



**Australian Government**

---

**Department of Health**

**Technical Guidelines for preparing  
assessment reports for the Medical  
Services Advisory Committee –  
Medical Service Type: Therapeutic  
(Version 2.0)**

March 2016

Medical Services Advisory Committee



# Foreword

---

The Medical Services Advisory Committee (MSAC) is an independent committee that provides advice to the Minister for Health on the strength of the evidence relating to the comparative safety, clinical effectiveness and cost-effectiveness of any new or existing medical services or technology, and the circumstances under which public funding should be supported through listing on the Medical Benefits Schedule (MBS).

To achieve this, MSAC undertakes Health Technology Assessments (HTA) using the best available evidence to assess proposals for their comparative safety, clinical effectiveness, and cost effectiveness.

Applications for therapeutic services provide applicants with a number of challenges, requiring them to prove that both the proposed service provides accurate, meaningful information and also that the information improves the subsequent treatment (and health outcomes) of patients.

This document provides detailed advice to assist applicants with determining content and presentation of submissions of evidence for consideration by MSAC and the Evaluation Sub-committee (ESC).

This document provides detailed advice to assist applicants with determining content and presentation of submissions of evidence for consideration by the MSAC and the ESC.

Chair  
Medical Services Advisory Committee

# Contents

---

<b>Foreword</b>	.....	<b>i</b>
<b>Contacts</b>	.....	<b>xi</b>
<b>Record of updates</b>	.....	<b>xii</b>
<b>Abbreviations</b>	.....	<b>xiii</b>
<b>PART I General information</b>	.....	<b>1</b>
<b>1 Medical Services Advisory Committee</b>	.....	<b>2</b>
1.1 Purpose and roles of MSAC	.....	2
1.2 Membership of MSAC	.....	2
1.3 MSAC sub-committees	.....	2
1.4 Overview of MSAC processes	.....	3
1.4.1 Regulatory framework	.....	3
1.4.2 The application and assessment process	.....	3
1.4.3 Sources of advice	.....	3
1.4.4 Publication of assessment report	.....	3
1.4.5 Timing of MSAC processes	.....	4
1.4.6 MSAC appraisal	.....	4
<b>2 Introduction to the Guidelines</b>	.....	<b>5</b>
2.1 Structure of these Guidelines	.....	5
2.2 Associated documents	.....	5
2.3 What is a therapeutic medical service?	.....	6
2.4 Writing and style conventions used in these Guidelines	.....	6
2.5 The future	.....	7
<b>3 Rationale and basis for the economic evaluation in the Australian context</b>	.....	<b>8</b>
3.1 Analysis of cost-effectiveness	.....	8
3.2 Australian context	.....	8
3.3 Relevant factors influencing MSAC decision making	.....	8
3.4 Flexibility in interpretation of these Guidelines	.....	9

<b>4</b>	<b>Organisation of a standard assessment report .....</b>	<b>10</b>
4.1	Choice of information .....	10
4.2	Overview of a standard assessment report .....	11
4.2.1	Sections of a standard assessment report .....	11
4.3	Presentation of the assessment report .....	11
4.3.1	Standard assessment report .....	11
4.3.2	Two stage approach to an assessment .....	13
4.3.3	Co-dependent/integrated assessment reports .....	13
<b>5</b>	<b>Lodging an assessment report .....</b>	<b>14</b>
5.1	Assessment report checklist .....	14
5.1.1	Information requirements .....	14
5.1.2	Information requests .....	14
5.2	Provision of information to allow independent verification of computer analyses .....	16
5.2.1	Information requirements .....	16
5.3	Provision of information after lodgement of the assessment report .....	16
5.3.1	Post-lodgement communication with MSAC .....	16
5.3.2	Provision of information sourced from the TGA after lodgement of the assessment report .....	16
<b>Part II</b>	<b>Preferred clinical and economic evidence for proposed medical services to be considered by MSAC.....</b>	<b>17</b>
<b>Section A</b>	<b>Details of the proposed therapeutic medical service and its intended use on the Medical Benefits Schedule (or for other public funding) .....</b>	<b>18</b>
A1	Address all items in the agreed PICO Confirmation.....	18
A2	Proposed medical service .....	18
A3	Proposed MBS listing or other public funding sought.....	18
A4	Proposed population.....	18
A5	Comparator details .....	18
A6	Clinical management algorithm(s).....	19
A6.1	Algorithms for intended and current contexts.....	19
A7	Differences between the proposed medical service and the main comparator .....	20
A8	Clinical claim .....	20
A9	Summarise the primary elements of the decision analysis (PICO).....	20

<b>Section B</b>	<b>Clinical evaluation for the proposed therapeutic medical service.....</b>	<b>21</b>
B1	Description of search strategies .....	23
B2	Listing all direct randomised trials.....	24
B3	Assessment of the measures taken by investigators to minimise bias .....	30
B4	Characteristics of the trials.....	33
B5	Outcome measures and analysis.....	34
B6	Systematic overview of the results.....	38
B7	Extended assessment of comparative harms .....	46
B8	Interpretation of the clinical evidence .....	46
<b>Section C</b>	<b>Translation Issues .....</b>	<b>49</b>
C1	Identification of issues to be addressed.....	51
C2	Focused analytical plan .....	56
C3	Results of pre-modelling studies .....	68
C4	Relationship of each pre-modelling study to the economic evaluation .....	69
<b>Section D</b>	<b>Economic evaluation for the main indication .....</b>	<b>71</b>
D.1	Overview of the economic evaluation.....	73
D2	Population and circumstances of use reflected in the economic evaluation .....	80
D3	Structure and rationale of the economic evaluation .....	85
D4	Variables in the economic evaluation .....	93
D5	Results of the economic evaluation.....	100
D6	Sensitivity analyses .....	107
<b>Section E</b>	<b>Estimated utilisation and financial implications .....</b>	<b>110</b>
E1	Justification of the selection of sources of data .....	112
E2	Estimation of use and costs of the proposed medical service .....	114
E3	Estimation of changes in use and cost of other medical services.....	117
E4	Estimated financial implications for the MBS .....	118
E5	Identification, estimation and reduction of uncertainty .....	119

<b>Section F</b>	<b>Options to present additional relevant information .....</b>	<b>122</b>
F1	Other relevant factors .....	122
<b>Part III</b>	<b>Alternative clinical evidence for proposed medical services to be considered by MSAC .....</b>	<b>125</b>
<b>Section B(i)</b>	<b>Clinical evaluation for the main indication: indirect comparison of randomised trials .....</b>	<b>126</b>
B(i)1	Description of search strategies .....	128
B(i)2	Listing all randomised trials considered for inclusion in indirect comparisons .....	128
B(i)3	Assessment of the measures taken by investigators to minimise bias .....	132
B(i)4	Characteristics of the trials.....	132
B(i)5	Outcome measures .....	133
B(i)6	Results.....	133
B(i)7	Extended assessment of comparative harms .....	134
B(i)8	Interpretation of the clinical evidence.....	135
<b>Section B(ii)</b>	<b>Clinical evaluation for the main indication: Presenting non-randomised studies.....</b>	<b>136</b>
B(ii)1	Description of search strategies .....	138
B(ii)2	Listing all non-randomised studies .....	138
B(ii)3	Assessment of the measures taken by investigators to minimise bias .....	140
B(ii)4	Characteristics of the studies.....	142
B(ii)5	Outcome measures .....	142
B(ii)6	Results of the comparison involving non-randomised studies.....	142
B(ii)7	Extended assessment of comparative harms .....	142
B(ii)8	Interpretation of the clinical evidence.....	143
<b>Section C(i)</b>	<b>Translating the clinical evaluation: indirect comparisons of randomised trials or non-randomised studies .....</b>	<b>144</b>
C(i)1	Relevant pre-modelling studies.....	145

<b>Section D(i)</b>	<b>Economic evaluation for the main indication: presenting a cost-minimisation approach .....</b>	<b>147</b>
D(i)1	Cost-minimisation analysis .....	147
D(i)2	Presentation of a cost-minimisation analysis or a cost analysis .....	149
<b>Appendices</b>	<b>.....</b>	<b>151</b>
<b>Appendix 1</b>	<b>Relevant factors influencing provision of advice by MSAC.....</b>	<b>152</b>
<b>Appendix 2</b>	<b>Expert opinion .....</b>	<b>154</b>
A2.1	Uses of expert opinion .....	154
A2.2	Presenting expert opinion.....	155
A2.3	Describing the collection and collation of expert opinion .....	155
<b>Appendix 3</b>	<b>Assessment of non-inferiority.....</b>	<b>158</b>
A3.1	Introduction.....	158
A3.2	Service delivery information.....	158
A3.3	Non-inferiority threshold .....	158
A3.4	Method of analysis .....	159
A3.5	Presenting an assessment of non-inferiority .....	159
A3.6	Assessing comparative harms in the context of non-inferiority.....	160
A3.7	Interpretation of the clinical evidence.....	160
<b>Appendix 4</b>	<b>Utility valuation of health outcomes.....</b>	<b>161</b>
A4.1	Use of health-related QALYs gained and cost-utility analysis .....	161
A4.2	Obtaining utility weights.....	162
A4.3	Trial-based utility valuation of health outcomes.....	163
A4.4	Scenario-based utility valuation of health outcomes .....	165
A4.5	Other methods for obtaining utilities .....	172
<b>Appendix 5</b>	<b>Monetary valuation of health outcomes.....</b>	<b>175</b>
A5.1	Preference for cost-utility analyses over cost-benefit analyses.....	175
A5.2	Scenario-based monetary valuation of health outcomes.....	176



<b>Appendix 6</b>	<b>Including non-health care resources and non-health outcomes in a supplementary analysis .....</b>	<b>181</b>
A6.1	Identifying, measuring and valuing non-health care resources .....	181
A6.2	Identifying, measuring and valuing non-health outcomes .....	181
A6.3	Resources and outcomes to be excluded .....	184
<b>Appendix 7</b>	<b>Measures taken by the investigators to minimise bias in non-randomised studies.....</b>	<b>185</b>
A7.1	Classical observational designs.....	185
A7.2	Quasi-experimental designs .....	187

## Tables and Figures

Figure 4.1	Sections of a standard assessment report .....	12
Table 5.1	Checklist of information to be included in a therapeutic assessment report .....	15
Figure B1	Key information requests for assessment report Section B of an assessment for MSAC .....	22
Table B2.1	Summary of identification of direct randomised trials from the search of the published literature .....	25
Table B2.2	Summary of identification of applicants' <sup>a</sup> direct randomised trials and information from the manual search of retrieved citations .....	26
Table B2.3	Trials (and associated reports) presented in the assessment report .....	27
Table B2.4	Reasons to exclude each trial from further detailed assessment .....	28
Table B2.5	Comparative summary of characteristics of each direct randomised trial .....	28
Table B2.6	Comparative summary of results of each direct randomised trial .....	29
Box B3.1	Checklist for assessing the quality (internal validity) of randomised trials .....	30
Table B3.1	Summary of the measures undertaken to minimise bias in the direct randomised trials .....	32
Table B3.2	Flow of participants through the direct randomised trials .....	32
Table B4.1	Eligibility criteria .....	33
Table B4.2	Characteristics of participants across randomised groups .....	33
Table B4.3	Therapeutic medical services compared in trials .....	34
Table B5.1	Primary outcomes and statistical analyses of the direct randomised trials .....	35
Table B5.2	Patient-relevant secondary outcomes and analyses .....	36
Table B6.1	Results of [patient-relevant outcome] (available as dichotomous data) across the direct randomised trials (relative risk) .....	41
Table B6.2	Results of [patient-relevant outcome] (available as dichotomous data) across the direct randomised trials (risk difference) .....	41
Table B6.3	Results of [patient-relevant outcome] (available as continuous data) across the direct randomised trials (end point) .....	42
Table B6.4	Results of [patient-relevant outcome] (available as continuous data) across the direct randomised trials (change) .....	42
Table B6.5	Results of [patient-relevant outcome] across the direct randomised trials (available as time-to-event data) .....	44
Table B8.1	Classification of the comparative effectiveness and safety of the proposed therapeutic medical service compared with its main comparator and guide to the suitable type of economic evaluation .....	47
Table C2.1	Assessment of treatment effect variation across subgroups .....	60
Table C2.2	Assessment of relative treatment effect variation across subgroups .....	61

Table C2.3	Assessment of absolute treatment effect variation across subgroups .....	62
Table C2.4	Summary of assessment of treatment effect variation across subgroups .....	63
Table C2.5	Reasons to exclude each direct randomised trial .....	67
Table C4.1	Summary of results of pre-modelling studies and their uses in the economic evaluation .....	70
Figure D1	Key information requests for assessment report Section D of a standard assessment for MSAC .....	72
Table D1.1	Classification of a service under MSAC consideration/Classification of the effectiveness of the proposed medical service over its main comparator and guide to the suitable type of economic evaluation .....	75
Table D1.2	Type of economic evaluation that should be presented for various classifications of a service under MSAC consideration.....	77
Table D2.1	Definitions for populations and circumstance of use that should be taken into account in the evaluation.....	80
Table D2.2	Comparison of characteristics of trial and requested populations and circumstances of use.....	82
Table D2.3	Example of a comparison of the characteristics of target, study and wider populations and settings.....	84
Figure D3.1	Example of the structure of a basic economic evaluation.....	87
Table D3.1	Key elements of the base-case economic evaluation .....	89
Table D4.1	List of health care resource items and unit costs included in the economic evaluation .....	97
Table D5.1	List of health care resource items and summary of cost impacts in the economic evaluation.....	102
Table D5.2	List of health care resource items and summary of cost impacts for each health state in a state transition model.....	102
Table D5.3	List of health states and summary of cost impacts included in the economic evaluation .....	102
Table D5.4	List of health states and summary of health outcomes included in the economic evaluation.....	103
Table D5.5	Assessment of the implications for the economic evaluation of applying the clinical evaluation (Step 1 then Step 2).....	106
Table D5.6	Assessment of the implications for the economic evaluation of extrapolating and transforming the clinical evaluation (Step 3) .....	107
Table D6.1	Analyses of the implications for the economic evaluation of usage beyond the requested populations and circumstances of use .....	109
Figure E1	Key information requests for assessment report Section E of a standard assessment for MSAC .....	111
Table E1.1	Categories of data sources.....	113

Figure B(i)1	Key information requests for assessment report Section B of a standard assessment for MSAC with clinical data from an indirect comparison of randomised trials .....	127
Table B(i)2.1	Trials (and associated reports) presented in the assessment report .....	130
Table B(i)2.2	Reasons to exclude each trial from the indirect comparison.....	131
Table B(i)2.3	Summary of randomised trials used to conduct the indirect comparison .....	132
Table B(i)6.1	Summary of results of the indirect comparison of randomised trials .....	134
Figure B(ii)1	Key information requests for assessment report Section B of a standard assessment for MSAC with clinical data from non-randomised studies .....	137
Table B(ii)2.1	Studies (and associated reports) presented in the assessment report .....	139
Table B(ii)2.2	Reasons to exclude each study .....	140
Figure C(i)1	Key information requests for assessment report Section C of a standard assessment for MSAC with clinical data from an indirect comparison of randomised trials (Section B(i)) or from non-randomised studies (Section B(ii)).....	144
Figure D(i)1	Key information requests for assessment report Section D of a standard assessment for MSAC in which the therapeutic conclusion from assessment report Section B is non-inferior .....	148
Table A1.1	Factors that are more readily quantified.....	152
Table A1.2	Examples of factors that are less readily quantified.....	153
Table A2.1	Methods to collect and collate expert opinion .....	156

# Contacts

---

The Medical Services Advisory Committee (MSAC) and its two sub-committees have secretariats within the Australian Government Department of Health.

Departmental Staff are available through the Health Technology Assessment (HTA) Team on the contact numbers and email below to discuss proposals for MSAC consideration or related matters. Any correspondence or assessment reports should also be lodged at via the address below. Staff within the HTA Team are also the first point of contact concerning the relevant committee or sub-committee's discussions and decisions.

HTA Team  
Australian Government Department of Health  
MDP 851  
GPO Box 9848  
CANBERRA ACT 2601

Or courier

Sirius Building  
23 Furzer Street  
PHILLIP ACT 2602

Phone: +61 2 6289 7550  
Fax: +61 2 6289 5540  
Website: <http://www.msac.gov.au>  
Email: [hta@health.gov.au](mailto:hta@health.gov.au)

## Record of updates

---

Date	Version	Summary of changes
March 2012	1.0	Draft Guidelines prepared by the MSAC Guidelines Working Group with technical and editing assistance from Biotext.
December 2012	1.1	Draft Guidelines separated into three Technical papers based on therapeutic, investigative and consultation types of medical services.
June 2013	1.2	Technical Guidelines edited for release. Changes to carry over to the Technical Guidelines: Investigative.
November 2013 Date when released?	1.3	Post public consultation adjustment based on internal review of stakeholder comments + amendments
February 2014	1.3	Internal review of stakeholder comments + amendments
August 2015	1.4	Post consultation
February 2016	1.5	Accessibility Standards applied
March 2016	2.0	Document renamed and released for publication Online

# Abbreviations

---

AHMAC	Australian Health Ministers' Advisory Council
ANZCTR	Australian New Zealand Clinical Trials Registry
AQoL	Assessment of Quality of Life
AR-DRG	Australian Refined Diagnostic Related Group
ARTG	Australian Register of Therapeutic Goods
CI	Confidence interval
CMA	Cost minimisation analysis
C-O	Cross-over
CUV	Cost-utility analysis
CV	Contingent valuation
DB	Double-blind
DOR	diagnostic odds ratio
DRG	diagnostic-related group
ESC	Evaluation Sub-committee
Guidelines	<i>Guidelines for Preparing Therapeutic Assessment Reports for the Medical Services Advisory Committee</i>
HTA	health technology assessment
HTA Review	<i>Review of Health Technology Assessment in Australia</i>
HUI	Health utilities index
ICER	Incremental cost-effectiveness ratio
ID	Identification
Investigative Guidelines	<i>Guidelines for Preparing Investigative Assessment Reports for the Medical Services Advisory Committee</i>
ITT	Intention-to-treat
MAUI	Multi- attribute utility instrument
MBS	Medicare Benefits Schedule
MCID	Minimal clinical important difference
MSAC	Medical Services Advisory Committee
NNT	Number needed to treat
PASC	PICO Advisory Sub-committee
PBAC	Pharmaceutical Benefits Advisory Committee
PBAC Guidelines	<i>Guidelines for Preparing Submissions to the Pharmaceutical Benefits Advisory Committee</i>
PBS	Pharmaceutical Benefits Scheme
PG	Parallel group
PLAC	Prostheses List Advisory Committee
PSD	public summary document
RR	Relative risk
SB	Single-blind
SD	Standard deviation

SG	Standard gamble
TGA	Therapeutic Goods Administration
TTO	Time trade-off
QALYs	Quality adjusted life years
Q-Twist	Quality- adjusted time without symptoms of the disease or toxicity
WTP	Willingness-to-pay



## **PART I**

### **General information**

# 1 Medical Services Advisory Committee

---

## 1.1 Purpose and roles of MSAC

The Medical Services Advisory Committee (MSAC) is a non-statutory committee established by the Australian Government Minister for Health in 1998. MSAC appraises new medical services proposed for public funding, and provides advice to Government about the level and quality of evidence relating to the comparative safety, clinical effectiveness, and cost-effectiveness of such services. Amendments and reviews of existing services funded by the Medicare Benefit Schedule (MBS) or other programs (for example, blood products or screening programs) are also considered by MSAC.

