA)	Type of analysis that measures benefits in utility units or utility-weighted life-years (QALYs); computes a cost per utility-measure ratio for comparison among programmes.

Term	Definition/Description

Decision analysis	An explicit quantitative approach for prescribing decisions under conditions of uncertainty.
Decision tree	A framework for representing alternatives for use in decision analysis.
Delphi panel method	A structured method of eliciting expert judgement, for obtaining data regarding effectiveness estimation.
Direct medical costs	Fixed and variable costs associated directly with a health care intervention (e.g. physician salaries).
Direct non-medical cost	A non-medical cost associated with provision of medical services (e.g. transportation of a patient to a hospital).
Discount rate	Rate of discount used to convert future costs and benefits into equivalent present values; typically 2 to 6% per annum for costs, and 0 to 6% for benefits. [It is, however, recommended in these Guidelines that the base case rate be 5% for both costs and outcomes, and that variations on this rate, including 0% and 3%, be provided as sensitivity analyses.]
Dominance	A comparison of the costs and effectiveness of each alternative treatment, which assists in defining the most appropriate comparators for use in the economic evaluation.
Effectiveness (of a drug)	The therapeutic outcome in a real world patient population (usually differs from efficacy determined in controlled clinical trials).
Efficacy	The therapeutic outcome determined in a randomized controlled trial.
Efficient frontier	In a graphical representation of the non-dominated comparators, the incremental cost-effectiveness or cost-utility ratios are formed along the efficient frontier.
Equity	Fairness in the allocation of resources or treatments among different individuals or groups.
Formulary	A list of drugs reimbursable under a health insurance plan or offered under a capitated or managed care programme or preferred in a particular clinical setting.

Term	**Definition/Description**

Friction cost method	A method of estimating the productivity costs by calculating the value of production losses during the friction period (i.e. between start of absence from work and replacement).
Future health care costs	Costs which result from the additional consumption of resources (via longer life span, etc.) due to a given intervention.
Health-related quality of life (HRQOL)	QOL measures that are likely to be influenced by health interventions.
Healthy years equivalent (HYE)	The hypothetical number of years spent in perfect health which could be considered equivalent to the actual number of years spent in a defined imperfect state of health.
Human Capital Approach (HCA)	A means of calculating the indirect cost of medical illness, based on the remaining lifetime economic value to society of a healthy individual of that age, measured by potential market earnings.
Incremental cost[a]	Difference between the cost of a programme (treatment) and the cost of the comparison programme.
Indirect medical cost	The cost of medical treatment (in life years) gained through an earlier intervention.
Indirect non-medical cost	The cost of reduced productivity resulting from illness or treatment (may be estimated by loss of wages and other means). New terminology: cost of lost time; productivity cost.
Intangible cost	The cost of pain and suffering occurring as a result of illness or treatment.
Marginal cost[a] (see also average cost)	The extra cost of one extra unit of product or service delivered (usually differs from average cost).
Markov model	A statistical representation of recurrent events over time that can be incorporated into decision analysis.
Meta-analysis	A systematic process for finding, evaluating and combining the results of sets of data from different scientific studies.
Net benefit	Benefit (in pecuniary units) minus total cost (in pecuniary units): a basic decision criterion in CBA.

Term	Definition/Description
Opportunity cost	The cost of using resources for some purpose, measured as their value in their next best alternative use.
Preference	Preference is a generic term and a concept that refers to the desirability of a health outcome. Both utility and value are special cases of the general term/concept of preference.
Quality of life (QOL)	Physical, social and emotional aspects of a patient's well-being that are relevant and important to the patient.
Quality-adjusted life year (gained) [QALY]	A common measure of health improvement used in CUA: combines mortality and HRQOL gains (outcome of a treatment measured as the number of years of life saved, adjusted for quality).
Revealed preference	Preferences revealed by the choices that individuals make. The choices may be those made by individuals in natural settings or responses to choices in questions posed by an investigator.
Sensitivity analysis	A process through which the robustness of an economic model is assessed by examining the changes in results of the analysis when key variables are varied over a specified range.
Standard gamble (SG)	A method of directly measuring utility, founded directly on the fundamental von Neumann-Morgenstern axioms of expected utility theory. A utility score is revealed by finding the probabilities in the gamble for which the respondent is indifferent between an uncertain alternative (the gamble) and a certain alternative.
Subgroup analysis	The process of analyzing data from subpopulations of patients which have been defined based on explicitly outlined parameters prior to the study.
Time trade-off (TTO)	A method of measuring value by finding the point at which the respondent is indifferent between two health states for different lengths of time. For chronic states, the choices are the index health state for time *t* followed by death, or perfect health for a shorter time followed by death. For temporary states, the choices are the index health state for time *t* followed by an explicitly specified outcome (usually healthy), or a worse health state for a shorter time followed by the same specified outcome.

Term	Definition/Description
Transfer payment	A payment (transfer of money) from one group to another without consumption of any physical resource; not recognized as a cost to society (e.g. taxation).
Unrelated costs	Costs that are not specifically attributable to the therapeutic pathway and its consequences.
Utility	In precise technical usage (common in Decision Science), utility is a cardinal measure of the preference for, or desirability of, a specific level of health status or a specific health outcome, measured under uncertainty. In the broader usage utility is often used (especially in Economics) interchangeably with the term preference.
Value	A cardinal measure of the preference for, or desirability of, a specific level of health status or a specific health outcome, measured under certainty.
Willingness to pay (WTP)	The maximum amount that a person is willing to pay: (i) to achieve a particular good health state or outcome, or to increase its probability of occurrence; or (ii) to avoid particular bad health state or outcome, or to decrease its probability.

[a] The concepts of incremental and marginal apply to costs, benefits and effects.

* Many of the definitions are reproduced, with permission, from **PharmacoEconomics**, published by Adis International Ltd.

22. QUALITY ASSURANCE TIPS

The following outlines the key quality elements in the process of conducting economic evaluations of pharmaceuticals.

Study Documents

The framework of the study should be defined in a protocol or research plan, including definition of the economic and clinical context, study objectives, study design, investigators, experts and/or patients selection criteria, and outline of the procedures for data collection, data management, statistical analysis, costing, economic analysis, modelling, whichever procedure applies to the study. Any data collection forms should be definable under one of the following formats: a questionnaire for expert interviews and quality of life or utility assessments; case report forms for primary data collection; and special forms for database extraction or data screens. The analysis plan should be written to detail the various statistical/economic analyses that will be used in the study. In the case of simple modelling studies, this analysis plan can be a separate section of the protocol.

The report should include the economic question, the clinical, epidemiological and economic context, the study objectives, the study design, the methods used in the study, the results and the discussion. All study documents should display a version number and date, and should be approved by a senior authority of the research group according to predefined procedures.

Data Collection

All data collection (primary and secondary) should be performed and documented following well-defined procedures. Economic evaluations conducted within or alongside a drug development program regulated by the Health Protection Branch (HPB) should follow regulatory requirements. If data collection is not regulated by HPB, ethical considerations and process indicators should be defined specifically in the study documents. In the case of secondary data collection (e.g. databases, interviews), studies should provide a description and validation of the data collection methods, as well as evidence of the means of review and approval by interviewees regarding the data collected.

Data Entry and Data Management

Data entry procedures should be defined within the research plan. Information pertaining to the type of data entry done (single versus double) and the software used should also be included. A data trail, an audit trail and an edit trail must be maintained, and adequate backup of the data should be available. Quality control of the data can be performed by running a coherency check program on the data collected. All data and programs from the evaluation should be kept either on a hard copy or as electronic files.

Data Analysis

Programs written and/or used for statistical, economic analysis or modelling should be kept either on hard copies or as electronic files.

Quality Control

Quality control should be applied at certain stages of the data entry and the analysis. A quality control verification of the tables included in the report should be done and documented. Reports should be reviewed by a senior investigator, along with documentation of an independent external review.

VI. REFERENCES

Appel LJ, Steinberg EP, Powe NR, Anderson GF, Dwyer SA, Faden RR. Risk reduction from low osmolality contrast media: what do patients think it is worth? **Medical Care** 1990;28(4):324-337.

Armitage PA, Berry G. **Statistical methods in medical research** 3rd ed. Oxford: Blackwell Science; 1994.

Bailey KR. Generalizing the results of randomized clinical trials. **Controlled Clinical Trials** 1994;15(1):15-23.

Baltussen R, Leidl R, Ament A. The impact of age on cost-effectiveness ratios and its control in decision-making. **Health Economics** 1996;5(3):227-239.

Begg C, Cho M, Eastwood S, Horton R, Moher D, Olkin I, et al. Improving the quality of reporting of randomized controlled trials: the CONSORT statement. **JAMA** 1996;276(8): 637-639.

Bellamy N, Buchanan WW, Goldsmith CH, Campbell J, Stitt LW. Validation study of WOMAC: a health status instrument for measuring clinically important patient relevant outcomes to antirheumatic drug therapy in patients with osteoarthritis of the hip or knee. **Journal of Rheumatology** 1988;15(12):1833-1840.

Bennett KJ, Torrance GW. Measuring health state preferences and utilities: Rating scale, time trade-off and standard gamble techniques. In: Spilker B., editor. **Quality of life and pharmacoeconomics in clinical trials**. 2nd ed. Philadelphia: Lippincott-Raven; 1996. P.253-265.

Berwick DM, Weinstein MC. What do patients value? Willingness to pay for ultrasound in normal pregnancy. **Medical Care** 1985;23(7):881-893.

Bleichrodt H. Health utility indices and equity considerations. **Journal of Health Economics** 1997;16(1):65-91.

Bloomfield Rubins H. From clinical trials to clinical practice: generalizing from participant to patient. **Controlled Clinical Trials** 1994;15(1):7-10.

Bootman JL, Townsend RJ, McGhan WF, editors. **Principles of pharmacoeconomics** 2nd ed. Cincinnati, OH: Harvey Whitney Books Company; 1996.

Briggs A. Handling uncertainty in the results of economic evaluation. **OHE Briefing** 1995;32(September):1-12.

Briggs A, Schulpher M, Buxton M. Uncertainty in the economic evaluation of health care technologies: the role of sensitivity analysis. **Health Economics** 1994;3:95-104.

Brooks R. EuroQol: the current state of play. **Health Policy** 1996;37:53-72.

Broome J. Utility. **Economics and Philosophy** 1991;7(1):1-12.

Broome J. QALYS. **Journal of Public Economics** 1993;50(2):149-167.

Brouwer WB, Koopmanschap MA, Rutten FF. Productivity costs measurement through quality of life? A response to the recommendations of the Washington Panel. **Health Economics** 1997;6(3):253-259.

Buckingham K. A note on HYE (Healthy Years Equivalent). **Journal of Health Economics** 1993;11:301-309.

Buxton MJ, Drummond MF, Van Hout BA, Prince RL, Sheldon TA, Szucs T, et al. Modelling in economic evaluation: an unavoidable fact of life. **Health Economics** 1997;6(3):217-227.

Cairns J. Discounting and health benefits: another perspective. **Health Economics** 1992;1(1): 76-79.

Canadian Coordinating Office for Health Technology Assessment (CCOHTA). **A guidance document for the costing process** Version 1.0. Ottawa: CCOHTA; 1996.

Carlin JB. Meta-analysis for 2 x 2 tables: a Bayesian approach. **Statistics in Medicine** 1992;11(2):141-158.

Chalmers I, Altman DG, editors. **Systematic reviews**. London: BMJ Publishing Group; 1995.

Chaudhary MA, Stearns SC. Estimating confidence intervals for cost-effectiveness ratios: an example from a randomized trial. **Statistics in Medicine** 1996;15:1447-1458.

Clinch JJ. The functional living index - cancer: Ten years later. In: Spilker B, editor. **Quality of life and pharmacoeconomics in clinical trials** 2nd ed. Philadelphia, PA: Lippincott-Raven; 1996. p. 215-225.