The MSAC advises the Minister for Health on medical services in relation to:

- the strength of evidence about the comparative safety, effectiveness, cost-effectiveness and total cost of the medical service;
- whether public funding should be supported for the medical service and, if so, the circumstances under which public funding should be supported;
- the proposed MBS item descriptor and fee for the service where funding through the MBS is supported; and
- other matters related to the public funding of health services referred by the Minister for Health.

MSAC also advises the Australian Health Ministers' Advisory Council (AHMAC) on health technology assessments referred under AHMAC arrangements.

There is no obligation on Government to accept or implement the advice MSAC provides.

## 1.2 Membership of MSAC

MSAC is an independent expert committee comprising professionals from the fields of clinical medicine, health economics and consumer matters. The Minister for Health determines the size and composition of MSAC. Members are drawn from a wide range of experts, constituted from time-to-time to address the likely type of applications for the committee's consideration. The current membership of MSAC is available on the MSAC website <http://www.msac.gov.au>.

## 1.3 MSAC sub-committees

MSAC currently has two sub-committees: the PICO Advisory Sub-committee (PASC) and the Evaluation Sub-committee (ESC). MSAC also has an Executive Committee (made up of the chairs of MSAC, ESC and PASC, and also the Deputy Chair of MSAC) to manage MSAC activities between formal committee meetings.

## **1.4 Overview of MSAC processes**

### **1.4.1 Regulatory framework**

All therapeutic goods used in the provision of medical services must be assessed by the Therapeutic Goods Administration (TGA) and included on the Australian Register of Therapeutic Goods (ARTG) before they can be marketed in Australia.

As a general rule, MSAC does not support public funding for a service that uses a therapeutic good for indications beyond those for which it was included on the ARTG.

An application to MSAC can be lodged before relevant therapeutic goods are included on the ARTG provided that the applicant has evidence that the relevant sponsor has commenced the TGA process. Confirmation of inclusion on the ARTG is required before MSAC can finalise its own appraisal of the corresponding medical service.

In considering whether to advise listing a service on the MBS, MSAC considers whether the service meets the criteria laid down in the *Health Insurance Act 1973*, and takes advice from the Department of Health on legal and policy matters as required.

### **1.4.2 The application and assessment process**

The approach to seeking MSAC advice to government for public funding is broken up into stages that provide stakeholders and the general public with opportunities to be actively engaged in the consultation phases, as well as opportunities for further applicant engagement throughout the process.

Unlike applications to the Pharmaceutical Benefits Advisory Committee (PBAC) and the Prostheses List Advisory Committee (PLAC), the costs incurred in submitting an application to MSAC is not recovered from the applicant. To ensure that only relevant information is collected, the scope of every application is determined before evidence is compiled.

MSAC may seek co-applicants or co-sponsors to broaden the scope of an application. In some instances, a professional body and more than one commercial company might be co-applicants in a combined application.

### **1.4.3 Sources of advice**

In formulating its advice, MSAC and its sub-committees may seek expert opinion from relevant professional bodies or appropriate specialists, and input from appropriate consumer bodies. Where external advice is obtained, the applicant is informed of the advice and given an opportunity to reply.

### **1.4.4 Publication of assessment report**

Reports will be published as submitted to ESC. Any agreed errors of fact will be provided separately as an errata. If commercial-in-confidence information has been identified by the applicant in the assessment report and where agreed confidentiality has been reached, applicants will have access to and will be required to confirm that the modified version of the assessment report (submitted or contracted), and the modified MSAC Public Summary Document (PSD) can be uploaded on the MSAC website.

#### **1.4.5 Timing of MSAC processes**

MSAC advises all interested stakeholders of the meeting dates for the following year, as well as the associated cut-off dates via the MSAC website.

Assessment reports should be presented on time, complete, in the format requested in the associated template and with the correct number of copies. No guarantee can be given that material supplied late will be incorporated into the assessment report or included in the agenda papers.

For PBAC co-dependent integrated reports (material being presented to both PBAC and MSAC), the PBAC requirement for report formatting and publication will prevail in acknowledgement of the different government public funding arrangements, Pharmaceutical Benefit Scheme (PBS) listing and established memorandum of understanding arrangements.

Initial advice of committee decisions for co-dependent applications to MSAC and PBAC will, where possible, follow the PBAC approach to provision of advice to applicants.

#### **1.4.6 MSAC appraisal**

MSAC will appraise the evidence presented in the assessment and ESC reports to inform its advice to government. MSAC prepares a detailed rationale for its conclusions in the form of a PSD.

Where specific material is agreed to be confidential, the PSD will be published with the confidential material redacted. The Department offers debrief meetings to applicants following the public release of MSAC's advice, if requested.

Following MSAC's consideration, the Department of Health is required to consider the financial impact to government, consult with relevant stakeholders, seek Cabinet agreement and draft and implement legislative change to amend or add an item to the MBS. As previously advised in Section 1.1, there is no obligation on Government to accept or implement the advice MSAC provides.

Please note that ESC and MSAC does not meet with or accept face-to-face presentations of evidence from applicants.

## 2 Introduction to the Guidelines

---

These *Guidelines for Preparing Assessment Reports for the Medical Services Advisory Committee* (referred to in this document as the ‘Guidelines’) provide practical information on how to present evidence to MSAC when seeking Australian Government funding of a medical service.

Although these Guidelines have been written for applicants from the medical profession and industry, they are also intended to provide information to other interested stakeholders, including clinical and patient groups, and the general community.

### 2.1 Structure of these Guidelines

These Guidelines are organised into four parts, as follows:

- **Part I General information**  
This part covers information on the preferred layout and style conventions, different types of applications and a checklist with a navigation aid of the information that is to be contained in reports for particular types of assessment reports.
- **Part II Clinical and economic evidence provided in the most preferred format**  
This part covers the evidence for public funding for the proposed medical service, when it is available in the most preferred format. The Sections in Part II follow the order in which information should be presented in the assessment report:
  - A Context (details of the proposed medical service and its intended use of the MBS)
  - B Clinical evaluation
  - C Translation issues
  - D Economic evaluation
  - E Financial implications
  - F Other
- **Part III Clinical and economic evidence provided in alternative formats**  
This part covers situations where the evidence is not available in the most preferred format.
- **Appendices** include additional information on various aspects of the assessment report.

Further information is available in the associated template for the therapeutic assessment report.

### 2.2 Associated documents

A template for the therapeutic assessment report is available on the MSAC website and should be used when developing a therapeutic assessment report in line with these Guidelines.

Applicants may also need to refer to the *Guidelines for Preparing Investigative Assessment Reports for the Medical Services Advisory Committee* (referred to in this document as the *Investigative Guidelines*).

## 2.3 What is a therapeutic medical service?

A therapeutic medical service is one that improves health outcomes **directly**; that is, no other intermediate medical service needs to be provided to achieve the improvement in health outcomes. Such a service might also directly harm the individual; however, this means that there needs to be a health ‘benefit to harm’ assessment of health outcomes.

Current health technology assessment (HTA) guidelines have comparative effectiveness as the primary measure of a health outcome. This is combined with comparative safety to determine net clinical benefit, and then combined with comparative cost to determine incremental cost-effectiveness (or cost-minimisation if the net clinical benefit of the proposed medical service is non-inferior rather than superior to the main comparator).

The proposed use of the term ‘therapeutic’ for this type of item is based on the *Macquarie Dictionary* definition, which covers both ‘*of or relating to the treatment of disease*’ and ‘*serving or performed to maintain health*’. However, caution is needed in the use of this term because ‘*therapeutic good*’ is defined more broadly by the *Therapeutic Goods Act 1989*.

Examples of this type of medical service include a novel surgical technique, the insertion of a stent or other therapeutic device. Most blood products (the main non-MBS type of application to MSAC) would also come under this subtype.

In some instances a therapeutic medical service is co-dependent on another medical service. A co-dependency occurs where the use of one health technology to directly improve health outcomes (e.g. a medicine, or medical device or procedure) is improved by the use of another health technology (e.g. a pathology or an imaging technology) and where both technologies require consideration for public funding. Possible co-dependencies involving investigative medical services include:

- investigative medical service + therapeutic medical service (both requiring funding approval through MSAC);
- investigative medical service (funding approval through MSAC) + therapeutic; or
- medical service (requiring funding approval through another committee e.g. co-dependent pharmaceutical that requires coordinated consideration for PBS funding by PBAC).

Further information on co-dependencies is available in Sub-section 4.3.3 of these Guidelines.

## 2.4 Writing and style conventions used in these Guidelines

Several conventions have informed the revision of these Guidelines to assist users of the document to navigate their way to the information needed when preparing their assessment reports.

These Guidelines include a series of requests for specific types of information. The aim is to provide an ordered series of reference points (requests for information) against which the specific information presented in an assessment report can be evaluated to ensure that the assessment report is complete.

The ‘default’ writing style for requests for information uses the imperative voice, as follows:

*‘Describe the proposed course of treatment’ or ‘Justify the exclusion of the study’.*

Readers should interpret these imperative statements as indicating what **should** be done. This allows requests for information that is known to be more persuasive or influential to be communicated as simply as possible in these Guidelines. Following these requests helps to improve the comparability of assessment reports considered by MSAC, and hence the consistency of decision making.

Within each Section, the main requests for information expected to be addressed by each standard assessment report are highlighted as **‘Information requests’**. Other subsidiary requests and background information are provided in normal text.

In some instances, the request includes the word **‘must’**. In each case, the requirement is included in the information request under the separate heading of **‘Information requirements’**. Failure to comply with these requirements is sufficient to render the assessment report unacceptable, and for the assessment report to be returned to the applicant.

In other instances, there is no basis to indicate a preference for one type of information over another. In these instances, options about what **could** be presented are usually given. MSAC is generally indifferent about which option is presented, although the context of a particular assessment report might suggest the basis for expressing a preference. The assessment report should therefore explain the basis for selecting the information presented.

## 2.5 The future

Future revisions of these Guidelines will be disseminated via the MSAC website. A summary of each change will be recorded at the front of the electronic version published on the website, and those involved in preparing assessment reports will be notified.

Further feedback on these Guidelines is welcome and should be forwarded to the HTA Team at [hta@health.gov.au](mailto:hta@health.gov.au).

# 3 Rationale and basis for the economic evaluation in the Australian context

---

## KEY POINTS — APPROACH TO ECONOMIC EVALUATION

- MSAC is required to assess the degree to which new, amended or revised medical services represent ‘value for money’ for the Australian community.
- The economic evaluation should focus on the effectiveness of the proposed medical service compared with existing medical services, its cost and the likely changes in the provision of health care resources after its introduction (including changes in the provision of other health care resources not funded through the MBS).
- Economic evaluations should be relevant to the Australian context.
- The practical aspects of the economic evaluation of the performance of medical services are challenging; therefore, there will be continued flexibility in the interpretation of these Guidelines.

Australia, like other countries, is faced with a steady increase in the total cost of medical services. Although the medical service budget is not ‘capped’ in Australia, choices must be made as to which medical service will be subsidised by the Australian Government. Economic evaluation is one factor to be considered when making choices among competing medical services. Other important factors that are considered include uncertainty, equity, extent of use and total costs.

### 3.1 Analysis of cost-effectiveness

MSAC considers the results of economic analyses in its decision making to assess the degree to which new or revised medical services represent ‘value for money’ for the Australian community.

### 3.2 Australian context

Although the results of clinical trials or studies of sufficient scientific rigour done overseas are a reasonable basis for economic evaluations relevant to the Australian health care system, an economic evaluation performed overseas will often not be relevant in Australia. This is because of standard differences in unit costs, the patterns of resource provision and the way in which health care is funded in other countries. Applicants are therefore encouraged to submit an economic evaluation that is relevant to the Australian context in Australian dollars.

### 3.3 Relevant factors influencing MSAC decision making

MSAC considers many factors when proposing that a medical service be publicly funded. Each of these factors might have a separate influence on the decision to list the proposed medical service and, depending on the circumstances of each consideration, might influence MSAC in favour of, or against, listing. More than one factor might be relevant to each consideration.



Tables A1.1 and A1.2 in Appendix 1 list relevant factors, which are divided into two groups: quantitative and qualitative. The qualitative factors (Table A1.2) include some of the underlying assumptions implicit in such concepts as quality-adjusted life-years (QALYs) and discounting. To enable consistency within an assessment report regarding these factors, MSAC has adopted a particular position (which is specified in these Guidelines in the Sections indicated by the cross-references in the tables). However, in certain circumstances, it might be reasonable to argue that a different position should be considered.

Individual factors are not weighted equally by MSAC in its decision-making process, and different factors might be more or less important in different situations. In other words, the importance of any particular factor cannot be quantified. The descriptions provided in Appendix 1 represent MSAC's understanding at the present time. MSAC continues to reflect on its processes and further develop its understanding of these matters.

### **3.4 Flexibility in interpretation of these Guidelines**

Despite the differences in data available and uncertainties that might exist in the base case, it is in the interests of the community, industry and MSAC that uniformity be maintained in the way that economic analyses are conducted and evaluated. However, the practical aspects of the economic evaluation of the performance of medical services are challenging for applicants, MSAC and the administrative arm of government. For this reason, although applicants should present the economic analysis as outlined in these Guidelines, there will continue to be the need to be flexible in the interpretation of these Guidelines.

## 4 Organisation of a standard assessment report

---

### KEY POINTS — ORGANISATION OF A STANDARD ASSESSMENT REPORT

- Assessment reports must consist of an executive summary, the main text of the assessment report and additional information (attachments and technical documents).
- Part II (for the majority of assessment reports) and Part III (for supplementary and alternative information in some assessment reports) of these Guidelines provide the preferred order for presenting information in the main text of standard assessment reports.
- The preferred order for presenting information consists of six Sections (A–F). If possible, do not present information in any other order, because this will reduce MSAC’s ability to effectively evaluate the assessment reports.
- Use frequent, accurate cross-referencing between the executive summary, main text and other technical documents.
- Use succinct, plain English wherever possible (while maintaining scientific rigour).
- Provide justification for any variations to the requested information.
- If using a new analytical technique, present the base case using both the requested methods and the new technique for comparison.

These Guidelines are designed to assist applicants to identify and present the basic information required by MSAC to determine its advice and to provide guidance to applicants on the most appropriate form of economic evaluation for the specific assessment reports.

This Section outlines the information that should be presented in a standard assessment report. A flowchart showing MSAC’s key decisions in evaluating standard assessment reports is also included, along with advice on presenting alternative information in particular circumstances.

### 4.1 Choice of information

The information should address the PICO Confirmation agreed by the PASC or the MSAC Executive, however, an applicant, if they choose to take the risk of not to adhere to the PICO Confirmation should note that this may impact the final decision made by MSAC. These Guidelines set out the information requested, and while additional information might be included, it must be clear that this additional information addresses matters that are outside the Final PICO Confirmation.

A wide array of information should be presented in a standard assessment report to MSAC. Some information is requested for all assessment reports, whereas some additional information requests only apply according to the type of service for which funding is being sought. In addition, a large number of information requests provide guidance on presenting the ‘next best’ option when it is not possible to provide the preferred information.

Each assessment report should be as succinct and informative as possible. MSAC and ESC are most likely to be influenced by arguments based on scientifically rigorous data rather than opinions. Assessment reports should use suitable scientific language, but avoid jargon.

## **4.2 Overview of a standard assessment report**

### **4.2.1 Sections of a standard assessment report**

To achieve a 'base case' estimate and decision analysis with uncertainty identified, a standard assessment report needs to include an executive summary and Sections A–F as shown in Figure 4.1.

These Guidelines are set out to provide a stepped approach for the presentation of the best and most persuasive evidence. The most preferred option is described in Part II and alternative, less preferred options are provided in Part III.

The order of the information requests indicate the preferred flow of information. The requests in parts II and III refer to all medical services and products.

## **4.3 Presentation of the assessment report**

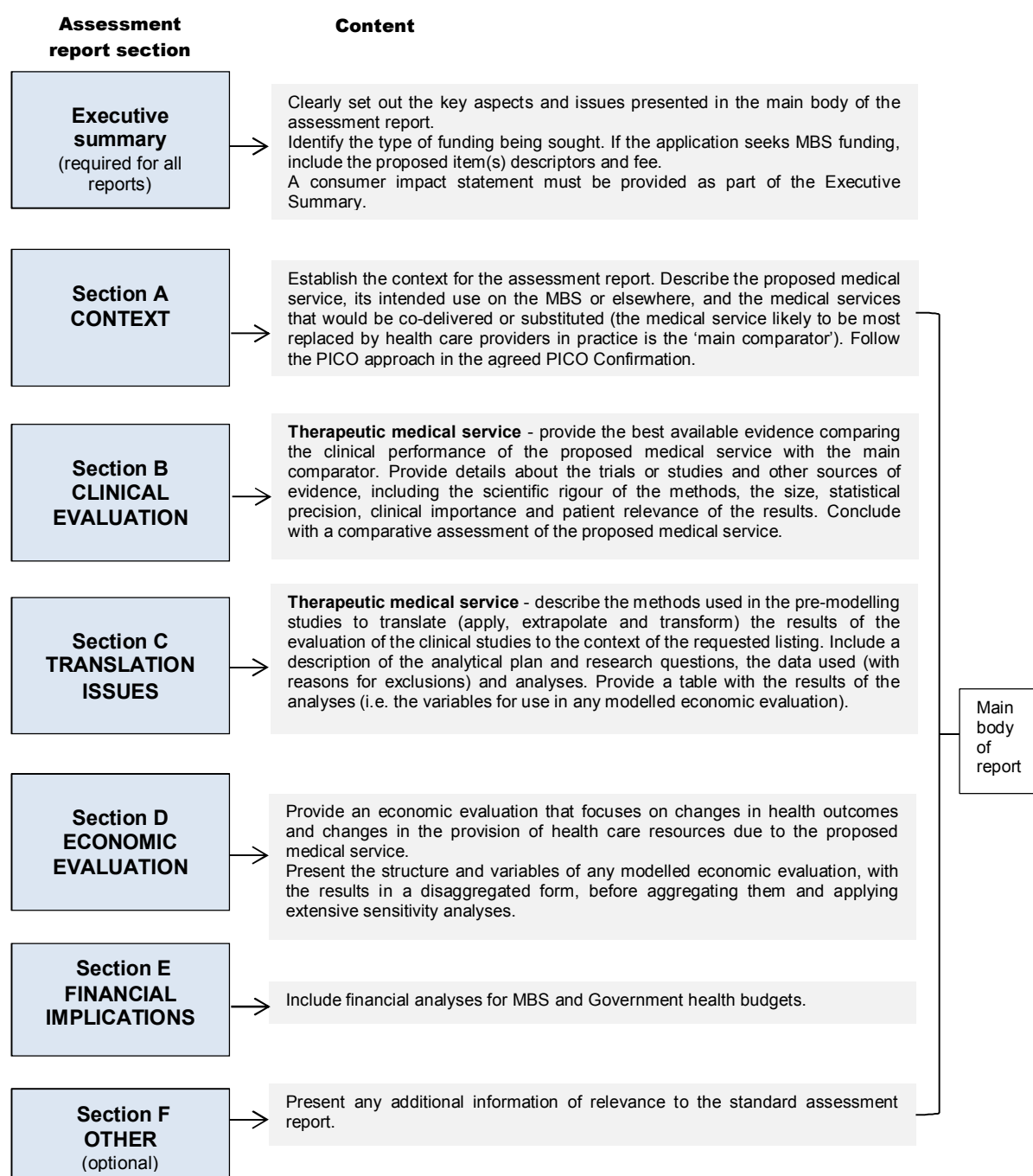
### **4.3.1 Standard assessment report**

The main body of the assessment reports **must** be presented according to the MSAC *Therapeutic Assessment Report Template* available on the MSAC website. Key reports of the relevant trials on which the report is based must be provided separately. Other information might be provided as attachments or technical documents. This other supplementary material is made available to committee members on request. Where the report relies on specific information, it should be referenced (if publicly available) or included within the report and available for publication. Where the conclusions in a report rely on agreed commercial-in-confidence material it might, by agreement, be provided as a not-for-public-release attachment.