Cook DJ, Mulrow CD, Haynes RB. Systematic reviews: synthesis of best evidence for clinical decisions. **Annals of Internal Medicine** 1997;126(5):376-380.

Cooper H, Hedges LV, editors. **The handbook of research synthesis** New York, NY: Russell Sage Foundation; 1994.

Culyer AJ. The normative economics of health care finance and provision. **Oxford Review of Economic Policy** 1989;5(1):34-58.

Culyer AJ. Commodities, characteristics of commodities, characteristics of people, utilities, and the quality of life. In: Baldwin S, Godfrey C, Propper C, editors. **Quality of life: perspectives and policies.** London: Routledge; 1990. p. 9-27.

Culyer AJ. Health, health expenditures and equity. In: van Doorslaer E, Wagstaff A, editors. **Equity in the finance of and delivery of health care: an international perspective**. Oxford, UK: Oxford

University Press; 1993. p. 299-319.

Culyer AJ, Wagstaff A. QALYs versus HYEs. **Journal of Health Economics** 1993;11:311-323.

Damiano AM. The Sickness Impact Profile. In: Spilker B, editor. **Quality of life and pharmacoeconomics in clinical trials** 2nd ed. Philadelphia: Lippincott-Raven; 1996. p. 347-354.

Davis CE. Generalizing from clinical trials. **Controlled Clinical Trials** 1994;15:11-14.

Detsky AS. Guidelines for economic analysis of pharmaceutical products: a draft document for Ontario and Canada. **PharmacoEconomics** 1993;3 (5):354-361.

Detsky AS, Naglie G. A clinician's guide to cost-effectiveness analysis. **Annals of Internal Medicine** 1990;113(2): 147-154.

Diamond PA, Hausman JA. Contingent valuation: is some number better than no number? **Journal of Economic Perspectives** 1994;8(4):45-64.

Ditto PH, Druley JA, Moore KA, Danks JH, Smucker WD. Fates worse than death: the role of valued life activities in health-state evaluations. **Health Psychology** 1996;15(5):332-343.

Dolan P, Gudex C, Kind P, Williams A. **A social tariff for EuroQol: results from a UK general population survey**. York, UK: University of York, Centre for Health Economics; 1995. Discussion Paper no. 138.

Donaldson C. Willingness to pay for publicly-provided goods: a possible measure of benefit? **Journal of Health Economics** 1990;9:103-118.

Doubilet P, Begg CB, Weinstein MC, Braun P, McNeil BJ. Probabilistic sensitivity analysis using Monte Carlo simulation: a practical approach. **Medical Decision Making** 1985;5 (2):157-177.

Drummond MF, Bloom BS, Carrin G, Hillman AL, Hutchings HC, Knill-Jones RP, et al. Issues in the cross-national assessment of health technology. **International Journal of Technology Assessment in Health Care** 1992;8(4):671-682.

Drummond MF, Jefferson TO. Guidelines for authors and peer reviewers of economic submissions to the BMJ. **BMJ** 1996;313(7052):275-283.

Drummond M, Mason J, Torrance G. Cost-effectiveness league tables: think of the fans. **Health Policy** 1995;31:231-238.

Drummond MF, O'Brien BJ, Stoddart GL, Torrance GW. **Methods for the economic evaluation of health care programmes**. 2nd ed. Oxford: Oxford University Press; 1997.

Drummond M, Torrance G, Mason J. Cost-effectiveness league tables: more harm than good? **Social Sciences and Medicine** 1993;37(1):33-40.

Estaugh SR. Valuation of the benefits of risk-free blood: willingness to pay for hemoglobin solutions. **International Journal of Technology Assessment in Health Care** 1991;7(1):51-57.

Eddy DM, Hasselblad V, Shachter R. **Meta-analysis by the confidence profile method: the statistical synthesis of evidence**. Boston, MA: Academic Press; 1992.

Eisenberg JM, Schulman KA, Glick H, Koffer H. Pharmacoeconomics: economic evaluation of pharmaceuticals. In: Strom BL, editor. **Pharmacoepidemiology**. 2nd ed. Chichester, UK: John Wiley & Sons; 1994. p. 469-493.

Essink-Bot ML, Stouthard ME, Bonsel GJ. Generalizability of valuations on health states collected with the EuroQolc-questionnaire. **Health Economics** 1993;2 (3):237-246.

EuroQol Group. EuroQol: a new facility for the measurement of health-related quality of life **Health Policy** 1990;16(3):199-208.

Feeny D, Barr RD, Furlong W, Torrance GW, Weitzman S. Quality of life of the treatment process in pediatric oncology: an approach to measurement. In: Osoba D, editor. **Effect of cancer on quality of life**. Boca Raton, FL: CRC Press;1991. p. 73-88.

Feeny D, Furlong W, Boyle M, Torrance GW. Multi-attribute health status classification systems: Health Utilities Index. **PharmacoEconomics** 1995;7(6):490-502.

Feeny D, Guyatt G, Tugwell P. **Health care technology: effectiveness, efficiency and public policy**. Montreal, QC: The Institute for Research on Public Policy; 1986.

Feeny DH, Torrance GW, Furlong WJ. Health utilities index. In: Spilker B, editor. **Quality of life and pharmacoeconomics in clinical trials** 2nd ed. Philadelphia, PA: Lippincott-Raven; 1996. p. 239-252.

Feeny DH, Torrance GW, Goldsmith CH, Furlong W, Boyle M. A multi-attribute approach to population health status. In: **Proceedings of the Social Statistics Section of the American Statistical Association, 1993**. Session 1X; 1993. p. 161-166.

Feeny DH, Torrance GW, Labelle R. Integrating economic evaluations and quality of life assessments. In: Spilker B, editor. **Quality of life and pharmacoeconomics in clinical trials**, 2nd ed. Philadelphia: Lippincott-Raven; 1996. p. 85-95.

Fleiss JL. The statistical basis of meta-analysis. **Statistical Methods in Medical Research** 1993;2:121-145.

Fretwell MD. Frail older patients: creating standards of care. In: Spilker B. editor. **Quality of life and pharmacoeconomics in clinical trials** 2nd ed. Philadelphia, PA: Lippincott-Raven; 1996. p. 809-817.

Freund DA, Dittus RS. Principles of pharmacoeconomic analysis of drug therapy. **PharmacoEconomics** 1992;1(1): 20-32.

Fryback DG. QALYs, HYEs, and the loss of innocence. **Medical Decision Making** 1993;13(4):271-272.

Fuchs VR. Time preference and health: an exploratory study. In: Fuchs VR, editor. **Economic aspects of health**. Chicago, IL: University of Chicago Press; 1982. p. 93-120.

Furlong W, Feeny D, Torrance G, Barr R, Horsman J. **Guide to design and development of health-state utility instrumentation** Hamilton, ON: McMaster University, Centre for Health Economics and Policy Analysis; 1990. Working Paper no. 90-9.

Furlong W, Torrance GW, Feeny D, Boyle M. McMaster health utilities index of health-related quality of life [abstract]. **Quality of Life Research** 1994;3:76-77.

Gafni, A. Willingness-to-pay as a measure of benefits: relevant questions in the context of public decision-making about health care programs. **Medical Care** 1991;29(12):1246-1252.

Gafni A. Time in health: can we measure individuals 'Apure time preferences'? **Medical Decision Making** 1995;15(1): 31-37.

Gafni A. HYEs: do we need them and can they fulfil the promise? **Medical Decision Making** 1996;16(3):215-216.

Gafni A, Birch S. Equity considerations in utility-based measures of health outcomes in economic appraisals: an adjustment algorithm. **Journal of Health Economics** 1991;10(3):329-342.

Gafni A, Birch S, Mehrez A. Economics, health and health economics: HYEs versus QALYs. **Journal of Health Economics** 1993;11:325-339.

Gafni A, Torrance GW. Risk attitude and time preference in health. **Management Science** 1984;30(4):440-451.

Garber AM, Phelps CE. Economic foundations of cost-effectiveness analysis. **Journal of Health Economics** 1997;16(1):1-31.

Genduso LA, Kotsanos JG. Review of health economic guidelines in the form of regulations, principles, policies and positions. **Drug Information Journal** 1996;30:1003-1016.

Gibaldi M. Meta-analysis: a review of its place in therapeutic decision-making. **Drugs** 1993;46(5):805-818.

Gold MR, Patrick DL, Torrance GW, Fryback DG, Hadorn DC, Kamlet MS, et al. Identifying and valuing outcomes. In: Gold MR, Russell LB, Seigel JE, Weinstein MC, editors. **Cost-effectiveness in health and medicine**. New York, NY: Oxford University Press; 1996. p. 82-134.

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

Goldsmith CH, Gafni A, Drummond MF, Torrance GW, Stoddart GL. Sensitivity analysis and experimental design: the case of economic evaluation of health care programmes. In: Horne JM, editor. **Proceedings of the third Canadian Conference on Health Economics 1986** Winnipeg, MB: University of Manitoba, Department of Social and Preventive Medicine; 1987. p. 129-148.

Guyatt GH, Feeny DH, Patrick DL. Measuring health-related quality of life. **Annals of Internal Medicine** 1993;118(8):622-629.

Guyatt GH, Jaeschke R, Feeny DH, Patrick DL. Measurements in clinical trials: choosing the right approach. In: Spilker B, editor. **Quality of life and pharmacoeconomics in clinical trials** 2nd ed. Philadelphia, PA: Lippincott-Raven; 1996. p. 41-48.

Guyatt GH, Veldhuyzen Van Zanten SJ, Feeny DH, Patrick DL. Measuring quality of life in clinical trials: a taxonomy and review. **Canadian Medical Association Journal** 1989;140(12):1441-1448.

Hanemann WM. Valuing the environment through contingent valuation. **Journal of Economic Perspectives** 1994;8(4):19-43.

Harvey CM. The reasonableness of non-constant discounting. **Journal of Public Economics** 1994;53(1):31-51.

Hillman AL, Kim MS. Economic decision-making in healthcare: a standard approach to discounting health outcomes. **PharmacoEconomics** 1995;7(3):198-205.

Hunter JE, Schmidt FL. **Methods of meta-analysis: correcting error and bias in research findings**. Newbury Park, CA: Sage Publications; 1990.

Huston P, Hoey J. CMAJ endorses the CONSORT statement. **Canadian Medical Association Journal** 1996;155(9):1277-1279.

Ikeda S, Ikegami N, Oliver AJ, Ikeda M. A case for the adoption of pharmacoeconomic guidelines in Japan. **PharmacoEconomics** 1996;10(6):546-551.

Jacobs P, Bachynsky J, Baladi JF. A comparative review of pharmacoeconomic guidelines. **PharmacoEconomics** 1995;8(3):182-189.

Jenkins GP. The public-sector discount rate for Canada: some further observations. **Canadian Public**

Policy 1981;7(3):399-407.

Johannesson M. Economic evaluation of hypertension treatment. **International Journal of Technology Assessment in Health Care** 1992;8(3):506-523.

Johannesson M. Economic evaluation of lipid lowering - a feasibility test of the contingent valuation approach. **Health Policy** 1992;20(3):309-320.

Johannesson M. The contingent valuation controversy in environmental economics and its relevance to health services research. **Journal of Health Services Research and Policy** 1996;1(2):116-117.

Johannesson M. A note on the relationship between *ex ante* and expected willingness to pay for health care. **Social Science and Medicine** 1996;42(3):305-311.

Johannesson M, Johansson PO, Jonsson B. Economic evaluation of drug therapy: a review of the contingent valuation method. **PharmacoEconomics** 1992;1(5):325-337.

Johannesson M, Jonsson B. Economic evaluation in health care: is there a role for cost-benefit analysis? **Health Policy** 1991;17(1):1-23.