It is vital that the assessment report includes frequent and accurate cross-references between the executive summary and the main body of the assessment report, and between the main body of the assessment report and reports of the key trials, attachments, technical documents and material in electronic formats. This will assist those who have to evaluate and consider the assessment report.

The key steps for preparation of a standard assessment report and how these decisions relate to the Sections of the assessment report, are set out in Parts II and III of these Guidelines. The order of the information requests in Part II and / or Part III indicates the preferred order for the information that should be presented to optimise its evaluation by the MSAC. Arranging the same information in another order has generally been found to be unhelpful.

**Figure 4.1      Sections of a standard assessment report**



PICO = population; intervention (medical service), comparator; outcomes

### **4.3.2 Two stage approach to an assessment**

Some applicants may advise their preference for submitting the clinical component (Sections A and B) of their submission based assessment before commencing their economic component (Sections C to F). This would provide the benefit of the ESC and MSAC's feedback on both the clinical evidence and the proposed structure of the economic model.

In commencing this pathway, applicants should be aware that it will require a two stage approach to the submission, ie. it will be considered by ESC and MSAC twice. The first stage would be presenting Sections A and B, and then the second stage, presenting Sections C to F.

If an applicant chooses this pathway, consideration should be given to submitting the draft financial component (Section E), alongside that of Sections A and B.

### **4.3.3 Co-dependent/integrated assessment reports**

Co-dependent applications is discussed in the *Investigative Guidelines* as most co-dependent applications relate to the use of an investigative service with a medicine. While this is the case, other combinations of co-dependent services will occur from time-to-time.

## 5 Lodging an assessment report

---

### 5.1 Assessment report checklist

#### 5.1.1 Information requirements

As indicated in the template, the therapeutic assessment report must consist of the following components:

- contents;
- executive summary (including a consumer impact statement);
- main body of report;
- attachments; and
- appendices.

Each hard copy of the main body of a standard assessment report must be suitably bound, identified, indexed, pages numbered and divided with labelled tabs.

All economic calculations must be provided in Australian dollars.

The investigator's summary of each trial report, the main published paper, and an adequate account of the methods and results for each trial or study must be included as attachments within the main body of the assessment report.

All submitted information must be legible and in English.

All assessment reports, unless otherwise specified, will be made public in the format in which they were lodged.

#### 5.1.2 Information requests

Before lodging an assessment report, the applicant must notify the Health Technology Assessment Team ([hta@health.gov.au](mailto:hta@health.gov.au) or (02) 6289 7550) of their intention to lodge an assessment report using the timeframe published on the MSAC website and to also receive up-to-date detail on the information requirements (i.e. the number of hard copies required etc).

A checklist is provided at Table 5.1 as an initial guide to assist applicants in this process.

**Table 5.1 Checklist of information to be included in a therapeutic assessment report**

<b>Component</b>	<b>Included?</b>
<ul style="list-style-type: none"> <li>The original, signed covering letter for the assessment report (with an attachment containing the complete index to the assessment report).</li> </ul>	Yes/No
<ul style="list-style-type: none"> <li>A comprehensive index attached to the covering letter, which serves as a checklist for all documentation and other materials comprising the assessment report and confirming: <ul style="list-style-type: none"> <li>the numbers of copies of the main body of the assessment report and details of its contents</li> <li>the numbers of copies of other parts of the assessment reports and details of their contents.</li> </ul> </li> </ul>	Yes/No
<ul style="list-style-type: none"> <li>The current TGA-approved product information with approval date (if and when available, with the latest draft product information in the meantime; each copied single-sided and stapled) (where relevant).</li> </ul>	Yes/No
<ul style="list-style-type: none"> <li>The letter of registration with details of marketing approval and registration (if and when available; each copied single-sided and stapled) (where relevant).</li> </ul>	Yes/No
<ul style="list-style-type: none"> <li>Any additional technical documents, attachments and references provided separately to the main body of the assessment report (where relevant), which should: <ul style="list-style-type: none"> <li>be suitably bound (i.e. each folder is robust enough to withstand regular use, with the width of its spine matching the number of pages it contains)</li> <li>have the contents identified on the cover</li> <li>be legible and in English (or accompanied by a reputable translation)</li> </ul> </li> </ul>	Yes/No
<ul style="list-style-type: none"> <li>Bound copies of the main body of the assessment report (using the agreed template), which <b>must</b>: <ul style="list-style-type: none"> <li>be suitably bound (i.e. each folder is robust enough to withstand regular use, with the width of its spine matching the number of pages it contains)</li> <li>have the contents identified on the cover</li> <li>have a clear and adequate index (which encompasses both the main body of the assessment report and the contents of all other documentation contained in separate volumes, and also identifies all other materials supplied as part of the assessment report, which is also attached to the covering letter of the assessment report)</li> <li>have consistent pagination throughout</li> <li>include dividers between each Section, attachments and references, with an appropriately labelled tab extending beyond the page width</li> <li>have all cost calculations in Australian dollars (A\$)</li> <li>incorporate attachments containing reports of each of the relevant randomised trials (or each of the relevant non-randomised studies, if necessary), which <b>must</b> be: <ul style="list-style-type: none"> <li>(i) the investigator's summary of each applicant's trial report and the main published paper (where available), together with adequate details of the trial methods, analysis and all trial results presented in the assessment report for use in the economic evaluation; OR the main published paper alone if the applicant has no access to a more detailed report</li> <li>(ii) legible and in English (or be accompanied by a reputable translation)</li> </ul> </li> </ul> </li> </ul>	Yes/No
<b>Electronic versions of the assessment report on a USB</b>	
<ul style="list-style-type: none"> <li>Supply the whole assessment report and any accompanying calculations and models in electronic format (with any spreadsheet compatible with Microsoft Excel 2010-13, RevMan, any word-processing document compatible with Word 2010-13, and any other software package consistent with Sub-section 5.2.1). Ensure that all components of these electronic documents, spreadsheets and analyses are fully accessible (e.g. do not have password protection); fully enabled to allow all document text, tables and figures to be accessed for copying; and fully executable to allow all spreadsheet cells and all statistical or decision analysis input variables to be changed.</li> </ul>	Yes/No
<ul style="list-style-type: none"> <li>Supply electronic copy of key articles that the conclusions in the report are based on.</li> </ul>	Yes/No

## **5.2 Provision of information to allow independent verification of computer analyses**

### **5.2.1 Information requirements**

Provide sufficient information to permit independent verification of computer-based analyses to generate information for the assessment report Sections C or D (e.g. input data, methods of analysis, outputs).

Provide an electronic copy of all computer-based analyses (including the economic evaluation) in the form in which it was conducted, together with any associated data files, and a technical document or an attachment with clear cross-references to the assessment report.

Use a software package that can be readily evaluated by MSAC or, before lodging the assessment report, discuss the arrangements with the HTA Team to ensure the acceptability for evaluation of any software that is not on the maintained list of software packages. Examples of software include Word (2010-2013 compatible), Excel (2010-2013 compatible), STATA, Triage, RevMan, Endnote, etc.

## **5.3 Provision of information after lodgement of the assessment report**

### **5.3.1 Post-lodgement communication with MSAC**

MSAC procedures provide post-lodgement opportunities for applicants to communicate with MSAC.

It is expected that applicant responses will address issues raised in the relevant papers rather than introduce substantive changes, such as a different population identified by a modification to the requested restriction, a different nomination for the main comparator, new data or new analyses. Such changes might result in an MSAC request for a standard reassessment to examine the implications of the substantive change.

Before the departmental papers are finalised, applicants might be approached by either the Department or an assessment group for further information or clarification of aspects of their assessment report. Applicants are expected to deal with these requests expeditiously.

### **5.3.2 Provision of information sourced from the TGA after lodgement of the assessment report**

Upon receipt of notification of TGA registration approval, applicants are requested to advise the HTA team (via [hta@health.gov.au](mailto:hta@health.gov.au)) immediately, in writing, of any aspect of an assessment report that is not consistent with the final TGA registration. At this time, also provide a copy of the TGA-approved product information, accompanied by a document highlighting any variation between the most recent draft provided with the assessment report and the subsequent TGA-approved product information that would have any bearing on the consideration of the assessment report or on the consideration of any subsequent MSAC recommendation to list.



## **Part II**

### **Preferred clinical and economic evidence for proposed medical services to be considered by MSAC**

# **Section A**

## **Details of the proposed therapeutic medical service and its intended use on the Medical Benefits Schedule (or for other public funding)**

---

### **Introduction**

Section A of the assessment report establishes the context for the report. It provides the information outlined in the PICO Confirmation that has been agreed to by PASC or the MSAC Executive in the pre-assessment phase of the application to MSAC.

### **A1 Address all items in the agreed PICO Confirmation**

- All items in the agreed PICO Confirmation should be addressed in the assessment report.
- If any items are not addressed this presents a risk to the applicant; these items should be identified and reasons provided for not addressing them.
- Confirm that the assessment report has fully addressed the questions defined in the agreed PICO Confirmation.
- Indicate if any additional information provided in the assessment report has been compared to the agreed PICO Confirmation.

### **A2 Proposed medical service**

Describe the key components of the proposed medical service as set out in the agreed PICO Confirmation, including mode of delivery and a broad description of the support infrastructure and type of facility required to deliver the service.

### **A3 Proposed MBS listing or other public funding sought**

Provide MBS or other public funding descriptors, as set out in the agreed PICO Confirmation. Differences between the proposed descriptor and the descriptor provided in the agreed PICO Confirmation should be highlighted and a justification provided in the assessment report.

### **A4 Proposed population**

- Identify the main population(s) described in the agreed PICO Confirmation including key inclusion and exclusion criteria. This may involve the results of prior tests to exclude or include patients in the proposed population.
- Include a high level summary of the frequency (prevalence and / or incidence) of the population or disease in question and where relevant the natural history / pathophysiology of the condition of interest.

### **A5 Comparator details**

- Identify the main comparator(s) described in the agreed PICO Confirmation.

- If there are any additional comparator(s) justify their selection.
- Identify any other factors that might affect the identification of the main comparator in the future.

## **A6 Clinical management algorithm(s)**

- Present the clinical management algorithm(s) described in the agreed PICO Confirmation.
- Present the clinical management algorithm that depicts the context of the intended use of the proposed medical service following a listing on the MBS or other public funding, as listed in the agreed PICO Confirmation.
- Present the corresponding algorithm depicting the current context as listed in the agreed PICO Confirmation.
- Highlight the differences between the two algorithms to summarise the changes in the patterns of resource provision, both those required by any requested indication and those that would be expected to follow as consequences of the requested listing.
- Indicate whether multiple-listing scenarios are presented.

### **A6.1 Algorithms for intended and current contexts**

Clinical management algorithms are most relevant to an assessment report presenting a modelled economic evaluation (see Sub-section D1). They are also helpful for estimating changes in use and cost of other medical services (see Sub-section E3). An assessment report not presenting a modelled economic evaluation might only need to present straightforward algorithms.

The objective of these clinical management algorithms is to help clarify the comparison addressed in the assessment report through the following three steps:

- Define the eligible patients and the circumstances of use if the listing or public funding were implemented as requested (algorithm 1).
- Identify the current situation in terms of the expected substitution of service options for these patients and their circumstances of use, both at the time of substitution and subsequently (algorithm 2).
- Identify the full nature of the comparison(s) being made in the assessment report and limit the comparison to these contexts (highlight the differences between algorithms 1 and 2).

The algorithms are expected to be of varying complexity, depending on the particular contexts to be described in each assessment report. Overall, ensure that the algorithms identify the nature of any and all differences across the full streams of resource provision consequences, both before and after the point(s) in the algorithm at which the proposed medical service is introduced. This ensures greater clarity about the context of the intended use of the proposed medical service in terms of patients and circumstances, from which the comparative health outcomes, comparative costs, comparative cost-effectiveness and financial implications can all be estimated.

In each algorithm, summarise all:

- relevant diagnostic and treatment steps, including all:

- required previous medical services; and
- diagnostic criteria and/or tests (including those demonstrating that one or more previous medical services cannot be used to manage the indication, and including those required to support any continuation criteria in the requested restriction);
- required co-delivered services; and
- consequences for subsequent service options.

Specify any other important characteristics of patients and types of circumstances of use. Examples include specifying the characteristics of the medical condition in the eligible patients (e.g. in terms of risk factors) and the aspects of the spectrum of the medical condition (e.g. in terms of severity of disease or remaining treatment options). Sub-section D2 provides further examples.

Justify the basis for the selection of the algorithm with reference to a literature review of relevant published clinical management guidelines. Provide a copy of those clinical management guidelines in an attachment or technical document. If expert opinion or survey has been used to help specify the clinical management algorithms.

## **A7 Differences between the proposed medical service and the main comparator**

Describe the main differences in the indications, contraindications, likelihood and severity of adverse events between the proposed medical service and the main comparator(s).

## **A8 Clinical claim**

Provide information about the clinical claim with respect to the proposed therapeutic medical service, as set out in the PICO Confirmation, against the main comparator. The clinical claim is to be tested in Section B.

## **A9 Summarise the primary elements of the decision analysis (PICO)**

Provide the PICO (population / problem, intention, comparator and outcome) criteria and decision option(s) for the proposed therapeutic medical service, as set out in the PICO Confirmation.

Primary elements for a therapeutic medical service include:

- population and medical condition including results of prior tests that would inform which patients are included or excluded from the proposed population;
- proposed therapeutic service;
- comparator service; and
- outcome claim.

## Section B

### Clinical evaluation for the proposed therapeutic medical service

---

#### Introduction

The purpose of Section B is to identify and present the best available clinical evidence for the main indication for a therapeutic medical service.

Sub-section B1 sets out the requested search strategy to identify all trials that can be used to compare the proposed therapeutic medical service with its main comparator. The MSAC has a strong preference for clinical and economic evaluations that are based on randomised trials that directly compare the proposed therapeutic intervention with the main comparator (referred to in these Guidelines as ‘direct randomised trials’, but also known as ‘head-to-head trials’). However, MSAC recognises that such trials are not always available. If this is the case, alternatives might be (in order of priority):

- an indirect comparison across two or more sets of randomised trials involving one or more common reference (indirect comparison of randomised trials); and
- non-randomised studies (including comparisons involving single arms extracted from randomised trials).

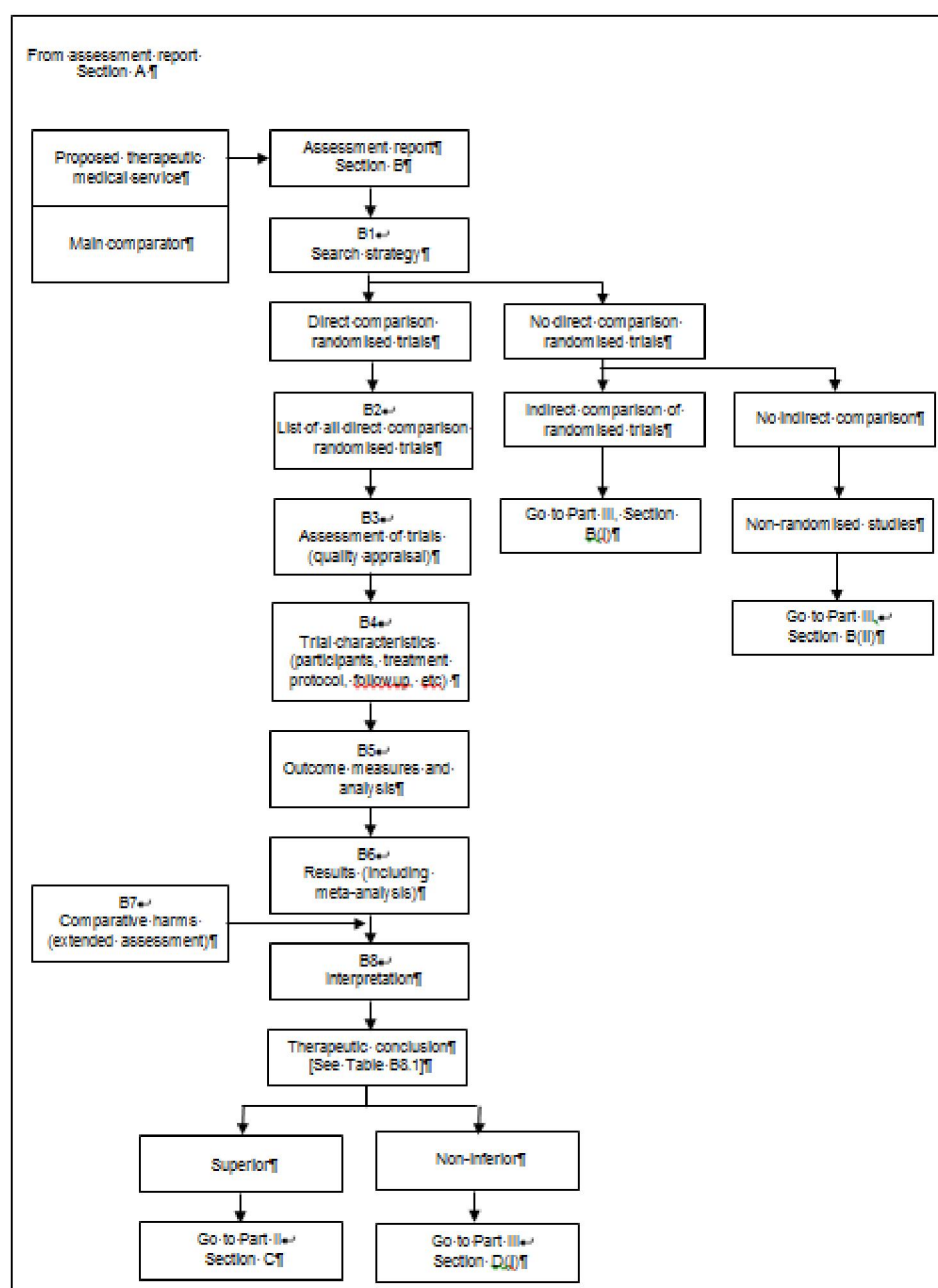
Part III, Section B(i) and Section B(ii) of these Guidelines provides guidance for presenting Section B in the assessment report based on these other types of studies. Figure B1 shows a flowchart of these options.