Johannesson M, Jonsson B, Borgquist L. Willingness to pay for antihypertensive therapy - results of a Swedish pilot study. **Journal of Health Economics** 1991;10: 461-473.

Johannesson M, Jonsson B, Karlsson G. Outcome measurement in economic evaluation. **Health Economics** 1996;5(4):279-296.

Johannesson M, Jonsson B, Kjekshus J, Olsson AG, Pedersen TR, Wedel H. Cost effectiveness of simvastatin treatment to lower cholesterol levels in patients with coronary heart disease. **New England Journal of Medicine** 1997;336(5):332-336.

Johannesson M, Pliskin JS, Weinstein MC. Are healthy-years equivalents an improvement over quality-adjusted life years? **Medical Decision Making** 1993;13(4):281-286.

Kaplan RM, Anderson JP. A general health policy model: update and applications **Health Services Research** 1988;23(2):203-235.

Kaplan RM, Anderson JP. The general health policy model: an integrated approach. In: Spilker B, editor. **Quality of life and pharmacoeconomics in clinical trials** 2nd ed. Philadelphia, PA: Lippincott-Raven; 1996. p. 309-322.

Katz DA, Welch HG. Discounting in cost-effectiveness analysis of healthcare programmes. **PharmacoEconomics** 1993;3(4):276-285.

Keeler EB, Cretin S. Discounting of life-saving and other nonmonetary effects.**Management Science** 1983;29(3):300-306.

Kind P. The EuroQol instrument: an index of health-related quality of life. In Spilker, B, editor. **Quality of life and pharmacoeconomics in clinical trials** 2nd ed. Philadelphia, PA: Lippincott-Raven; 1996. p. 191-201.

Koopmanschap MA, Rutten FF. Indirect costs in economic studies: confronting the confusion. **PharmacoEconomics** 1993;4(6):446-454.

Koopmanschap MA, Rutten FF. A practical guide for calculating indirect costs of disease. **PharmacoEconomics** 1996;10(5):460-466.

Koopmanschap MA, Rutten FF, van Ineveld BM, van Roijen L. The friction cost method for measuring indirect costs of disease.**Journal of Health Economics** 1995;14:171-189.

Koopmanschap MA, Turren FFH, van Ineveld BM, van Roijen L. The friction cost method for measuring indirect costs of disease. **Journal of Health Economics** 1995;14:171-189.

Koopmanschap MA, van Ineveld BM. Towards a new approach for estimating indirect costs of disease. **Social Science and Medicine** 1992;34 (9):1005-1010.

Krahn M, Gafni A. Discounting in the economic evaluation of health care interventions**Medical Care** 1993;31(5): 403-418.

Krahn MD, Naglie G, Naimark D, Redelmeier DA, Detsky AS. Primer on medical decision analysis: Part 4 - Analyzing the model and interpreting the results. **Medical Decision Making** 1997;17(2):142-151.

Kuppermann M, Shiboski S, Feeny D, Elkin EP, Washington AE. Can preference scores for discrete states be used to derive preference scores for an entire path of events? An application to prenatal diagnosis. **Medical Decision Making** 1997;17(1):42-55.

L'Abbé KA, Detsky AS, O'Rourke K. Meta-analysis in clinical research. **Annals of Internal Medicine** 1987;107(2):224-233.

Labelle RJ, Hurley JE. Implications of basing health-care resource allocations on cost-utility analysis in the presence of externalities.**Journal of Health Economics** 1992;11(3):259-277.

Laska EM, Meisner M, Seigel C. Statistical inference for cost-effectiveness ratios. **Health Economics** 1997;6(3):229-242.

Laupacis A, Feeny D, Detsky A, Tugwell PX. How attractive does a new technology have to be to warrant adoption and utilization? Tentative guidelines for using clinical and economic evaluation. **Canadian Medical Association Journal**1992;146(4):473-481.
Lipscomb J. Time preference for health in cost-effectiveness analysis.**Medical Care** 1989;27(3

Suppl):S233-S253.

Lipscomb J, Weinstein MC, and Torrance GW. Time preference. In: Gold MR, Seigel JE, Russell LB and Weinstein MC, editors. **Cost-effectiveness in health and medicine** New York: Oxford University Press; 1996. p. 214-246.

Llewellyn-Thomas HA, Sutherland HJ, Thiel EC. Do patients' evaluations of a future health state change when they actually enter that state?**Medical Care** 1993;31(11):1002-1012.

Loomes G. The myth of the HYE.**Journal of Health Economics** 1995;14(1): 1-7.

Loomes G, McKenzie L. The use of QALYs in health care decision-making.**Social Science and Medicine** 1989;28 (4):299-308.

Luce BR, Manning WG, Seigel KE and Lipscomb J. Estimating costs in cost-effectiveness analysis. In: Gold MR, Seigel JE, Russell LB and Weinstein MC , editors. **Cost-effectiveness in health and medicine**. New York: Oxford University Press; 1996. p. 176-213.

MacKeigan LD, Larson LN, Draugalis JR, Bootman JL, Burns LR. Time preference for health gains versus health losses. **PharmacoEconomics** 1993;3(5):374-386.

Mandelblatt JS, Fryback DG, Weinstein MC,Russell LB, Gold MR, Hadorn DC. Assessing the effectiveness of health interventions. In: Gold MR, Seigel JE, Russell LB and Weinstein MC, editors. **Cost-effectiveness in health and medicine** New York: Oxford University Press; 1996. p. 135-175.

Mangione CM, Lee PP, Hayse RD. Measurement of visual functioning and health-related quality of life in eye disease and cataract surgery. In: Spilker B, editor. **Quality of life and pharmacoeconomics in clinical trials** 2nd ed. Philadelphia, PA: Lippincott-Raven; 1996. p. 1045-1051.

Mangione CM, Phillips RS, Seddon JM, Lawrence MG, Cook EF, Dailey R, et al. Development of the 'Activities of Daily Vision Scale'. A measure of visual functional status.**Medical Care** 1992;30(12):1111-1126.