The clear preference for evidence from the most scientifically rigorous sources does not imply that a minimum standard must be met. MSAC has considered and will continue to consider all levels of evidence. However, MSAC will be most influenced by the results of direct randomised trials as the most rigorous data source.

Therefore, the remainder of Section B relates to assessing the characteristics of direct randomised trials, and interpreting the results.

Figure B1

## Key information requests for assessment report Section B of an assessment for MSAC



## B1 Description of search strategies

### INFORMATION REQUEST

- Describe the search strategies and characteristics used to locate reports of potentially relevant trials from the published literature, registers of randomised trials and unpublished sources held by the applicant.

#### Search strategies

The primary objective of the search strategies is to locate *all* randomised trials that, for the main indication, compare the proposed therapeutic medical service directly with the main comparator for participants with characteristics that overlap with patients who would be eligible to use the proposed therapeutic medical service.

The search should involve at least four approaches:

- a search of the published literature, including reviews by an overseas regulatory body (e.g. the US Food and Drug Administration) or by an overseas health technology assessment (HTA) agency (e.g. the English National Institute for Health and Clinical Excellence);
- a search of registers of randomised trials;
- if relevant, an examination of the dossier seeking marketing approval submitted to the TGA, supplemented by checks with the applicant's head office and subsidiaries of the company (and any other original applicant or co-licensed companies) for any further randomised trials (which might be unpublished); and
- manual checking of reference lists of all relevant articles that are obtained by other means.

When describing the search strategies and characteristics, sufficient detail should be provided so that an independent replication of the search would yield the same results.

The methods used to search the published literature are pivotal to assessing the completeness of the overall search. Therefore, specify the following characteristics of the search strategy:

- the specific databases and registers of clinical trials searched, including at least MEDLINE, EMBASE, the Cochrane Library (including the Cochrane Database of Systematic Reviews, and the Cochrane Central Register of Controlled Trials), the National Institute of Health and the Australian New Zealand Clinical Trials Registry (ANZCTR). The search should also include databases internal to the company and any other known registers of randomised trials relevant to the therapeutic area;
- the date the search was conducted;
- the date span of the search (which should include the most recent update of each database searched);
- the complete search strategies used, including the search terms (key or MeSH words) and the relationship (sets and Boolean) between the search terms; and
- any supplementary searches, especially manual checking of references in the retrieved papers from the database searches.

## B2 Listing all direct randomised trials

### INFORMATION REQUEST

- The assessment report must identify and list all relevant direct randomised trials.
- If no relevant direct randomised trials are found in the searches, a 'nil return' must be included in the assessment report.
- Present tables listing all citations of the direct randomised trials identified from the search of the published literature, marketing dossier and other sources. Show the inclusion and exclusion criteria for identifying relevant trials, and state which trials have been published.
- On the hard copy of each of the search printouts supplied as technical documents with the assessment report, annotate each citation to indicate excluded citations with the reason for the exclusion.
- Collate all reports of each direct randomised trial to create a master list and indicate the preferred identification (ID) for each trial to be used throughout the assessment report for consistency.
- Justify the exclusion of any relevant direct randomised trials. Tabulate a summary that highlights key aspects of the identified trials, presenting included and then excluded trials.
- Separately identify any meta-analysis of randomised trials and assess their exclusion or inclusion using the same criteria as above. Include any relevant systematic reviews; for example, the Cochrane Database of Systematic Reviews.
- Identify any direct randomised trial that was designed prospectively as a non-inferiority trial and/or whether the therapeutic conclusion presented in response to Sub-section B8 is one of non-inferiority or equivalence.
- Include copies (or sufficient details) of the included trials as attachments in the main body of the assessment report and ensure that the location of each item is shown clearly in the assessment report index.

The listing of relevant direct randomised trials must be complete to satisfactorily address publication bias, duplication bias and outcomes reporting bias. The assessment group will run an independent literature search and if this search retrieves relevant trials that were not listed in the assessment report, processing of the assessment report will stop until the matter has been resolved.

If no relevant direct randomised trials are found in the searches, the assessment report **must** include a statement to this effect with the results of the searches.

### Search results

Assess all citations retrieved by the searches (see Sub-section B1) to extract all trials that meet each of the following inclusion criteria for direct randomised trials:

- a) the trial included a randomisation procedure in its design;
- b) the trial compared the proposed therapeutic medical service and the main comparator in separate arms; and
- c) the trial recruited participants with characteristics that overlap with those of patients who would be eligible for the main indication.



Of these criteria, only (c) requires an element of judgment. If there is any uncertainty about whether to include or exclude a direct randomised trial, it is usually wiser to include it.

Tables B2.1 and B2.2 provide a suggested format (where appropriate) for presenting the search results to summarise the inclusion and exclusion of citations from the results of searches reported in response to Sub-section B1.

**Table B2.1**      **Summary of identification of direct randomised trials from the search of the published literature**

	MEDLINE	EMBASE	Trial registries	Other databases
Number of citations retrieved by search				
Number of citations excluded after title/abstract review: <ul style="list-style-type: none"> <li>not a randomised trial</li> <li>randomised trial does not include the proposed therapeutic medical service and the main comparator in separate arms</li> <li>characteristics of the recruited participants do not overlap with the main indication</li> </ul> <b>TOTAL</b>				
Number of citations excluded after full text review: <ul style="list-style-type: none"> <li>not a randomised trial</li> <li>randomised trial does not include the proposed therapeutic medical service and the main comparator in separate arms</li> <li>characteristics of the recruited participants do not overlap with the main indication</li> </ul> <b>TOTAL</b>				
Number of citations of direct randomised trials included from each database				
Consolidated number of citations of direct randomised trials (removing exact duplicates across different databases)				
Number of multiple (additional) citations of direct randomised trials identified				
Number of published direct randomised trials included				

**Note:** Present columns that correspond with submitted printouts (e.g. if the printouts combine MEDLINE and EMBASE, these results can be combined in the table).

**Table B2.2 Summary of identification of applicants<sup>a</sup> direct randomised trials and information from the manual search of retrieved citations**

	<b>TGA dossier (where appropriate)</b>	<b>Other 'in-house' trials</b>	<b>Manual search</b>	<b>Total</b>
Number of reports or citations of randomised trials retrieved				
Number of randomised trials excluded: <ul style="list-style-type: none"> <li>randomised trial does not include the proposed therapeutic medical service and the main comparator in separate arms</li> <li>characteristics of the recruited participants do not overlap with the main indication</li> </ul> <b>TOTAL</b>	<b>b</b>			
Number of direct randomised trials included from these searches	<b>b</b>			
Number of these direct randomised trials identified in Table B2.1	<b>b</b>			
Number of other direct randomised trials identified in Table B2.1				
<b>Total number of direct randomised trials included</b>				

TGA = Therapeutic Goods Administration

**a** For the purposes of the search for relevant randomised trials, 'applicant' is the entity responsible for lodging the assessment report and includes any original applicant (including head office and all subsidiaries) and/or any co-licensing entity responsible for the proposed therapeutic medical service in addition to the applicant.

**b** Separately list and identify each of these trials using the identifying nomenclature used for the trials in the TGA evaluation reports to enable a cross-check against the trials considered by the TGA.

**Note:** If the only source of a direct randomised trial relevant to the assessment report is located by a manual search within an independently conducted meta-analysis (preferably published in a peer-reviewed journal and incorporating all important trials listed in this Section B), count the trial here and list the trial with the master list as shown in Table B2.3.

### **Annotated search printouts**

On the hard copy of each of the search printouts supplied as technical documents with the assessment report, annotate each citation as appropriate with the letter (a), (b) or (c) to indicate which of the above criteria were invoked to exclude that citation. Each citation without an annotation should thus be a report of a direct randomised trial included in the assessment report.

### **Master list of trials**

Table B2.3 provides a suggested format for presenting a master list of all the direct randomised trials identified in the search.

For any trial that has not yet reported any results, indicate when results are expected to become available.

**Table B2.3 Trials (and associated reports) presented in the assessment report**

<b>Trial</b>	<b>Reports</b>
Identification (ID) of trial used in remainder of assessment report	Internal study report title. Date. Author(s). Title. <i>Journal</i> Year; Vol(No):pages Author(s). Title. <i>Journal</i> Year; Vol(No):pages

If there are no direct trials, see Figure B1 for the next step in the clinical evaluation.

### **Option to present supplementary randomised trial data**

Where data from one or more direct randomised trials are available, the presentation of an indirect comparison is generally not encouraged. However, in certain circumstances, it might be reasonable to justify the inclusion of supplementary randomised trial data. The following list shows possible situations where this might apply:

- A supplementary indirect comparison of two or more sets of trials involving one or more common references that is based on much larger participant numbers (particularly if the direct randomised trials available are underpowered overall); see also Part III, Section B(i) for further guidance on presenting an indirect comparison.
- A meta-analysis comparing all trials of the proposed therapeutic medical service against several therapeutic medical services widely accepted as equivalent to the main comparator in terms of effectiveness and safety, as well as the direct randomised trials.
- A meta-analysis comparing all trials of the main comparator against several therapeutic medical services widely accepted as equivalent to the proposed therapeutic medical service in terms of effectiveness and safety, as well as the direct randomised trials.

Separately identify and list the supplementary randomised trials as part of the response to Sub-section B2 and include reports of these trials with other references to the assessment report. Present these supplementary trials in Sub-sections B3–B6. Clearly label this supplementary information to distinguish it from the information from the relevant direct randomised trial(s).

### **Meta-analyses**

Separately identify any meta-analysis of randomised trials from the suite of searches above and assess their exclusion or inclusion using the criteria above. This should include any relevant systematic reviews from the Cochrane Database of Systematic Reviews.

If a published meta-analysis of direct randomised trials is the principal source of the presented clinical evaluation, provide a copy of the publication as an attachment in the main body of the assessment report. Assess whether the published meta-analysis has a well-defined clinical question relevant to the intended listing of the proposed therapeutic medical service, a reproducible literature search strategy and appropriate criteria for any exclusions of identified direct randomised trials. Assess the meta-analysis using the framework of this Sub-section alongside the presentation of the individual trials. Where there is more than one such meta-analysis, tabulate these assessments.

## Exclusion of trials

Justify the exclusion of any direct randomised trial included in the master list in Table B2.3 from further detailed assessment in the assessment report. The grounds for exclusion might include any aspect reported in Sub-sections B3–B5 (i.e. the quality of the trials, the patient characteristics and circumstances of use, and the outcomes reported in the trials). This might minimise observable differences across the randomised trials, or examine and explain, where possible, their contribution to heterogeneity across all the trials.

It is not possible to give unequivocal guidance on the exclusion of direct randomised trials at this stage. If a decision to exclude or include one or more randomised trials is likely to be controversial, it is usually wiser to also present a sensitivity analysis examining whether that decision makes a difference to the conclusions from the overall clinical evaluation.

If one or more trials are to be excluded, identify those aspects of each trial that result in the exclusion (see Table B2.4). Indicate whether each reason relates to the quality of the trials, the patient characteristics and circumstances of use, and/or the outcomes reported in the trials. Present greater detail of each aspect (as a minimum, to the extent requested in the relevant text below in Sub-sections B3–B5). If there is more than one type of reason for exclusion, arrange the excluded trials in Table B2.4 by the reason for exclusion.

**Table B2.4 Reasons to exclude each trial from further detailed assessment**

Trial ID	Ground(s) for seeking exclusion	Details <sup>a</sup>
<b>Quality of the trial</b>		
Trial A		
Etc.		
<b>Patient characteristics and circumstances of use in the trial</b>		
Trial B		
Etc.		
<b>Outcomes reported in the trial</b>		
Etc.		

ID = identification

<sup>a</sup> Cross-reference each set of details to the source of information (specifying the trial report with page, table, figure number).

Tabulate a brief summary highlighting key aspects of the identified trials, presenting included and then excluded trials (see Tables B2.5 and B2.6).

**Table B2.5 Comparative summary of characteristics of each direct randomised trial**

Trial ID	Design characteristics <sup>a</sup>	Compared interventions (N, service <sup>b</sup> )	Summary of main population characteristics	Main outcomes	
				Primary	Secondary
Included trials					
Trial 1					
Etc.					
Excluded trials					
Trial A					
Etc.					

ID = identification

<sup>a</sup> C-O = cross-over; DB = double-blind; ITT = intention to treat; PG = parallel group; SB = single-blind

<sup>b</sup> include dose, frequency, duration of service as relevant

**Table B2.6 Comparative summary of results of each direct randomised trial**

Trial ID	Primary outcome (95% CI)	Secondary outcomes (95% CI)	Major adverse events
<b>Included trials</b>			
Trial 1			
Etc.			
<b>Excluded trials</b>			
Trial A			
Etc.			

CI = confidence interval; ID = identification

### **Presentation of non-inferiority (equivalence) trials**

Most randomised trials are designed to show a difference between the compared therapeutic medical services. If any direct randomised trial was designed prospectively as a non-inferiority trial, and/or the therapeutic conclusion presented in Sub-section B8 is non-inferiority or equivalence, refer to the additional guidance on presenting the direct randomised trial in Appendix 3.

Non-inferiority means that, in terms of effectiveness, the proposed therapeutic medical service is no worse than its main comparator. It is used to support a claim of equivalence, because it is not adequate to demonstrate the *absence* of a statistically significant difference between the treatments to claim equivalence; such a lack of a significant difference might occur when the trials are too small to demonstrate a real difference in the effects of the interventions. The appropriate comparison to present is the point estimate of the difference with its 95% confidence interval. This allows MSAC to assess whether the confidence interval contains the minimal clinically important difference.

### **Trial details**

Include sufficient details of the relevant randomised trials as attachments in the main body of the assessment report. Where there is more than one report of a randomised trial (e.g. one or more published papers) and one or more trial reports internal to the applicant, provide both the published paper(s) and key extracts from the applicant's internal trial report. The results might vary between reports of the same randomised trial. If so, justify and cross-reference the selection of the source of results extracted for the assessment report. Provide a copy of each other publication reporting data from a listed randomised trial. Ensure that the assessment report index shows the location of all submitted papers, both in the main body of the assessment report and in the attachments.

For any relevant trial identified from a meta-analysis, include the individual trial report or publication(s) as above. If no separate report is available, indicate the efforts made to retrieve them and to obtain any missing information from the authors of the published meta-analysis.

Provide reputable translations of trial reports printed in other languages.

## B3 Assessment of the measures taken by investigators to minimise bias

### INFORMATION REQUEST

- For each direct randomised trial listed, provide information on the measures taken to minimise bias, using the checklist provided.
- For each checklist response, specify the source document in the reports or papers accompanying the main body of the assessment report, together with the page or table from which the information was extracted

Assessing measures of minimising bias provides the applicant and MSAC with a clear idea of which trials are of greater scientific rigour. There is no minimum standard, but MSAC is most likely to be persuaded by the data of the highest scientific rigour.

The checklist in Box B3.1 includes three sets of methodological topics that help to assess the methodological quality of each trial. Table B3.1 shows a suggested approach to presenting the information in a summary format. This is a useful guide to help MSAC and the applicant review the scientific rigour of the evidence by assessing the measures taken by the investigators to minimise bias. It is not intended to discourage the presentation of data.

#### Box B3.1 Checklist for assessing the quality (internal validity) of randomised trials

Methodological topic	Quality issue
(a) Randomisation	(i) How was the randomisation sequence concealed during the allocation process?
(b) Blinding	(i) Were the following groups blinded to the treatment allocation? 1. Trial participants 2. Investigators 3. Personnel assessing the outcomes (ii) If any of the groups in (b)(i) were blinded to treatment allocation, how was blinding achieved? (iii) If any of the groups in (b)(i) were <i>not</i> blinded to treatment allocation, why was blinding not possible?
(c) Follow-up	(i) What was the basis of the analysis? (ii) How many participants were randomised to each arm of the trial? (iii) How many participants in each arm of the trial did not receive the allocated intervention? (iv) How many participants in each arm of the trial were lost to follow-up? (v) How many participants in each arm of the trial discontinued the intervention? (vi) How many participants in each arm of the trial contributed data to the primary analysis?

#### Notes for trial quality checklist

##### (a) Randomisation

Randomisation distributes both known and unknown confounders by the play of chance, providing a good basis for comparison between randomised groups in a treatment trial because the groups differ only by the treatment allocation and the play of chance. Statistical methods then help determine whether observed differences can credibly be attributed to the treatment(s) under investigation rather than to chance. Secure

randomisation minimises selection bias. To ensure that randomisation remains secure, it is important that the personnel responsible for enrolling participants into a trial are unable to predict which treatment a participant would receive before a final decision is made regarding entry to the trial. Provide details of the methods of concealing the randomisation sequence, such as decentralised or ‘third party’ assignment, or sequentially numbered envelopes or containers.

### **(b) Blinding**

Blinding of participants, investigators or those responsible for assessing the outcomes helps prevent several important biases in randomised trials. Blinding of participants and investigators might influence several aspects of the trial, including the response to treatments, the use of co-interventions and withdrawal rates from the trial. Blinding of outcome assessors might also influence the reported response to treatment. The influence of blinding is most important where the outcome is subjective, such as the evaluation of pain or preference of treatment.

If blinding of treatment allocation was used, describe the methods used, such as identical tablets or capsules. Blinding of treatment allocation might not always be possible; for example, in a comparison between an open and laparoscopic surgery. Where the comparator is distinguishable by some means or there is a high chance of ‘unblinding’ (e.g. some sham surgical procedures), it is important that the observer responsible for measuring the trial outcomes remains unaware of the treatment assignment. State the reasons for not blinding the participants, investigator(s) or outcome assessors. Discuss the effect, if any, that the absence of blinding might have had on the measurement of the primary and secondary outcomes of the trial.

### **(c) Follow-up**

Follow-up is important, and it is also important that an attempt is made to summarise the trial outcomes for all participants. A full intention-to-treat (ITT) analysis is preferred for trials designed to demonstrate a therapeutic difference (and related incremental cost-effectiveness analysis) to minimise bias in the follow-up of participants. Specify how the ITT analysis dealt with missing data.

### **Tabulate responses**

If there is more than one trial, tabulate the responses in the main body of the assessment report, with the detailed responses to the above questions in an accompanying attachment or technical document. In this detailed presentation, also provide adequate cross-references to the trial report (including page, table or figure numbers of the source document) from which each aspect of the information was extracted.

Tables B3.1 and B3.2 provide a suggested format for the presentation of the summary in the main body of the assessment report.

**Table B3.1 Summary of the measures undertaken to minimise bias in the direct randomised trials**

Trial ID	Concealment of randomisation <sup>a</sup>	Blinding			Basis of analysis <sup>b</sup>
		Participants	Investigators	Outcomes assessors	
Trial 1	A/B/C/None	Yes/No	Yes/No	Yes/No	D/E/F/G
Trial 2	A/B/C/None	Yes/No	Yes/No	Yes/No	D/E/F/G
Etc.					

ID = identification

**a** A = central telephone randomisation service; B = third-party randomisation service; C = sequentially labelled, fully opaque, sealed envelopes

**b** D = intention-to-treat (all randomised participants, specify how the analysis dealt with missing data); E = all treated participants (specify how the analysis dealt with missing data); F = per protocol participants; G = other (specified)

**Table B3.2 Flow of participants through the direct randomised trials**

Trial ID • Intervention arm	No. randomised	Did not receive intervention	Lost to follow-up	Dis-continued	Analysed	Source of information
Trial 1 • Proposed therapeutic medical service	<i>N</i>	<i>n</i> (%)	<i>n</i> (%)	<i>n</i> (%)	<i>n</i> (%)	(Add this column to tables and submit in a separate technical attachment)
• Main comparator	<i>N</i>	<i>n</i> (%)	<i>n</i> (%)	<i>n</i> (%)	<i>n</i> (%)	
Trial 2 • Proposed therapeutic medical service	<i>N</i>	<i>n</i> (%)	<i>n</i> (%)	<i>n</i> (%)	<i>n</i> (%)	
• Main comparator	<i>N</i>	<i>n</i> (%)	<i>n</i> (%)	<i>n</i> (%)	<i>n</i> (%)	
Etc.						

ID = identification; *n* = number of participants with event; *N* = total participants in group

## Source documents

For each of the responses provided in Tables B3.1 and B3.2, specify the source document in the reports or papers accompanying the main body of the assessment report. Provide adequate detail of cross-referencing to page, table or figure number of the relevant trial report(s) in a way that does not detract from the presentation of the requested results.

For the presentation of a complex systematic overview, consider re-presenting the tables from the main body of the assessment report in a technical document or attachment, and add an additional column to each table to provide adequate detail of cross-referencing (as illustrated by the shaded column in Table B3.2). Alternatively, if it is clearer for some tables, identify the source of information cell by cell, using footnotes.



## B4 Characteristics of the trials

### INFORMATION REQUEST

- For each direct randomised trial, provide the following details of the trial protocols and participants:
  - the eligibility criteria for participants considered for recruitment into the trial;
  - the baseline demographic and clinical characteristics of each randomised group; and
  - the duration of follow-up (median and range) and whether the trial has been completed or is ongoing.
- For each response, specify the source document in the reports or papers accompanying the main body of the assessment report, together with the page or table from which the information was extracted.

### Details of trials

If there is more than one direct randomised trial, tabulate the responses in the main body of the assessment report. Tables B4.1–B4.3 provide a suggested format.

**Table B4.1 Eligibility criteria**

Trial ID	Inclusion criteria	Exclusion criteria
Trial 1	Typical inclusion criteria might relate to age, sex and clinical diagnosis.	Exclusion criteria are often used to ensure participant safety.
Trial 2		
Etc.		

Indicate any significant differences in the baseline characteristics of randomised groups across the trials and discuss any impact this might have on the interpretation of the trial results, including those to be examined in Sub-section C1. Table B4.2 provides a suggested format for this information.

**Table B4.2 Characteristics of participants across randomised groups**

Trial ID Baseline characteristic	First randomised group	Second randomised group	Third randomised group	Etc.
Trial 1				
Age				
Sex (etc.)				
Trial 2				
Age				
Sex (etc.)				
Etc.				

ID = identification

**Table B4.3 Therapeutic medical services compared in trials**

<b>Trial ID</b>	<b>Treatment</b>	<b>Treatment regimen</b>	<b>Duration of treatment</b>	<b>Duration of follow-up</b>
Trial 1	Proposed therapeutic medical service		Median (range)	Median (range)
	Main comparator		Median (range)	Median (range)
Trial 2	Proposed therapeutic medical service		Median (range)	Median (range)
	Main comparator		Median (range)	Median (range)
Etc.				

ID = identification

Present a separate table for any cross-over randomised trials (to report additional details such as a period of wash-out between treatment periods) and indicate how the results of the cross-over have been included in the systematic overview (see Sub-section B6).

Provide any additional information about the trial or participant characteristics that is not requested elsewhere in Sub-sections B3–B5, but that is relied on to assess the applicability of the direct randomised trial evidence to the listing requested (see Sub-section C1). For example, if it is considered that the settings and locations where the interventions were provided modify the treatment effect, summarise the details of this characteristic across all the trials and cross-reference to Sub-section C1.

If the proposal is to limit the use of the proposed therapeutic medical service to a last line of therapy so that placebo for standard medical management is the nominated main comparator, identify whether the participants in the direct randomised trials reflected a similar positioning in the clinical management algorithm. If the trials recruited participants earlier in the clinical management algorithm, discuss the implications for the assessment report.

### **Source documents**

For each of the responses provided above, provide adequate cross-referencing to page, table or figure numbers of the relevant trial report(s), if necessary in a separate technical document or attachment, as described for Sub-section B3.

## **B5 Outcome measures and analysis**

### **INFORMATION REQUEST**

- For each direct randomised trial, describe the primary outcome and how it was analysed.
- For each direct randomised trial, describe the patient-relevant secondary outcomes (including any quality-of-life outcomes) and how they were analysed.
- Discuss the clinical importance of the primary outcome and secondary outcomes listed in response to the requests above.
- Assess each instrument used to measure quality of life.
- For each direct randomised trial, indicate whether a multi-attribute utility instrument (MAUI) was used and, if so, how it was used and how its results were analysed.
- For each response, specify the source document in the reports or papers accompanying the main body of the assessment report, together with the page or table from which the information was extracted.

## Primary outcomes

List and clearly define the primary outcome measure for each direct randomised trial, including its units of measurement. Specify enough details of the outcome measurement for MSAC to assess its clinical importance (e.g. supine/erect blood pressure). State the difference specified as worth detecting in any power calculation. For each primary outcome, describe the statistical methods used in the primary analysis to compare across the randomised groups. State whether the primary outcome was assessed at several time points after randomisation. If so, indicate the pre-specified time point of the primary analysis and describe the methods of adjusting for multiple interim analyses.

Table B5.1 provides a suggested format for presenting and comparing primary outcomes from several trials.

**Table B5.1 Primary outcomes and statistical analyses of the direct randomised trials**

Trial ID	Definition of primary outcome	Method of primary statistical analysis
Trial 1		
Trial 2		
Etc.		

ID = identification

Ensure that each primary outcome is reported as being truly independent, or that the statistical analysis appropriately adjusts for clustering. This issue has most often occurred where a single patient can experience multiple events (e.g. fractures, hypoglycaemic events, hospitalisation episodes) during the follow-up of the trial.

## Secondary outcomes

For each direct randomised trial, list and define each secondary outcome and analysis that is patient relevant, including the units of measurement. This might include secondary analyses of the primary outcome. Include any data collected for resources provided (economic outcomes) as well as health outcomes gained, because they are relevant both to patients and the economic evaluation. For each patient-relevant secondary outcome, describe the statistical methods used to compare across randomised groups. State the number of pre-specified secondary outcomes and any methods used to address the multiplicity of analyses across outcomes. Increasing the number of multiple comparisons increases the odds that, through chance alone, a statistically significant difference will emerge in one of these comparisons, assuming the null hypothesis is true.

Patient-relevant outcomes are those outcomes that are perceptible to the patient; the more important the outcome is to the patient, the more relevant it becomes. Examples of patient-relevant outcomes include quality-of-life measures, preference weights (see Appendix 4), and economic inputs and outcomes (see Sub-section D4).

Table B5.2 provides a suggested format for presenting and comparing patient-relevant secondary outcomes and analyses when more than one trial is included in the assessment report.

**Table B5.2 Patient-relevant secondary outcomes and analyses**

Trial ID	Definition of secondary outcome	Method of statistical analysis
Trial 1		
Trial 2		
Etc.		

ID = identification

For the primary outcomes, ensure that each outcome is reported as being truly independent, or that the statistical analysis appropriately adjusts for clustering.

### **Clinical importance**

Discuss the clinical importance of the primary outcome and secondary outcomes listed in Tables B5.1 and B5.2. For primary outcomes, this might be informed by the basis given in the trial protocol for the minimal clinically important difference used in the power calculation. Discuss clinical importance in terms of both relative and absolute changes.

### **Composite outcomes**

If one or more of the reported outcomes is a composite outcome, discuss and compare the clinical importance of each of its component outcomes. Report whether the definition of the composite outcome was pre-specified explicitly. Explain the justification, and provide the key literature for the inclusion of the components in the composite outcome and for the exclusion of any components that were considered but rejected as components in the composite outcome. Disaggregate the composite outcome to present the results (e.g. comparative rates) of each component as a secondary outcome in Sub-section B6. To avoid double-counting, a composite outcome is usually defined as having happened when the trial participant experiences the first component outcome in the composite (such as disease progression), even though other component outcomes in the composite (such as death) might be subsequently experienced. This needs to be appropriately handled in disaggregating the composite outcome so that, where possible, all subsequent first experiences of any other component outcome in the composite are also included.

### **Quality-of-life instruments**

Consider using a quality-of-life measure where a change in quality of life is the principal intended final outcome (see Sub-section D4). This is true for some indications (e.g. relief of pain, treatment of depression, treatment of some cancers) in which improved quality of life is the principal aim of therapy. Alternatively, quality of life might be impaired by the proposed therapeutic medical service or by its main comparator (or other intervention). Quality-of-life measures might supplement other clinical measures.

Quality-of-life instruments include generic ('global') health-related quality-of-life scales and disease-specific rating scales (e.g. for pain, disability or depression), which might themselves be the primary measure of outcome in the trials. Increasingly, trials are collecting data using both types of quality-of-life instruments.

Where a quality-of-life instrument is used, details should be provided on the instrument. Controversy remains about which quality-of-life instruments are most acceptable, so special attention should be paid to the following parameters:

- the validity of the instrument;

- the reliability of the instrument;
- the responsiveness of the instrument to differences in health states between individuals and to changes in health states over time experienced by any one individual; and
- the clinical importance of any differences detected by the instrument.

Where possible, provide any supportive data and references assessing these parameters of the instrument in a technical document or an attachment to the assessment report (provide clear cross-references between these data and the main body of the assessment report).

For therapeutic medical services that cure or prevent short-term illnesses (e.g. infections), outcomes might not always be measurable on a quality-of-life instrument. It might also be reasonable to assume that certain events that might themselves be serious do not greatly impair quality of life in survivors (e.g. pneumonia).

### **Use of a multi-attribute utility instrument**

Appendix 4 describes the use of health-related quality-adjusted life-years (QALYs) gained as a measure of health outcomes that is comparable across health states. It also provides background information on the generally preferred method of measuring QALYs, which is through the repeated application of a valid, reliable and responsive MAUI questionnaire to participants in a randomised double-blind trial, together with the application of an appropriate scoring algorithm (see Sub-section B6).

The MAUI should be used to collect information from trial participants at baseline, and at one or more time points during the trial follow-up (see advantages of relying on the trial based MAUI data (h) in Appendix 4).

Because health-related quality of life is inherently subjective, its assessment in a randomised trial as a basis for then estimating utility weights using a MAUI algorithm is more persuasive if the trial design blinded the observers of the outcome being measured to the treatment assigned (see Sub-section B3 and advantages of relying on the trial based MAUI data (c) in Appendix 4).

Acceptable MAUIs are the Health Utilities Index (HUI2 or HUI3), the EQ5D ('EuroQol'), the SF-6D (a subset of the Short Form 36, or SF-36) or the Assessment of Quality of Life (AQoL) instrument. Currently, there is insufficient basis for a preference to be expressed between these MAUIs. All are based on acceptable scaling techniques of the standard gamble (SG) or time trade-off (TTO), and some have different scoring algorithms for different countries. Studies directly comparing these MAUIs suggest that each MAUI yields different results for the same health state, so their utility weight results cannot be compared with complete confidence. The MAUIs listed above vary in their coverage of important health domains, but they all cover the main areas of health-related quality of life that patients would be willing to trade for increased survival. HUI2 is designed for use in childhood conditions.

All the MAUIs have strengths and weaknesses. For example, as a general observation, the EQ5D has fewer possible health states, which means that it has been perceived as relatively unresponsive or insensitive compared with the other MAUIs listed above. Another feature of the EQ5D is that when a difference is detected, the numerical value can appear disproportionately large compared with the more gradual increments of the other MAUIs listed above.

The use of any other possible preference-based instrument, such as the Quality of Well-Being Scale (QWB) or the 15 Dimensions (15D), needs to be particularly justified, including with reference to the above criteria of comparability, acceptable scaling techniques and responsiveness.

If a MAUI has been used in a relevant randomised trial for reporting utility weights, provide details of the selected MAUI. Justify the selection of any MAUI used in the trial, but not listed above as acceptable by assessing:

- the validity of the instrument;
- the reliability of the instrument;
- the responsiveness or sensitivity of the instrument to differences in health states between individuals who are likely to be affected by the proposed therapeutic medical service and its main comparator;
- the responsiveness or sensitivity of the instrument to changes in health states over time experienced by any one individual;
- the duration of the period assessed when responding to the MAUI questionnaire compared with the duration of the condition of interest; and
- the applicability to the general Australian population of the scoring algorithm applied to the responses reported with the MAUI questionnaire to calculate utility weights.

Include any data and references that support the selection of the MAUI in a technical document or an attachment to the assessment report (provide clear cross-references between these data and the main body of the assessment report).

### **Source documents**

For each of the responses provided above, provide adequate cross-referencing to page, table or figure numbers of the relevant trial report(s), if necessary in a separate technical document or attachment, as described in Sub-section B3.

## **B6 Systematic overview of the results**

### **INFORMATION REQUEST**

- For each direct randomised trial, present the results of the primary analysis for that trial.
- Present an analysis of the results for each type of patient-relevant outcome in terms of its natural units in tables with graphed forest plots. Include results reporting quality-of-life outcomes.
- Where there is more than one randomised trial reporting a particular outcome, statistically combine (meta-analyse) the results.
- For each meta-analysis of each outcome, assess the potential for outcomes reporting bias by reporting (in a footnote) on the presentation of the forest plot for each outcome:
  - the number of trials contributing to the forest plot ; and
  - the proportion of these trials over the total number of trials included in Table B2.3.
- Present the results of any MAUI used in any of the direct randomised trials.
- For each response, specify the source document in the reports or papers accompanying the main body of the assessment report, together with the page or table from which the information was extracted.

The presentation of the trial results in Sub-section B6 serves two purposes:

- the presentation of the results of the primary analyses as established for each direct randomised trial is part of the assessment of the scientific rigour of the trial dataset and becomes a reference point for interpreting other patient-relevant outcomes for that trial; and
- the presentation of the results of common outcomes across more than one trial enables an assessment to be made of the comparative effectiveness of the proposed therapeutic medical service and the main comparator under the circumstances of the trials as designed and conducted.

Sub-section B6 is not directly concerned with the application of the available trial evidence to the listing requested. Section C addresses this important issue.

### **Primary analysis**

For each direct randomised trial listed in Sub-section B2, present the results for the primary outcome according to the design of the pre-specified primary analysis for that trial. Justify and discuss any early stopping of a trial or reliance on any interim analysis in the interpretation of the primary outcome. Analyse (including meta-analyse, if necessary) all patient-relevant outcomes.

Present a meta-analysis for each patient-relevant outcome listed in Sub-section B5 (which might include one or more primary outcomes). First, present the results (preferably analysed on an ITT basis) for each randomised group of each randomised trial listed in Sub-section B2 reporting that particular outcome. Then present the measured direction and the magnitude of the treatment effect across groups of each trial (also preferably analysed on an ITT basis).

Guidance is provided below on the preferred method of displaying results, depending on the way the data are reported (see also Tables B6.1–B6.5).

Where there is more than one randomised trial reporting a particular outcome, the presentation of a meta-analysis, which statistically combines (pools) results across trials, is generally preferred, where appropriate. Collate the results of each trial reporting the outcome into a meta-analysis and present the results of each meta-analysis in a table and as a graphed forest plot, including the pooled results across the trials. As an example, the software from the Cochrane Collaboration quickly and succinctly conveys the requested array of meta-analysed information in a format suitable for including in the main body of the assessment report.

Where a meta-analysis is based on a subset of all available direct randomised trials, identify the trials in the subset. Report the number of trials in the subset and the proportion that this number represents of the total number of trials listed in Sub-section B2. This includes situations where there is only one randomised trial reporting a particular patient-relevant outcome; in this case, the number of trials in the subset is one and there is no basis to meta-analyse the data any further. Examine whether there are any differences between the results of the subset and the total set of trials using group-level data, and assess the impact of any bias (such as outcomes reporting bias) across any differences detected.

Meta-analysis is useful because it might increase the precision of the estimates of differences between the proposed therapeutic medical service and the main comparator. It is also useful when there are conflicting results from trials of similar scientific rigour. Meta-analysis can also highlight advantages of a proposed therapeutic medical service that are too small to be detected reliably in individual randomised trials, but might be clinically important. Justify any decision *not* to present a meta-analysis whenever there is more than one direct randomised trial reporting a common, patient-relevant outcome.

Explain and justify the methods used for statistically combining cross-over trials in a meta-analysis of parallel group trials. Clearly document and reference the methods used to make them independently reproducible and verifiable.

Where a meta-analysis of group-level data is supplemented by individual patient data, provide an appropriate summary of these data for each trial and for the pooled results overall. Where individual patient data are used in a pooled analysis, ensure that the trial in which each individual was randomised is included as a covariate in the analysis.

Explain and justify any other method used for statistically combining the results of the direct randomised trials and any additional statistical tests used. Clearly document and reference the methods used to make them independently reproducible and verifiable. Provide adequate detail of all sources of information relied on for these other analyses (see Part I, Section 5), and then present their results.

## **Dichotomous data**

For each outcome measured as dichotomous data (e.g. with or without the event), present for each group in each trial:

- the number with the event;
- the number in the group;
- the percentage with the event; and
- the period of time after randomisation at which these data were collected in the trial (which is usually the median duration of follow-up).

Then present the relative risk, risk difference and number needed to treat (NNT) with their associated 95% confidence intervals for each trial reporting the outcome.

Where there is more than one randomised trial reporting a particular dichotomous outcome, tabulate the results (point estimates and 95% confidence intervals) of the individual trials as the relative risk and the risk difference. Also present these results for the individual trials on a graphed forest plot.

Statistically combine the results for the relative risk and risk difference using the DerSimonian–Laird random effects model and include the pooled results in each table and graphed forest plot, together with their associated 95% confidence intervals.