Manning WG, Fryback DG, Weinstein MC. Reflecting uncertainty in cost-effectiveness analysis. In: Gold MR, Seigel JE, Russell LB and Weinstein MC, editors. **Cost-effectiveness in health and medicine**. New York: Oxford University Press; 1996. p. 247-275.

Mason J. The generalisability of pharmacoeconomic studies. **PharmacoEconomics** 1997;11(6):503-514.

Mason J, Drummond M, Torrance G. Some guidelines on the use of cost effectiveness league tables. **BMJ** 1993;306(6877):570-572.

McDowell I, Newell C. **Measuring health: A guide to rating scales and questionnaires, 2nd ed**.

New York: Oxford University Press; 1996.

McEwen J, McKenna SP. Nottingham Health Profile. In: Spilker B, editor. **Quality of life and pharmacoeconomics in clinical trials** 2nd ed. Philadelphia, PA: Lippincott-Raven; 1996. p. 281-286.

McIntosh MW. The population risk as an explanatory variable in research synthesis of clinical trials. **Statistics in Medicine** 1996;15(16):1713-1728.

Mehrez A, Gafni A. Quality-adjusted life years, utility theory, and healthy-years equivalents. **Medical Decision Making** 1989;9(2):142-149.

Mehrez A, Gafni A. The healthy-years equivalents: how to measure them using the standard gamble approach. **Medical Decision Making** 1991;11(2):140-146.

Mehrez A, Gafni A. Preference based outcome measures for economic evaluation of drug interventions: quality adjusted life years (QALYs) versus healthy years equivalents (HYEs). **PharmacoEconomics** 1992;1(5):338-345.

Mehrez A, Gafni A. Healthy-years equivalents versus quality-adjusted life years: in pursuit of progress. **Medical Decision Making** 1993;13(4):287-292.

Meltzer D. Accounting for future costs in medical cost-effectiveness analysis. **Journal of Health Economics** 1997;16(1):33-64.

Miller GA. The magical number seven, plus or minus two: some limits on our capacity for processing information. **Psychological Review** 1956;63(2):81-97.

Mishan EJ. **Cost -Benefit Analysis**. London, UK: Allen and Unwin; 1971.

Mitchell RC, Carson RT. **Using surveys to value public goods: the contingent valuation method.** Washington, DC: Resources for the Future; 1989.

Mullahy J. What you don't know can't hurt you? Statistical issues and standards for medical technology evaluation. **Medical Care** 1996;34(12 Suppl):DS124-DS135.

Mullahy J, Manning W. Statistical issues in cost-effectiveness analysis. In: Sloan FA, editor. **Valuing health care: costs, benefits and effectiveness of pharmaceuticals and other medical technologies**. New York, NY: Cambridge University Press; 1995. p. 149-184.

Murray CJL. Quantifying the burden of disease: the technical basis for disability-adjusted life years. **Bulletin of the World Health Organization** 1994;72(3):429-445.

Murray CJ, Lopez AD, editors. **The global burden of disease: a comprehensive assessment of mortality and disability from diseases, injuries, and risk factors in 1990 and projected to 2020** Cambridge: Harvard University Press; 1996.

Neumann PJ, Johannesson M. The willingness to pay for in vitro fertilization: a pilot study using contingent valuation. **Medical Care** 1994;32(7):686-699.

Nord E. The person-trade-off approach to valuing health care programs. **Medical Decision Making** 1995;15(3):201-208.

Nord E. Health status index models for use in resource allocation decisions: a critical review in the light of observed preferences for social choice. **International Journal of Technology Assessment in Health Care** 1996;12(1):31-44.

Nord E, Richardson J, Macarounas-Kirchmann K. Social evaluation of health care versus personal evaluation of health states: evidence on the validity of four health-state scaling instruments using Norwegian and Australian surveys. **International Journal of Technology Assessment in Health Care** 1993;9(4):463-478.

O'Brien BJ. A tale of two (or more) cities: geographic transferability of pharmacoeconomic data. **American Journal of Managed Care** In press 1997.

O'Brien BJ, Drummond MF. Statistical versus quantitative significance in the socioeconomic evaluation of medicines. **PharmacoEconomics** 1994;5(5):389-398.

O'Brien BJ, Drummond MF, Labelle RJ, Willan A. In search of power and significance: issues in the design and analysis of stochiastic cost-effectiveness studies in health care. **Medical Care** 1994;32(2):150-163.

O'Brien B, Gafni A. When do the "dollars" make sense? Toward a conceptual framework for contingent valuation studies in health care. **Medical Decision Making** 1996;16(3):288-299.

O'Brien BJ, Goeree R, Gafni A, et al. Assessing the value of a new pharmaceutical: a feasibility study of contingent valuation in managed care. **Medical Care**. In press 1997.

O'Brien BJ, Heyland D, Richardson WS, Levine M, Drummond MF. Users' Guides to the Medical Literature: XIII. How to use an article on economic analysis of clinical practice. Part B: What are the results and will they help me in caring for my patients? **JAMA** 1997:277(22):1802-1806.

O'Brien B, Viramontes JL. **Willingness to pay: a valid and reliable measure of health state preference?** Hamilton, ON: McMaster University, Centre for Health Economics and Policy Analysis; 1993. CHEPA Working Paper no. 93-7.

Oliver RM, Smith JQ, editors. **Influence diagrams, belief nets, and decision analysis** New York, NY: Wiley; 1990.

Olsen JA. On what basis should health be discounted? **Journal of Health Economics** 1993;12:39-53.

Ontario Ministry of Health. **Ontario guidelines for economic analysis of pharmaceutical products**. Toronto: Ontario Ministry of Health, Drug Programs Branch; 1994.

Oxman AD. Meta-statistics: help or hindrance? **ACP Journal Club** 1993; (May-June):A13-A15.

Oxman AD. Preparing and maintaining systematic reviews. In: Sackett D, Oxman A, editors. **The Cochrane Collaboration handbook** Oxford, UK: The Cochrane Collaboration; 1996.

Oxman AD, Guyatt GH. A consumer's guide to subgroup analyses. **Annals of Internal Medicine** 1992;116 (1):78-84.