Report results for statistical heterogeneity as the Cochran  $Q$  with a chi-square test for heterogeneity and the  $I^2$  statistic with its 95% uncertainty interval. If heterogeneity is present, consider examining it in Section C of the assessment report.

Tables B6.1 and B6.2 provide a suggested format for presenting and comparing dichotomous outcome data from several trials.



**Table B6.1 Results of [patient-relevant outcome] (available as dichotomous data) across the direct randomised trials (relative risk)**

Trial ID	Proposed therapeutic medical service	Main comparator	Forest plot here	Relative risk (95% CI)
Trial 1	<i>n</i> with event/ <i>N</i> (%)	<i>n</i> with event/ <i>N</i> (%)		
Trial 2				
Etc.				
Pooled result from random effects model				
Chi-square ( <i>Q</i> ) for heterogeneity: <i>P</i> = <i>I</i> <sup>2</sup> statistic with 95% uncertainty interval =				

CI = confidence interval; ID = identification; *n* = number of participants with event; *N* = total participants in group

**Note:** Provide number and percentage of the identified relevant direct randomised trials that contributed data to this meta-analysis.

**Table B6.2 Results of [patient-relevant outcome] (available as dichotomous data) across the direct randomised trials (risk difference)**

Trial ID	Proposed therapeutic medical service	Main comparator	Forest plot here	Risk difference (95% CI)
Trial 1	<i>n</i> with event/ <i>N</i> (%)	<i>n</i> with event/ <i>N</i> (%)		
Trial 2				
Etc.				
Pooled result from random effects model				
Chi-square ( <i>Q</i> ) for heterogeneity: <i>P</i> = <i>I</i> <sup>2</sup> statistic with 95% uncertainty interval =				

CI = confidence interval; ID = identification; *n* = number of participants with event; *N* = total participants in group

**Note:** Provide number and percentage of the identified relevant direct randomised trials that contributed data to this meta-analysis.

### Continuous data

For each outcome measured as continuous data, present for each group in each trial the mean at baseline, the mean at end point (or other justified time point) and the mean change, each with its standard deviation. Then present, for each trial reporting the outcome, the mean difference at end point and the mean difference of the change, each with its 95% confidence interval. Report the number of participants in each randomised group of the trial contributing data to each analysis of a continuous outcome.

Where there is more than one trial, tabulate the results (point estimates and 95% confidence intervals) of the individual trials. On a graphed forest plot, plot the results (point estimates and 95% confidence intervals) of the individual trials as the weighted mean difference at end point and the weighted mean difference of the change.

Statistically combine the results for the weighted mean difference using the DerSimonian–Laird random effects model and include the pooled result in each table and graphed forest plot, together with its associated 95% confidence interval.

Report results for statistical heterogeneity as the Cochran *Q* with a chi-square test for heterogeneity and the *I*<sup>2</sup> statistic with its 95% uncertainty interval. If heterogeneity is present, consider examining it in Section C of the assessment report.

Tables B6.3 and B6.4 provide a suggested format for presenting and comparing continuous outcomes data from several trials.

**Table B6.3 Results of [patient-relevant outcome] (available as continuous data) across the direct randomised trials (end point)**

Trial ID	Proposed therapeutic medical service		Main comparator		Forest plot here	Mean difference (95% CI)
	<i>n</i> reporting data/ <i>N</i> (%)	End point <sup>a</sup> mean (SD)	<i>n</i> reporting data/ <i>N</i> (%)	End point <sup>a</sup> mean (SD)		End point <sup>a</sup>
Trial 1						
Trial 2						
Etc.						
Pooled result from random effects model						
Chi-square (Q) for heterogeneity: <i>P</i> = <i>I</i> <sup>2</sup> statistic with 95% uncertainty interval =						

CI = confidence interval; ID = identification; *n* = number of participants reporting data; *N* = total participants in group; SD = standard deviation

<sup>a</sup> Or other justified time point

**Note:** Provide number and percentage of the identified relevant direct randomised trials that contributed data to this meta-analysis.

**Table B6.4 Results of [patient-relevant outcome] (available as continuous data) across the direct randomised trials (change)**

Trial ID	Proposed therapeutic medical service (mean values)			Main comparator (mean values)			Forest plot here	Mean difference (95% CI)
	Base-line (SD)	End point <sup>a</sup> (SD)	Change (SD)	Base-line (SD)	End point <sup>a</sup> (SD)	Change (SD)		Change
Trial 1								
Trial 2								
Etc.								
Pooled result from random effects model								
Chi-square (Q) for heterogeneity: <i>P</i> = <i>I</i> <sup>2</sup> statistic with 95% uncertainty interval =								

CI = confidence interval; ID = identification; SD = standard deviation

<sup>a</sup> Or other justified time point

**Note:** Provide number and percentage of the identified relevant direct randomised trials that contributed data to this meta-analysis.

## Ordinal or categorical data

A similar approach to the above for continuous data should be attempted if the trial results are available as ordinal or categorical data (e.g. a Likert scale reporting quality-of-life data). Expert biostatistical advice will be helpful in such circumstances, particularly to meta-analyse such data.

## Time-to-event data

Whenever time-to-event data are reported for the overall population in a direct randomised trial, present a graphical plot of the relevant Kaplan–Meier curves (if necessary, reproduce the graphical plot directly from the cited work, because these data are only reported in a published citation).

Present a separate graphical plot for each such trial and for each time-to-event outcome, displaying a separate curve for each randomised group, preferably on an ITT basis. On each graphical plot, also display the median duration of follow-up and the remaining sample size for each curve at each of a series of time points along the x-axis. Analyse

differences between event curves using the logrank test. If the Wilcoxon signed-rank test is also presented, justify why it is appropriate (e.g. because of its emphasis on early event times).

Where the analysis is based on a Cox proportional hazards model, present the hazard ratios, together with their 95% confidence intervals. Discuss whether the results are consistent with the assumption of constant proportional hazards.

In the analysis of time-to-event data from the direct randomised trials, censoring usually precludes the estimation of a mean time-to-event. Thus, for any trial reporting time-to-event data where the trial follow-up is insufficient to record all events, the result is a restricted or truncated time-to-event analysis. If the integrals between the two truncated Kaplan–Meier curves are compared, the result is a difference in the truncated means. Therefore, present differences in times-to-event as comparisons of medians (where possible) and of truncated means (with their 95% confidence intervals), with the latter preferably calculated both:

- from the beginning of the trial to the end of the most recent available follow-up of the trial; and
- for the median duration of follow-up across the trial population, where follow-up for each individual is defined to be the duration of time from the date of randomisation to the date of the clinical cut-off (for a completed trial) or to the date of the most recent data snapshot (for an ongoing trial). Assuming a constant rate of accrual into the trial, a similar duration can be estimated as being from the start of the trial to time  $t$ , where  $t$  occurs at a point in time equivalent to half the accrual period before the most recent available follow-up of the trial.

Where there is more than one randomised trial reporting a particular time-to-event outcome, present the pooled results across the trials, together with the number of trials contributing to the forest plot and the proportion of those trials over the total number of trials included in the assessment report. Data from multiple trials involving a particular time-to-event outcome might be statistically combined in a number of ways. Justify and reference the method(s) selected for pooling time-to-event data. Specify and describe this method in a short technical document or attachment to the assessment report and provide sufficient data to allow the results to be reproduced and verified independently (see Part I, Section 5).

The preferred method would be to pool individual patient data from a Cox proportional hazards model, with the pooling method including the trial as a covariate. If individual patient data are not available, then pool the hazard ratios from the trial-level data to present the pooled hazard ratio with its 95% confidence interval. If hazard ratios with their standard errors are not all available, it might be possible to pool dichotomised data based on a common duration of follow-up. Expert biostatistical advice will be helpful for pooling the integral between Kaplan–Meier curves.

Table B6.5 provides a suggested format for presenting and comparing time-to-event outcomes from several trials.

**Table B6.5 Results of [patient-relevant outcome] across the direct randomised trials (available as time-to-event data)**

Trial ID	Hazard ratio (95% CI)	Logrank test <i>P</i> -value
Trial 1		
Trial 2		
Etc.		
Pooled		Not applicable

CI = confidence interval; ID = identification

**Note:** Provide number and percentage of the identified relevant direct randomised trials that contributed data to this meta-analysis.

The assessor should assess the evidence by tabulating the *P*-values for each study. It is important to consider the validity of the statistical approaches used to obtain each *P*-value. Marginally significant *P*-values can be difficult to interpret, and particular consideration needs to be given to the sample size of the study and the number of statistical comparisons that have been carried out in the analysis.

### Adverse event data

As a minimum, report important adverse events as the number of participants reporting:

- any adverse event;
- any adverse event resulting in discontinuation of the randomised treatment;
- any adverse event resulting in hospitalisation;
- any adverse event resulting in death; and
- each and every other type of adverse event where the frequency or severity differs substantially across randomised groups, for each randomised trial listed in Sub-section B2, preferably on an ITT basis.

For each important adverse event, present these results based on proportions of participants reporting each type of adverse event (i.e. as for dichotomous data above), therefore also presenting relative risks and risk differences with their 95% confidence intervals across the randomised groups for each trial separately. In addition, where appropriate, pool these results across all trials using the random effects model. Where the average period at risk per participant varies substantially between treatment groups, the relative adverse event rates (events/period-at-risk) should also be analysed using Poisson regression, with pooling across trials as necessary using the random effects model. See Sub-section B7 for further discussion of adverse outcomes reported from other sources.

### Present the results of a multi-attribute utility instrument

Ideally, report MAUI results as the difference (with 95% confidence interval) in the integrals between the mean utility weights obtained over time up to the median period of follow-up in the trial for the proposed therapeutic medical service and its main comparator. This directly estimates the incremental QALYs gained. Also report the results analysed as specified in the trial protocol, particularly if the difference between integrals cannot be generated directly.

Ideally, the scoring algorithm of the acceptable MAUIs listed in Sub-section B5 would be derived from the general population in Australia (see advantages of relying on the trial based MAUI data (e) in Appendix 4), because this would assist in generating Australia-specific utility weights from responses to the MAUI questionnaires generated in international trials. However, there are few Australian-based scoring algorithms for MAUIs generated from an appropriately defined population sample and, in the absence of these, it might be justifiable to use scoring algorithms from other countries with similar cultural or political backgrounds and economic circumstances (e.g. Canada and England). Where more than one scoring algorithm exists for a MAUI questionnaire, but no Australian scoring algorithm, consider presenting an analysis to examine the sensitivity of the trial results to using different scoring algorithms. Similarly, if more than one MAUI questionnaire is used in a trial, present an analysis to examine the sensitivity of the trial results to changing the MAUI. Evidence suggests that differences in preferences as measured using different country scoring algorithms might be smaller than those measured by different MAUIs.

Discuss the interpretation of these QALY results. Assess the results against other outcomes measured in the trial. This could include reference to the consistency or inconsistency with any concomitantly assessed disease-specific quality-of-life and/or generic quality-of-life measure. This comparison across outcomes could help address questions of the sensitivity or responsiveness of the MAUI, and the plausibility of any argument that the evidence from the measure should be ignored as not being sensitive enough (rather than that the measure is correctly reflecting low strength of preference for the difference across the interventions and/or trade-offs due to adverse outcomes).

Also assess:

- whether the technique of measurement at baseline and during the trial is valid and likely to be free from bias (e.g. whether the results correlate with clinical or other measures of health outcomes in the trial);
- whether the results of the exercise are reliable (e.g. whether there is a high variance in results or inconsistencies in responses, or a high number of missing observations); and
- what attributes of health-related quality of life and other patient attributes are being valued.

## Source documents

For each of the responses provided above, provide adequate cross-referencing to page, table or figure numbers of the relevant trial report(s), as for Sub-section B3. For a complex systematic overview, consider re-presenting the tables from the main body of the assessment report in a technical document or attachment, as described Sub-section B3, including additional columns or footnotes for each table to indicate the source of the data in each row or cell, as appropriate.

## **B7 Extended assessment of comparative harms**

### **INFORMATION REQUEST**

- State whether there is any evidence beyond the direct randomised trials of delayed or rare adverse outcomes reported for the proposed therapeutic medical service, or whether there is any biological or clinical basis to suspect that such delayed or rare adverse outcomes might be anticipated.
- Specify and justify the search strategy used to identify suitable sources of evidence.
- Succinctly present any such evidence identified, with appropriate cross-referencing to any source documents provided in a technical document or attachment to the assessment report.
- Provide appropriate cross-referencing to any source documents provided in a technical document or attachment to the assessment report.
- Indicate how this extended harms profile compares with that of the main comparator.

Direct randomised trials are often an inadequate source of data on comparative harms. Thus, a wider basis of assessment of comparative harms from other sources (i.e. beyond the results of direct randomised trials) is encouraged to complement rather than replicate the assessment of comparative harms presented in response to Sub-section B6. This wide assessment is especially important for serious adverse outcomes that might occur in the long term or rarely, or when the proposed therapeutic medical service has a new mechanism of action, or if the mechanism of action and/or evidence of early physiological or biochemical changes suggests an increased potential for subsequent harms. Specify and justify the search strategy used to identify suitable sources of information about any such reactions. Extend the scope of this strategy beyond that presented in response to Sub-section B1.

Where these complementary data are from non-comparative sources, an overall comparative conclusion should be drawn. If the therapeutic conclusion in the assessment report is that the proposed therapeutic medical service is no worse than the main comparator in terms of effectiveness but is significantly less harmful, or there is an expectation that selection bias might have an influence, it is preferred that the advantage in terms of comparative harms is demonstrated as a pre-specified outcome in the context of direct randomised trials.

## **B8 Interpretation of the clinical evidence**

### **INFORMATION REQUEST**

- Provide a summary assessment of the overall trial evidence presented.
- Use this assessment to state the category from Table B8.1 that (in terms of comparative effectiveness and comparative safety) best reflects the therapeutic conclusion of the proposed therapeutic medical service over its main comparator, supported by the evidence presented.

Include in this assessment of the evidence a consideration of:

- the level of the evidence (Sub-section B2);
- the quality of the evidence (Sub-section B3 and B4);
- the clinical importance and patient relevance of the effectiveness and safety outcomes (Sub-section B5);
- the statistical precision of the evidence (Sub-sections B6 and B7);

- the size of the effect (Sub-sections B6 and B7); and
- the consistency of the results over the trials presented (Sub-sections B6 and B7).

The interpretation of the clinical data presented in Section B of the assessment report is crucial in determining the success of the assessment report. It is important to classify the therapeutic profile of the proposed therapeutic medical service in relation to its main comparator (i.e. whether it is therapeutically superior, inferior or equivalent to the comparator). Table B8.1 sets out a framework for this classification.

**Table B8.1 Classification of the comparative effectiveness and safety of the proposed therapeutic medical service compared with its main comparator and guide to the suitable type of economic evaluation**

Comparative safety	Comparative effectiveness			
	Inferior	Uncertain <sup>a</sup>	Non-inferior <sup>b</sup>	Superior
Inferior	Health forgone: need other supportive factors	Health forgone possible: need other supportive factors	Health forgone: need other supportive factors	? Likely CUA
Uncertain <sup>a</sup>	Health forgone possible: need other supportive factors	?	?	? Likely CEA/CUA
Non-	Health forgone: need other supportive factors	?	CMA	CEA/CUA
Superior	? Likely CUA	? Likely CEA/CUA	CEA/CUA	CEA/CUA

CEA = cost-effectiveness analysis; CMA = cost-minimisation analysis; CUA = cost-utility analysis

? = reflect uncertainties and any identified health trade-offs in the economic evaluation, as a minimum in a cost-consequences analysis

<sup>a</sup> 'Uncertainty' covers concepts such as inadequate minimisation of important sources of bias, lack of statistical significance in an underpowered trial, detecting clinically unimportant therapeutic differences, inconsistent results across trials, and trade-offs within the comparative effectiveness and/or the comparative safety considerations

<sup>b</sup> An adequate assessment of 'non-inferiority' is the preferred basis for demonstrating equivalence

The essential difference between assessing whether the proposed therapeutic medical service is superior or non-inferior to the main comparator is that the 95% confidence interval for superiority excludes the possibility that there is no difference between the therapies, whereas the 95% confidence interval for non-inferiority excludes the possibility that the proposed therapeutic medical service is inferior to a clinically important extent. Discuss any results to support a conclusion for non-inferiority in the context of the similarity or otherwise of the mechanism of action(s) of the proposed therapeutic medical service and the main comparator to assess whether this conclusion is supported by any other argument.

Categorising the proposed therapeutic medical service helps guide the selection of the more suitable options for the type of economic evaluation (see Sub-section D1). This includes the unusual circumstance of an assessment report for a proposed therapeutic medical service that is therapeutically inferior to its main comparator. It is theoretically possible to construct an economic evaluation if its overall cost of therapy is cheaper than that of its main comparator.

If the proposed therapeutic medical service is no worse than (non-inferior) the main comparator, there is no basis in terms of health outcomes to justify a higher price (unless there are cost offsets due to a different method of administering the proposed therapeutic medical service). A cost-minimisation analysis is therefore appropriate.

If the therapeutic medical service is superior to the main comparator, a cost-effectiveness analysis or cost-utility analysis is appropriate to determine whether the increase in health outcomes (and any cost offsets) justifies the increase in therapeutic medical service costs (and hence increased price) in terms of being acceptably cost-effective. If there are uncertainties and/or trade-offs across health outcomes (e.g. both increased effectiveness, and reduced safety or differing safety profiles), a cost-consequences analysis is appropriate to present these results in a disaggregated way against the costs and, if it helps to reduce the uncertainty and/or quantify the trade-offs, a cost-utility analysis would also be appropriate.



# Section C

## Translation Issues

---

### Introduction

The primary purpose of Section C of the assessment report is to guide the presentation of analyses conducted to translate the systematic overview of the results of evidence to the listing requested, and thus to the framework of the economic evaluation (Section D of the assessment report). This is particularly important when one or more variables incorporated into the economic evaluation are derived from, but not directly based on, the clinical evaluation presented in Section B of the assessment report. These variables might be derived using a number of analyses that modify the results of the clinical evaluation to help construct a modelled economic evaluation. Such analyses are referred to in these Guidelines as ‘pre-modelling studies’.

The need for pre-modelling studies arises because the study protocols for the trials used for the clinical evaluation might differ from the proposed clinical practice setting for the main indication in one of the following ways:

- The participants and circumstances of use in the trial might not be the same as the intended population for treatment in Australia (and might therefore have a different profile of risks of future events and circumstances of use). In this case, the clinical evaluation would need to be applied from the baseline risk of the sample of trial participants and their circumstances of use to the expected absolute risks of future events of the intended Australian population and their circumstances of use. Examples of pre-modelling studies of applicability include subgroup analyses and surveys of the patterns of health care resource provision in Australia corresponding to one or more health states included in a modelled economic evaluation.
- The length of follow-up (time horizon) of participants in the trial might be less than the expected duration of therapy or expected duration of overall health and health care resource impacts. In this case, the clinical evaluation would need to be extrapolated to the intended duration of therapy or expected health and resource impacts. Examples of pre-modelling studies of extrapolation include extrapolating integrals of time-to-event analyses and a review of the literature for single-arm follow-up studies of the natural history of the condition to estimate rates of disease progression.
- The outcomes measured in the trial might not be the patient-relevant final outcomes of treatment. In this case, the clinical evaluation would need to be transformed to take account of the patient-relevant final outcomes (in terms of quality-adjusted life-years [QALYs] gained). Examples of pre-modelling studies of transformation include transforming comparative treatment effects measured on surrogate outcomes to final outcomes and scenario-based studies to value health outcomes using utilities.