Parsonage M, Neuburger H. Discounting and health benefits. **Health Economics** 1992;1:71-79.

Patrick DL, Erickson P. **Health status and health policy: quality of life in health care evaluation and resource allocation**. New York, NY: Oxford University Press; 1993.

Patrick DL, Starks HE, Cain KC, Uhlmann RF, Pearlman RA. Measuring preferences for health states worse than death. **Medical Decision Making** 1994;14(1):9-18.

Petitti D. **Meta-analysis, decision analysis, and cost-effectiveness analysis: methods for quantitative synthesis in medicine** New York, NY: Oxford University Press; 1994.

Polsky D, Glick HA, Willke R, Schulman K. Confidence intervals for cost-effectiveness ratios: a comparison of four methods. **Health Economics** 1997;6(3):243-252.

Portney PR. The contingent valuation debate: why economists should care. **Journal of Economic Perspectives** 1994;8(4):3-17.

Richardson J. Cost-utility analysis: what should be measured? **Social Science and Medicine** 1994;39(1):7-21.

Richardson S. **Cost-utility analyses: what should be measured: utility, value or healthy years equivalents**. Fairfield, Victoria: Monash University, Centre for Health Program Evaluation; 1990. Working Paper no. 5.

Rittenhouse BE. Healthy years equivalents versus time trade-off: ambiguity on certainty and uncertainty. **International Journal of Technology Assessment in Health Care** 1997;13(1):35-48.

Rosenthal R. **Meta analytic procedures for social research** rev. ed. Newbury Park, CA: Sage Publications; 1991.

Sackett DL, Torrance GW. The utility of different health states as perceived by the general public. **Journal of Chronic Diseases** 1978;31(11):697-704.

Sacristan JA, Day SJ, Navarro O, Ramos J, Hernandez JM. Use of confidence intervals and sample

size calculations in health economic studies. **Annals of Pharmacotherapy** 1995;29(7-8):719-725.

Schubert F, editor. **Canadian collaborative workshop on pharmacoeconomics (1993 : Saint Adèle, Quebec, Canada). Proceedings.** Princeton, NJ: Excerpta Medica; 1993.

Seigel JE, Weinstein MC, Torrance GW. Reporting cost-effectiveness studies and results. In: Gold MR, Seigel JE, Russell LB and Weinstein MC, editors. **Cost-effectiveness in health and medicine** New York: Oxford University Press; 1996. P. 276-303.

Selvin S. **Statistical analysis of epidemiologic data** New York, NY: Oxford University Press; 1991.

Sen A. Utility: ideas and terminology. **Economics and Philosophy** 1991;7 (2):277-283.

Sorofman BA, Milavetz G. Meta-analysis and the scientific research that precedes it [editorial]. **Annals of Pharmacotherapy** 1994;28:1290-1291.

Speigelhalter DJ, Freedman LS, Parmar MK. Bayesian approaches to randomized trials. **Journal of the Royal Statistical Society A** 1994;157(Part 3):357-416.

Spilker B, editor. **Quality of life and pharmacoeconomics in clinical trials** 2nd ed. Philadelphia, PA: Lippincott-Raven; 1996.

Su XY, Li Wan Po A. Combining event rates from clinical trials: comparison of Bayesian and classical methods. **Annals of Pharmacotherapy** 1996;30(5):460-465.

Thompson MS. Willingness to pay and accept risks to cure chronic disease **American Journal of Public Health** 1986;76(4):392-396.

Torrance GW. Health states worse than death. In: v.Eimeren W, Engelbrecht R, Flagle CD, editors. **Proceedings of the Third International Conference on Systems Science in Health Care** Berlin: Springer-Verlag; 1984. p . 1085-1089.

Torrance GW. Measurement of health-state utilities for economic appraisal: a review **Journal of Health Economics** 1986;5:1-30.

Torrance GW. Designing and conducting cost-utility analyses. In: Spilker B, editor **Quality of life and pharmacoeconomics in clinical trials** 2nd ed. Philadelphia: Lippincott-Raven; 1996. p. 1105-1111.

Torrance GW, Feeny D. Utilities and quality-adjusted life years. **International Journal of Technology Assessment in Health Care** 1989;5(4):559-575.

Torrance GW, Feeney DH, Furlong WJ, Barr RD, Zhang Y, Wang Q. Multiattribute utility function for a comprehensive health status classification system. Health Utilities Index Mark 2. **Medical**

Care 1996;34(7):702-722.

Torrance GW, Furlong WJ, Feeny DH, Boyle M. Multi-attribute preference functions: Health Utilities Index. **PharmacoEconomics** 1995;7(6):503-520.

Torrance GW, Seigel JE, Luce BR. Framing and designing the cost-effectiveness analysis. In: Gold MR, Seigel JE, Russell LB and Weinstein MC, editors. **Cost-effectiveness in health and medicine** New York: Oxford University Press; 1996. P. 54-81

Tugwell P, Bennett KJ, Sackett DL, Haynes RB. The measurement iterative loop: a framework for the critical appraisal of need, benefits and costs of health interventions.**Journal of Chronic Diseases** 1985;38(4):339-351.

Tweedie RL, Scott DJ, Biggerstaff BJ, Mengersen KL. Bayesian meta-analysis, with application to studies of ETS and lung cancer. **Lung Cancer** 1996;14 (Suppl 1):S171-S194.

Ubel PA, DeKay ML, Baron J, Asch DA. Cost-effectiveness analysis in a setting of budget constraints: is it equitable? **New England Journal of Medicine** 1996;334(18):1174-1177.

Ubel PA, Loewenstein F, Scanlon D, Kamlet M. Individual utilities are inconsistent with rationing choices: a partial explanation of why Oregon's cost-effectiveness list failed.**Medical Decision Making** 1996;16 (2):108-116.

United States General Accounting Office. **Cross design synthesis: a new strategy for medical effectiveness research**. Washington,DC: General Accounting Office(US); March 1992. Report No. GAO/PEMD-92-18.