Thus, the results of the trials might need to be applied, extrapolated and transformed (collectively referred to in these Guidelines as ‘*translated*’) into a decision analysis appropriate for the intended clinical use of the proposed therapeutic medical service if publicly funded in Australia, taking into account the above issues. These pre-modelling studies provide a clearer and more systematic basis to support the necessary variables for inclusion in the economic evaluation (see Section D). As indicated by the examples above, the types of pre-modelling studies relevant to this process of translation can vary widely.

The methods of translation are described in Sub-section C2. The methods also help examine any impact of reintroducing sources of random error (the play of chance) and systematic error (bias), which were minimised in the systematic overview of the direct randomised trials presented in Section B of the assessment report. Given that these sources of error cannot be minimised to the same extent for indirect comparisons of randomised trials and non-randomised studies (see Part III, Sections B(i) and B(ii)), there is less basis to guide corresponding analyses in these circumstances. For assessment reports based on these types of studies, the information requests in Section C of the assessment report are shown in Part III, Section C(i) .

The results of pre-modelling studies are intended to inform:

- the underlying structure of the model and the selection of options for examination in an analysis of the structure of the model, and the scenarios it is examining; and
- the selection of values for variables in the economic evaluation, and ranges of plausible extremes to include in the associated sensitivity analyses.

Importantly, Section C requests a consistent format for the presentation of all pre-modelling studies. Each presentation has the following components:

- a succinct question to address a particular issue (Sub-section C1);
- a focused analytical plan that is presented and justified (Sub-section C2);
- a set of results (Sub-section C3); and
- an explanation of how these results contribute to the economic evaluation presented in Section D of the assessment report (Sub-section C4).

Presentation of Section C in the assessment report would be assisted by listing the issues to be addressed in pre-modelling studies in a single response to Sub-section C1, preferably with a concluding tabulated summary. Then present the pre-modelling studies sequentially in a series of Sub-section C2 and C3 pairs (i.e. the focused analytical plan in response to Sub-section C2 requests and the results in response to Sub-section C3 requests). A single response to Sub-section C4 should then summarise the main results of the pre-modelling studies together and indicate how their results are to be used in the economic evaluation presented in assessment report Section D.

## C1 Identification of issues to be addressed

### INFORMATION REQUEST

- **Define application issues:** Describe any ways in which the participants and circumstances of use in the trial differ from the proposed population for treatment (including the baseline risk of participants and circumstances of use).
- **Define extrapolation issues:** State whether there is a need to extrapolate the outcomes reported in the clinical evaluation beyond the trial or study horizon.
- **Define transformation issues:** State whether there is a need to transform the nature of the outcome measured in the clinical evaluation (i.e. taking a surrogate or intermediate endpoint, and transforming it to a QALY or equivalent).
- **Define any other translation issues:** State whether there is any other need to translate from the clinical evaluation.
- **Convert** each defined translation issue into a succinct question that can be addressed in a pre-modelling study.

In many circumstances, the direct randomised trial evidence identified in assessment report Section B can be used to directly support the listing requested; for example, in the context of a therapeutic conclusion where the proposed therapeutic medical service is no worse than the main comparator. However, in other circumstances, additional argument and associated analyses are needed to translate the evidence more rigorously to the listing requested.

The following guidance is intended to help an applicant decide whether additional analyses are needed and to identify methodological options that might be considered. It is recognised that not all the necessary information will be available to inform every aspect of each circumstance and the resulting analyses. Methodological experts might also disagree about the most appropriate methodological option to pursue in particular circumstances. However, this detailed guidance is warranted because many assessment reports have had difficulties in this area.

The issues identified in response to Sub-section C1 should focus on those for which pre-modelling studies are presented in Section C of the assessment report. At the end of the response to Sub-section C1, tabulate a summary list of these material translation issues in the order identified. Separately tabulate a summary list of any other translation issues identified, but for which pre-modelling studies are not presented. In each case, summarise in the table why a pre-modelling study is not presented (e.g. not expected to make a material difference).

### Applicability issues

Define any issues that indicate a need to apply the trial data to the intended population and circumstances of use. Applicability issues might arise due to differences between participants enrolled in the trials and patients who would be likely to obtain the proposed therapeutic medical service if publicly funded in Australia, and between the circumstances of use in the trials and those that would occur if publicly funded in Australia.

Table B4.2 identifies some important patient factors that might affect outcomes. There might also be important differences in the mix of patients who would receive the proposed therapeutic medical service if publicly funded in Australia. For example, it is a concern of MSAC that there might be patients in the community who have a disease that is less severe than that of participants in the randomised trials. There might also be patients in the community for whom the main comparator can be expected to perform better than in the trials. Both could diminish the difference in effectiveness between the proposed therapeutic medical service and the main comparator, and therefore make the incremental cost-effectiveness ratio less favourable for the proposed therapeutic medical service.

Table B4.3 identifies some factors relating to the circumstances of use. These factors might also include extrapolating results of trials conducted in hospitals to use outside the hospital and the effect of more rigorous follow-up, which might swamp important differences in the convenience and acceptability of the proposed therapeutic medical service compared with alternative treatments. This might have resulting effects on patient compliance and subsequent response to treatment.

The fact that one or more differences might be demonstrated does not necessarily raise an applicability issue, because the differences might not help to predict any variation in treatment effect. However, the demonstration of such differences does identify areas that could be examined, such as in the examples given in the following Sub-sections.

### **Population characteristics**

There might be evidence within the trials and/or other sources to indicate that patients vary in their expected risk of adverse major clinical outcomes. In such cases, which are common for many medical conditions, additional analysis of the comparative treatment effect detected in the trials, presented as a pre-modelling study, might indicate that this effect is best summarised as a constant relative reduction in the risk of these outcomes across the trial population of varying baseline (expected) risks.

If this is the case, such an analysis forms an acceptable basis to apply the trial data to specific subgroups. For example, this evidence would be sufficient to justify targeting a requested restriction to those patients with a greater expected absolute risk of future events at the point of deciding whether to start therapy with the proposed therapeutic medical service (i.e. a poorer prognosis) as being the patients likely to benefit most from the proposed therapeutic medical service. Any thresholds of greater expected absolute risk to identify the population that would be eligible to start the proposed therapeutic medical service according to the requested restriction (see Section A) would need to be justified and supplemented by sensitivity analyses on different thresholds. The absolute or incremental treatment effect would then be calculated by multiplying the expected absolute risks across the eligible population by the estimated overall relative treatment effect. As a check, present the results of the targeted subgroup that might be recruited in the randomised trials as the absolute risk difference, or explain why this is not possible.

The comparative treatment effect detected in the trials might indicate that this effect is best summarised as a varying relative reduction in the risk of these outcomes across the trial population of varying baseline risks. In this case, which is less common than the previous example, the pre-modelling analysis would need to identify treatment effect variation when measured in relative terms (e.g. relative risk, hazard ratio, odds ratio). This analysis of the relative treatment effect would need to show sufficient heterogeneity within the set of direct randomised trials available to support statistically a claim regarding the nature (qualitative or quantitative) and extent of each treatment effect variation, and thus any resulting subgroup analysis.

Variations in the relative treatment effect might arise with varying characteristics of the patient, the therapeutic medical service(s) or the medical condition. Together with a justification of any thresholds as necessary (supplemented by sensitivity analysis on different thresholds), this evidence contributes to an argument to target a requested restriction to these patients (see Section A) and to calculate the absolute treatment effect by applying the estimated relative treatment effect for the subgroup to the expected risk for the subgroup.

### **Circumstances of use**

- One or more of the direct randomised trials might include methods of delivery and/or co-administered therapies that are not recommended by the TGA, or that might otherwise have an impact on the direction and/or magnitude of the treatment effect.
- One or more of the direct randomised trials might have been conducted in settings that are not applicable to the requested listing on the MBS or with some trial participants who would not be eligible for the proposed medical service according to the requested restriction.
- One or more of the direct randomised trials might deliver the proposed therapeutic medical service in a way that differs from how it would be delivered if publicly funded in Australia.

In addressing this last point, consideration will need to be given as to whether:

- the effectiveness of the proposed therapeutic medical service is operator dependent;
- the proposed therapeutic medical service consists of many components;
- the components of this proposed therapeutic medical service would all be available; and
- any infrastructure required would also be available (e.g. monitoring the proposed therapeutic medical service with regular blood tests).

There is no limit to the types of difference in populations and circumstances of use, but only a small number of these might modify the extent of treatment effect detected by the overall results of the trial or meta-analysis. Thus, the general rule is to apply the overall treatment effect from the ITT population, rather than to explore for possible variations in treatment effects in subgroups.

As discussed in Sub-section C2, an analysis to support a claim of treatment effect variation according to a particular patient characteristic or circumstance of use is more convincing if it was pre-specified with a biologically plausible rationale before the collection of any data in the trial(s) providing the source data for the analytical plan. Thus, for each analytical plan relying on direct randomised trial(s) and examining an applicability issue, state whether the data was collected before or after finalisation of the analytical plan (see below).

If an applicability issue involves introducing one or more diagnostic criteria or tests specifically to identify patients who are eligible according to the requested restriction that was not relied on in the trials, then separately present additional information on the validity (specificity, sensitivity, positive predictive value and negative predictive value), reliability and comparability of these criteria and tests, both across all trials presented and in regular Australian practice. This is necessary to examine the impact of false positive and false negative identification of eligible patients, as well as the impact of false positive and false negative identification of treatment response, on the application of the trial results. This is particularly the case if the latter are used in any proposed continuation criteria in the requested restriction. Sections A and Sub-section D4 provide further advice on specifying and costing these diagnostic criteria and tests in the diagnostic and treatment algorithm, and on the implications of misclassification for estimating incremental effectiveness and incremental cost-effectiveness.

If there is no applicability issue, state this.

### **Extrapolation issues**

Define any issues that indicate a need to extrapolate the within-trial patterns of resource provision (cost) and within-trial health outcome results, including time-to-event data, beyond the time horizon of the direct randomised trials. Such extrapolation might be considered necessary in the context of a modelled economic evaluation, to determine comparative effectiveness and cost-effectiveness beyond the median duration of therapy and/or follow-up in the presented direct randomised trials.

If there is no need to extrapolate the evidence from the clinical evaluation, state this.

### **Transformation issues**

Define any issues with outcomes that indicate a need to transform the nature of the outcome(s) measured in the direct randomised trials to those relied on in the economic evaluation. For example, the direct randomised trials might only report outcomes that are of less patient relevance than intended final outcomes of treatment. These less relevant outcomes are known as *surrogate outcomes*. Arguably, the closer a surrogate outcome is to the final outcome, the more useful it is, but generally the more difficult it is to measure accurately.

To transform the surrogate outcomes measured in the trials to final outcomes and to extend the range of outcomes (e.g. the number of patients with unhealed peptic ulcers who eventually need surgery), the trial results might need to be supplemented by estimates obtained from other sources (see Sub-section C2).

For most proposed therapeutic medical services, the ultimate outcome of therapy is to improve quality of life and/or survival. In theory, all outcomes could be expressed as QALYs gained (see Appendix 4). In practice, few randomised trials have measured the impact of a proposed therapeutic medical service on QALYs, because few are large enough or long enough to measure changes in final outcomes directly.

Another common need is to transform the outcome(s) measured in the clinical evaluation to value them in utility terms for the economic evaluation (see Appendix 4 for more information on utility terms). If this transformation supplements any other transformation (e.g. from surrogate outcomes measured in the direct randomised trials to patient-relevant outcomes), present the links between these two transformations and any assumptions involved in combining them.

Other transformations that have been considered include:

- converting outcomes reported as continuous data to dichotomous data; and
- converting outcomes reported as dichotomous data to time-to-event data to estimate periods of time in one or more health states, or periods of time free from being in one or more health states.

Although these transformations increase uncertainty, they can allow for a more readily interpretable health outcome (see Sub-section C2).

If there is no need to transform the outcomes measured in the direct randomised trials, state this.

### **Other translation issues**

Define any other issues that required pre-modelling studies to justify an aspect of the economic evaluation (see Section D). Examples of other issues that might be included here are as follows:

- One or more of the direct randomised trials was less successful in minimising bias (e.g. inadequate concealment of randomisation, inadequate blinding of subjective outcomes, unable to reconstruct full ITT analysis).
- One or more of the direct randomised trials reported fewer patient-relevant outcomes or no patient-relevant outcomes.
- One or more of the direct randomised trials was of insufficient duration to detect the most patient-relevant outcomes.
- The patterns of resource provision measured in the direct randomised trials did not closely reflect those in Australia (and/or the likely changes in patterns of resource provision were not measured in the trials).

Randomised trials performed overseas are an acceptable basis for an economic evaluation relevant to Australian practice. However, although the overall estimate of the change in a final or surrogate outcome might be transferable to Australia, estimates of the costs of resources provided (such as further investigations, procedures or operations) are often not readily transferable.

- It is usually apparent that the unit costs are quite different.

- Less apparent, but also important, is the fact that the frequency or patterns of resource provision might not be relevant to Australia because of major differences in medical practice or different incentives in different economies and health care systems.
- Sometimes assumptions need to be made during the translation of overseas randomised trials to create a modelled economic evaluation that is relevant to the Australian context.
- The trials did not measure provision of all types of health care relevant resources (which might change and therefore would need to be added in a model).
- The protocols of the trials required more resources to be provided than would be typical in normal management of the medical condition (such as extra monitoring to demonstrate safety or effectiveness). In this case, only resources provided or avoided in regular clinical practice need to be included in a model.

If there are no other issues that require pre-modelling, state this.

## **C2 Focused analytical plan**

### **INFORMATION REQUEST**

- Present and justify a focused analytical plan (which specifies details of data, sources, methods and analyses) to address each translation issue identified.

For each translation issue identified in Sub-section C1, provide a focused analytical plan that clearly describes the:

- issue;
- specific question to be addressed by a pre-modelling study;
- data to be used and their sources; and
- methods of the pre-modelling study (with sufficient details to enable independent verification of the analysis).

A range of methods that might inform the development of an analytical plan are shown below. Justify the choice of method where more than one option exists. Comment on any implications of this choice for the results of the pre-modelling study, including how the choice of the method will be assessed; for example, in the sensitivity analyses of the economic evaluation.

### **Methods to address applicability issues**

Addressing applicability issues might involve investigations of heterogeneity, treatment-effect variation, subgroup analysis and/or meta-regression.

### ***Heterogeneity analysis***

Assess the statistical analyses of heterogeneity in the meta-analyses presented in Sub-section B6. For dichotomous outcomes, separately assess these analyses for the relative risk and the risk difference. The results of a Mantel–Haenszel fixed effect model could be presented in addition to the DerSimonian–Laird random effects model to help examine the assessment of heterogeneity.



Discuss and explain any suggested heterogeneity of trial results. Reasons for heterogeneity might include differences in trial population or design. If there are strong biological or methodological grounds for heterogeneity, consider presenting a pre-modelling study to examine the impact of these grounds for heterogeneity by comparing relevant pooled analyses with the overall estimate. Unexplained heterogeneity, depending on its direction and magnitude, generally makes the summary estimator less meaningful.

Assessment of heterogeneity is an important aspect of interpreting meta-analyses where there are a large number of trials. Refer to biological or clinical reasoning as appropriate when justifying the inclusion of further analyses in pre-modelling studies to take into account heterogeneity when considering the application of the results of the trials.

Explain and justify the presentation of any additional meta-analyses in which trials listed in response to Sub-section B2 are excluded (e.g. on the grounds of inadequately minimising bias or of reporting fewer patient-relevant outcomes) and examine the impact each exclusion has on the overall meta-analysis. Similarly, explain and justify the presentation of any additional meta-analyses in which trial groups are excluded and examine the impact each exclusion has on the overall meta-analysis.

Support any claimed treatment effect variation on the basis of observed heterogeneity with reference to the excluded trials and/or trial groups, and the covariate that predicts the treatment effect variation, such as:

- varying duration of use;
- settings of use;
- patient baseline characteristics, including risk factors and disease severity; and
- radiation dose-response or surgical experience considerations.

If any heterogeneity is thought to be due to the trials having different periods of follow-up, presenting the pooled incidence rate differences might be useful.

Assessment of possible publication bias, where there are sufficient trials, might be assisted by presentation of a funnel plot.

### ***Presenting and justifying a subgroup analysis or a meta-regression***

In general, an estimate of treatment effect is interpretable with respect only to the **whole** population of a randomised trial (or whole population of randomised trials within a meta-analysis) rather than by testing within each individual subgroup. Subgroup analysis, to determine whether a treatment effect varies across patient groups, should be interpreted with caution if it is not adequately pre-specified. This would occur if, before any data were collected, the subgroups were not defined, treatment allocation was not stratified or an alpha-spending plan was not formally included in the trial design. Justify any decision to identify the treatment effect obtained from a patient subgroup as the basis for the estimate of treatment effect for a requested listing.

Information presented in support of any presentation of a subgroup analysis or meta-regression in Section C should include each of four elements:

- a discussion of the plausibility of a variation in treatment effect;

- an indication of whether the hypothesis underpinning the analysis was developed before or after the trial data were collected;
- a statistical analysis of the variation in treatment effect; and
- an account of the number of pre-specified subgroup analyses conducted.

In isolation, no single element is convincing either in support of or against a subgroup analysis or meta-regression based on a claim of substituting the comparative treatment effect from this analysis for the estimate from the whole population in the trial or meta-analysis. Congruence of support across these elements (which are outlined in more detail below) strengthens the claim; conflicting conclusions across the elements weaken the claim. Each claim and its supporting information need to be judged on a case-by-case basis, and this judgment can be influenced by other relevant factors.

These elements apply when subgroups consist of participants within randomised trials, a single randomised trial, or groups of randomised trials within a meta-analysis. However, some of the underlying principles cannot be used to translate a treatment effect from a first-line to a second-line setting, although subgroup analyses might be constructed if separate subgroups of trial participants in both treatment arms are treated in either the first or second-line setting. Similarly, as discussed in Section A for continuation criteria in restrictions, the underlying principles might not readily apply to groups of patients who become identifiable after therapy has started (e.g. patients who achieve an early marker of response to therapy or who withdraw early from therapy). Such patients might appear to generate comparatively important impacts on an economic evaluation. However, these early effects also introduce a range of confounders (such as regression to the mean), which means that it is difficult to attribute the impacts to the substitution of the proposed therapeutic medical service for the main comparator.