Velanovich V. Meta-analysis for combining Bayesian probabilities. **Medical Hypotheses** 1991;35(3):192-195.

Von Neumann J, Morgenstern O. **Theory of games and economic behaviour**. Princeton, NJ: Princeton University Press; 1944.

Wakker P. A criticism of healthy-years equivalents.**Medical Decision Making** 1996;16(3):207-214.

Ware JE. The SF-36 health survey. In: Spilker B, editor.**Quality of life and pharmacoeconomics in clinical trials**. 2nd ed. Philadelphia, PA: Lippincott-Raven; 1996. p. 337-345.

Weinstein MC. Principles of cost-effective resource allocation in health care organizations. **International Journal of Technology Assessment in Health Care** 1990;6 (1):93-103.

Weinstein MC, Manning WG. Theoretical issues in cost-effectiveness analysis [editorial].**Journal of Health Economics** 1997;16(1):121-128.
Weinstein MC, Pliskin JS. Perspectives on healthy-years equivalents: HYEs: what are the issues? **Medical Decision Making** 1996;16(3):205-206.

Weinstein MC, Stason WB. Foundations of cost-effectiveness analysis for health and medical practices. **New England Journal of Medicine** 1977;296(13):716-721.

Willan AR, O'Brien BJ. Confidence intervals for cost-effectiveness ratios: an application of Fieller's theorem. **Health Economics** 1996;5(4):297-305.

Williams A. Ethics and efficiency in the provision of health care. In: Bell JM, Mendus S, editors. **Philosophy and medical welfare** Cambridge: Cambridge University Press; 1988. p. 111-126.

Williams A. Economics, QALYs, and medical ethics: a health economist's perspective **Health Care Analysis** 1995;3 (3):221-226.

Williams A. QALYs and ethics: a health economist's perspective. **Social Science and Medicine** 1996;43(12):1795-1804.

Williams A. Intergenerational equity: an exploration of the "fair innings" argument. **Health Economics** 1997;6(2):117-132.

Winston WL. **Operations research: applications and algorithms** Boston, MA: PWS-Kent Publishing; 1991.

Yusuf S, Wittes J, Probstfield J, Tyroler HA. Analysis and interpretation of treatment effects in subgroups of patients in randomized clinical trials. **JAMA** 1991;266(1):93-98.

VII. INDEX

Aggregate Impact Analysis 37, 60
Consequences 7, 11-13, 26, 39, 41, 58, 59, 62
Contingent Valuation 3, 12, 31-34, 62
Cost Identification 4, 39
Cost Measurement 4, 43, 59, 62
Cost of Lost Time 40, 59, 62, 64
Cost Valuation 4, 44
Cost(s) 1, 4, 7, 9, 11-18, 28, 31, 36, 38-44, 46-48, 50-53, 56, 58-60, 62-66
Decision Analysis 30, 52, 63, 64
Decision Tree 22, 39, 49, 59, 63
Direct Cost(s) 59
Disability-adjusted Life Years 2, 25, 27
Disclosure of Relationships 5, 55, 57
Discount Rate 4, 26, 38, 45-46, 63
Discounting 4, 38, 45, 59
Distributional Considerations 60
Dominance 17, 63
Effectiveness 1, 2, 5, 7-9, 13, 14, 17, 20, 21, 23, 24, 46-48, 50-53, 58-60, 63
Efficacy 2, 5, 8, 12, 20, 21, 51-53, 58, 63
Equity 3, 7, 36, 37, 60, 63
Financial Impact Analysis 2, 18
Friction Cost Method 31, 41, 42, 64
Future Health Care Costs 13, 40, 42, 64
Healthy Years Equivalent, HYE 2, 25, 27, 30, 64
Horizon 2, 18, 58
HRQOL Instruments 2, 23, 24, 28, 30, 58
Human Capital Approach, HCA 3, 31, 32, 42, 64
Incremental Analysis 11, 57
Incremental Cost(s) 12, 17, 31, 48, 60, 64
Indication(s) 1, 14, 57, 60
Indirect Cost(s) 13, 40, 42, 59, 64
Management of Studies 1, 10, 67
Meta-analysis 21, 46, 58, 64
Models, Modelling 2, 14, 18-22, 24, 46, 49, 50, 52, 57, 59, 64, 65, 67
Monte Carlo Simulation 5, 46-49
Net Benefit 11, 62, 64
Ontario Guidelines 56
Opportunity Cost(s) 4, 7, 41, 44, 65
Outcome(s) 2-5, 7, 11-13, 15, 18, 20, 23-26, 28, 30-32, 34, 35-39, 42, 45, 48, 50, 52, 58, 59, 62, 63, 65, 66
Perspective 2, 18, 19, 25-27, 30, 32-34, 37, 39-41, 46, 50, 57, 62
Person Trade-off, PTO 27, 30
Probability Tree 4, 5, 39, 50
Reporting Results 5, 50, 56

Research Design....11, 33, 34, 57, 67
Results....2-5, 10-12, 17, 25, 36, 38, 44, 46-48, 50-54, 56-61, 67
Role of Studies....7-9
Sampling Error....5, 46, 47, 59
SavedYoung Life Equivalents, SAVE....2, 25, 27
Sensitivity Analysis....2, 4, 15, 18-20, 22, 25, 28, 38-40, 43-51, 53, 58-60, 65
Standard Gamble....24, 27-29, 65
Subgroup Analysis....1, 14-16, 22, 57-60, 65
Target Audience....1, 7, 8, 17, 18, 57
Target Population....1, 14, 19, 36
Time Trade-off....27-29, 65
Timing of Studies....1, 9
Total Analysis....1, 11
Transfer Payment(s)....9, 39, 66
Uncertainty....5, 8, 15, 19, 43, 46-51, 59, 63, 66
Utility....13, 23-25, 28, 30, 31, 36, 40, 51, 58, 62, 65-67
Viewpoint(s)....2-4, 18, 34, 35, 39, 42, 53, 57, 62
Willingness to Pay....3, 12, 31-33, 62, 66