### ***Plausibility of treatment effect variation***

Discuss the biological and clinical plausibility of the claim for sufficient variation of comparative treatment effect to justify the use of results other than for the whole population. An unexplained variation is difficult to accept, but reliance on plausible explanations can often be misplaced.

### ***Pre-specification of treatment effect variation***

A conclusion of sufficient variation of treatment effect to justify the use of results other than for the whole population is strengthened if the subgroup analysis arises from an explicit hypothesis relating to the given subgroup included in the pre-specified analytical plan of the trial protocol. This is related to the previous element because it is difficult to specify implausible subgroups before collecting and analysing randomised trial data, whereas it is relatively easy to develop a plausible explanation for an unpredicted variation observed in the relative treatment effect data. A subsequent trial can be conducted to test a subgroup hypothesis generated from an earlier trial. If this is relevant to the assessment report, respond with reference to the most recent trial. The first statistical finding of treatment effect variation is usually sufficient to generate a hypothesis; its confirmation in a pre-specified analysis in a subsequent trial is more persuasive.

## **Statistical analysis of variation of the comparative treatment effect**

An important distinction exists between absolute treatment effect variation (e.g. of the absolute risk difference or weighted mean difference) and relative treatment effect variation (e.g. of the relative risk, relative risk reduction, odds ratio or hazard ratio).<sup>1</sup> Absolute treatment effect variation is common and has been observed more frequently than relative treatment effect variation. In several disease states, treatment effect variation has been observed across varying expected risks at baseline (i.e. the predicted risks of events before treatment) for the absolute effects, but not for the relative effects. This supports a conclusion of constant relative risk and has formed an accepted basis for targeting therapy to patients likely to benefit most (i.e. those with the greatest absolute risk difference) on the grounds that they have the greatest predicted risks of events at the point of deciding whether to start therapy with the proposed therapeutic medical service. This is calculated by multiplying the predicted risks of events in the intended subgroup(s) of the population at this decision point by the relative risk estimated from the whole population of the randomised trial(s) to calculate the absolute risk difference in the subgroup(s) for whom therapy with the proposed therapeutic medical service might be targeted.

In any presentation of a subgroup analysis or meta-regression, present tests for variation of the absolute and relative treatment effects, where possible, using appropriate tests for interaction between the treatment effect and the subgroup populations. The test should support and quantify the association between the treatment effect and the covariate defining the subgroup. This covariate provides a threshold that defines the restricted population; if a continuous variable is used, perform a sensitivity analysis on the threshold value chosen to define the subgroup.

For a subgroup analysis using dichotomous data from a single randomised trial, the test for interaction should compare across the nominated subgroup and its complement of all other participants in each arm of the trial. Present the treatment effects (measured on the pre-specified primary outcomes and any relevant secondary outcomes) as the relative risk and the risk difference, each with the chi-square test (presented as the *P*-value), using the Cochran *Q* statistic. Present the *I*<sup>2</sup> statistic with its 95% uncertainty interval. As discussed above, statistically significant variation of *relative* treatment effects is a more unusual finding. Statistically significant variation of *absolute* treatment effects is more common and might simply reflect constant relative treatment effect with varying baseline (expected) risks across the trial population.

### **Presentation of analysis**

Table C2.1 shows a suggested format to present tests for interaction across subgroups on treatment effects from a single randomised trial.

---

<sup>1</sup> Absolute treatment effect variation is also known as ‘treatment effect variation on the additive scale’ and relative treatment effect variation is also known as ‘treatment effect variation on the multiplicative scale’.

**Table C2.1      Assessment of treatment effect variation across subgroups**

Outcome	Proposed therapeutic medical service <i>n</i> with event/ <i>N</i> (%)	Main comparator <i>n</i> with event/ <i>N</i> (%)	Relative risk (95% CI)	Risk difference (95% CI)
<b>Primary outcome</b>				
Identified subgroup				
Complement of subgroup				
Meta-analysis of subgroups using random effects model	–	–		
Test for treatment effect variation	–	–	<i>P</i> =	<i>P</i> =
<i>I</i> <sup>2</sup> statistic with its 95% uncertainty interval	–	–		
Overall trial results as reported				
<b>Each other outcome</b>				
Identified subgroup				
Complement of subgroup				
Meta-analysis of subgroups using random effects model	–	–		
Test for treatment effect variation	–	–	<i>P</i> =	<i>P</i> =
<i>I</i> <sup>2</sup> statistic with its 95% uncertainty interval	–	–		
Overall trial results as reported				

– not applicable; CI = confidence interval; *n* = number of participants with event; *N* = total participants in group; *P* = probability

To extend this to more than one randomised trial in a meta-analysis, adopt a similar approach. Pool the subgroups and then their complements across trials, each using a random effects model, and analyse the chi-square test (presented as the *P*-value), using the Cochran *Q* statistic across the pooled results. Present the *I*<sup>2</sup> statistic with its 95% uncertainty interval. Tables C2.2 and C2.3 show a suggested format to present tests for interaction across subgroups on a treatment effect from a pooled analysis of randomised trials. The presentation includes a forest plot showing the individual trials, followed by a pooled analysis for each of the two subgroups. In this case, the vertical line for the forest plot should run through the point estimate of the overall treatment effect (rather than the null), and some indication of the 95% confidence interval around this estimate of treatment effect should be highlighted (e.g. by shading). Finally, present a pooled analysis across the subgroups and compare this with the results for the overall population.

Where there are many analyses of outcomes for a subgroup, present a summary table as shown in Table C2.4.

**Table C2.2      Assessment of relative treatment effect variation across subgroups**

Outcome	Proposed therapeutic medical service <i>n</i> with event/ <i>N</i> (%)	Main comparator <i>n</i> with event/ <i>N</i> (%)	Forest plot here	Relative risk (95% CI)
<b>Primary outcome</b>				
Identified subgroup • Trial 1 • etc. Meta-analysis of subgroup using random effects model	–	–		
Complement of subgroup • Trial 1 • etc. Meta-analysis of subgroup using random effects model	–	–		
Meta-analysis of subgroups using random effects model	–	–		
Test for treatment effect variation	–	–	–	<i>P</i> =
<i>I</i> <sup>2</sup> statistic with its 95% uncertainty interval	–	–	–	
Meta-analysis of whole population using random effects model as reported	–	–		
<b>Other outcomes</b>				
Identified subgroup • Trial 1 • etc. Meta-analysis of subgroup using random effects model	–	–		
Complement of subgroup • Trial 1 • etc. Meta-analysis of subgroup using random effects model	–	–		
Meta-analysis of subgroups using random effects model	–	–		
Test for treatment effect variation	–	–	–	<i>P</i> =
<i>I</i> <sup>2</sup> statistic with its 95% uncertainty interval	–	–	–	
Meta-analysis of whole population using random effects model as reported	–	–		

– = not applicable; CI = confidence interval; *n* = number of participants with event; *N* = total participants in group;  
*P* = probability

**Table C2.3      Assessment of absolute treatment effect variation across subgroups**

Outcome	Proposed therapeutic medical service <i>n</i> with event/ <i>N</i> (%)	Main comparator <i>n</i> with event/ <i>N</i> (%)	Forest plot here	Risk difference (95% CI)
<b>Primary outcome</b>				
Identified subgroup • Trial 1 • etc. Meta-analysis of subgroup using random effects model	–	–		
Complement of subgroup • Trial 1 • etc. Meta-analysis of subgroup using random effects model	–	–		
Meta-analysis of subgroups using random effects model	–	–		
Test for treatment effect variation	–	–	–	<i>P</i> =
<i>I</i> <sup>2</sup> statistic with its 95% uncertainty interval	–	–	–	
Meta-analysis of whole population using random effects model as reported	–	–		
<b>Other outcomes</b>				
Identified subgroup • Trial 1 • etc. Meta-analysis of subgroup using random effects model	–	–		
Complement of subgroup • Trial 1 • etc. Meta-analysis of subgroup using random effects model	–	–		
Meta-analysis of subgroups using random effects model	–	–		
Test for treatment effect variation	–	–	–	<i>P</i> =
<i>I</i> <sup>2</sup> statistic with its 95% uncertainty interval	–	–	–	
Meta-analysis of whole population using random effects model as reported	–	–		

– not applicable; CI = confidence interval; *n* = number of participants with event; *N* = total participants in group; *P* = probability

**Table C2.4 Summary of assessment of treatment effect variation across subgroups**

Outcome	Relative risk (95% CI)	Risk difference (95% CI)
<b>Primary outcome</b>		
Pooled results for identified subgroup using the random effects model		
Pooled results for complement of subgroup using the random effects model		
Meta-analysis of subgroups using random effects model		
Test for treatment effect variation	<i>P</i> =	<i>P</i> =
<i>I</i> <sup>2</sup> statistic with its 95% uncertainty interval		
<b>Other outcomes</b>		
Pooled results for identified subgroup using the random effects model		
Pooled results for complement of subgroup using the random effects model		
Meta-analysis of subgroups using random effects model		
Test for treatment effect variation	<i>P</i> =	<i>P</i> =
<i>I</i> <sup>2</sup> statistic with its 95% uncertainty interval		

CI = confidence interval; *P* = probability

As discussed above, a test for interaction is more likely to suggest a possible signal for variation across the absolute risk difference (i.e. on the additive scale). However, given that this is more likely to be explained by varying baseline (expected) risk across the subgroups, the results for the subgroup should generally not be used where the test for interaction for the relative risk (i.e. on the multiplicative scale) does not suggest treatment effect variation. In this circumstance, it is usually more reasonable to conclude an overall constant relative risk and therefore apply the results of the trial(s) from the full (ITT) trial population to any subgroup identified by a greater expected risk. It is less common for the test for interaction to suggest a possible signal for variation across the relative risk (i.e. on the multiplicative scale). In this circumstance, it might be appropriate to apply the results from the subgroup analysis rather than the full (ITT) trial population. A strong basis is needed to justify substituting the results of a subgroup analysis for the full population because of the greater risk of random error (play of chance) due to smaller sample sizes in the subgroups and the impact of multiple analyses.

Indicate whether the results of the identified subgroup and its complement are qualitatively different from the primary analysis of the trial(s) and/or the corresponding secondary analysis for the full trial population (i.e. a different conclusion on treatment effect might be drawn), or whether they are quantitatively different (i.e. a similar conclusion on treatment effect might be drawn, but the magnitude of effect might be different).

Meta-regression refers to analyses in which the characteristics of the randomised trials or of the participants in the randomised trials are used as explanatory variables (covariates) in a multivariate regression analysis with the relative effect size (or some measure of deviation from the summary measure of effect) as the dependent variable. Meta-regression has a potential advantage when compared to the stratified analyses, based on subgroups described above, in that it examines more than one covariate simultaneously to determine whether there is more than one potential explanation of treatment effect variation. The data can be analysed at the trial level (more commonly done, but potentially confounded) or at the individual patient level (with the trial as a covariate). In meta-regression, the unit of observation is the trial or the subgroup. Where meta-regression is used, clearly describe the method.

If a regression-based approach is adopted, then to minimise over-fitting, enough data points are required to detect any underlying relationships between the covariate defining the subgroup and the treatment effect measured as the absolute risk difference and the relative risk. At the trial level, this approach is only useful where the number of trials is large. It cannot be sensibly attempted when small numbers of trials are being combined (e.g. at least five to ten trials are needed for each covariate examined).

### ***Multiplicity of treatment effect variation analyses***

Report the number of pre-specified subgroup analyses conducted. If a subgroup analysis or a meta-regression is presented that was not pre-specified, report the number of such subgroup analyses or meta-regressions conducted of the data in total. Report any adjustment for multiple comparisons.

### **Methods to address extrapolation issues**

#### ***Extrapolating time-to-event data***

Several different methods might be used to extrapolate time-to-event data, and a range of assumptions need to be tested in an extrapolation of survival or time-to-event data beyond the horizon of the trial. Justify the assumption (whether made directly or indirectly) in relation to the hazard ratio reflecting the comparative treatment effect beyond the time horizon of the trial(s). This should be consistent with the duration of therapy and should be biologically plausible with its expected impact on the medical condition being managed. Provide particularly strong justification to maintain a hazard ratio more favourable than one beyond the trial follow-up and duration of therapy.

Examine several alternative methods of extrapolation. Present the results of each method of extrapolation superimposed on the corresponding Kaplan–Meier curves from the direct randomised trials (see Sub-section B6). Present goodness-of-fit tests as part of the justification of the choice of the preferred extrapolation method of these curves and examine the sensitivity of any extrapolation that relies on observed data beyond the median duration of follow-up. Also apply these extrapolations to 95% confidence limits of each of these curves to reflect appropriately the uncertainty of the unextrapolated curves.

If the economic evaluation is based on an extrapolation of time-to-event data, also present the within-trial case (i.e. within the time horizon of the trial evidence) alongside the extrapolation, because this allows an at-a-glance assessment of the extent to which the incremental gains arise within the time horizon of the trial compared with the extrapolated time horizon. Similarly, if the proposed approach to extrapolating the time-to-event results does not result in a convergence of the two extrapolated curves, present an analysis that incorporates a linear triangulation from each of the observed curves at the point of median duration of follow-up to a single common maximum end point justified as being clinically plausible. Another method to converge these curves would be to project the curve representing the outcome with the main comparator beyond the median duration of the trial follow-up, and apply a hazard ratio of one to estimate the projection of the curve representing the outcome with the proposed therapeutic medical service from this time point. Particular justification would be needed to apply a hazard ratio representing a continued differential treatment effect beyond the median duration of the trial.



### ***Use of data from non-randomised studies to extrapolate beyond the evidence from randomised trials***

Data from non-randomised studies are sometimes useful to extrapolate beyond the results of direct randomised trials. This is because the trials might have been of insufficient size or duration to capture the full impact of therapy on the outcomes of the disease, or the typical resource provision measured in an overseas trial might need adjustment to reflect patterns of resource provision in Australia. In contrast, the non-randomised studies might involve longer follow-up for an active main comparator, or the natural history of the medical condition if the main comparator is not an active intervention. Given that the data from non-randomised studies are subject to bias, assumptions based on those data made during a modelling exercise should be cautious.

When presenting data from non-randomised studies for extrapolation purposes in a modelled economic evaluation, demonstrate that a systematic approach has been taken to search for, locate and select the non-randomised studies for presentation. The selection process should be presented and justified. Provide a report of each study in a technical document or attachment. The results of the non-randomised study might contribute to finding and justifying a variable in the economic evaluation. This variable might vary from a single point estimate to a regression formula. The results of the non-randomised study might also help identify risk factors that contribute to the expected risks of the comparator arm in a model.

When indicating which results are being extrapolated, explain how the extrapolations are achieved by the model for the streams of costs and outcomes for the proposed therapeutic medical service and the main comparator. In particular, if non-comparative data are used (e.g. from single-arm studies), it is necessary to make an assumption about how the other arm in the model would change. The usual practice, in the absence of empirical evidence to the contrary, is to assume that the comparator arm would change so that the relative risk between the two arms measured in the randomised trial(s) remains constant across the duration of therapy. Justify the use of this (or any other) assumption in the model presented in the assessment report.

### **Methods to address transformation issues**

#### ***Use of surrogate outcomes to estimate final outcomes***

The claim that an incremental treatment effect on a surrogate outcome measured in respect to the proposed therapeutic medical service quantitatively predicts a subsequent incremental treatment effect on a final outcome is more persuasively shown if attention is given to the following issues:

- **Step 1** — Present a systematic review of the literature to examine whether epidemiological evidence and biological reasoning has established that there is a relationship between the surrogate outcome and the final outcome independent of any intervention. In a few instances, relationships have been established, or have been proposed, between surrogate outcomes and final outcomes. Examples include blood left ventricular ejection fraction and survival after myocardial infarction, or viral load and cure of viral hepatitis.

- **Step 2** — Present a systematic review of the literature to examine whether direct randomised trial evidence using other active medical services has shown that there is a basis to conclude that a treatment effect on the surrogate outcome has satisfactorily predicted a treatment effect on the final outcome. (If there is evidence of this type for the proposed therapeutic medical service, this might help support a biological argument for the therapy.) Based on this evidence, quantify the relationship between these treatment effects with an assessment of the uncertainty of the relationship. Discuss the reproducibility of these findings (e.g. whether they have been consistently shown across more than one trial and for more than one alternative active medical service).
- **Step 3** — Explain why this relationship between the treatment effects on these outcomes with these other active medical services is likely to apply to the proposed therapeutic medical service. At present, it is difficult to give categorical advice. Consider which outcomes are most appropriate and most feasible, given the data available. The clinical importance and patient relevance of the outcomes should be established and, where possible, supported with data.

Having addressed the three steps above in transforming a treatment effect on a surrogate outcome to a treatment effect on a final outcome, explain in response to Sub-section D4 how this is included in the economic evaluation, including by specifying and referencing the sources of the longer term natural history (e.g. longitudinal population studies) as well as the transformed treatment effects.

### ***Valuing health outcomes***

Where the final outcome of the proposed therapeutic medical service is a change in quality of life (with or without a change in the number of projected life-years gained), a separate utility analysis is appropriate to transform this change into a preference-based measure. Appendix 4 provides further guidance on the presentation of a pre-modelling study to elicit the utility valuations.

### ***Other useful transformations of outcomes measured in direct randomised trials***

Outcomes that are expressed as dichotomous outcomes measured on a per patient basis (e.g. proportion of participants in response to treatment or for whom blood pressure was ‘controlled’ following the stated period of time after randomisation at which these data were collected in the trial) are easier to interpret and to incorporate into an economic evaluation than a difference in means for a quality-of-life scale or a physiological variable. Further, converting these proportions, as appropriate, to estimate periods of time free of an event, time with an event or time in a health state allows for a more interpretable incremental cost-effectiveness ratio if there is no limit to the duration of therapy. Consider providing a technical document or an attachment to the assessment report to give the details of the methods of these transformations.

### **Methods to address other translation issues**

#### ***Examination of exclusion of trials from the meta-analyses presented in Sub-section B6***

Examination of the impact of removing trials from a meta-analysis can sometimes suggest explanations for translating the clinical evaluation. If one or more trials are to be excluded from a meta-analysis, identify the aspect(s) of each trial that justify the exclusion (see

Table C2.5). Indicate whether each reason relates to an applicability, extrapolation or transformation issue (see above), or whether a translation issue arises because one or more of the direct randomised trials was less successful in minimising bias, or reported fewer or no patient-relevant outcomes. Present greater detail of each aspect (as a minimum, to the extent requested in the relevant text in Sub-section B2) or refer to the information provided in Table B2.4.

If there is more than one type of reason for exclusion, arrange the trials for exclusion in Table C2.5 by reason for exclusion. Present each relevant meta-analysis both with and without the trial(s) excluded. Discuss any implications of the exclusions for the interpretation of the results of the meta-analysis.

**Table C2.5 Reasons to exclude each direct randomised trial**

<b>Trial ID</b>	<b>Ground(s) for seeking exclusion</b>	<b>Details <sup>a</sup></b>
Trial 1		
Etc.		

ID = identification

<sup>a</sup> Cross-reference each set of details to the source of information (specifying the trial report with page, table, figure number).

### ***Adjustment of resource provision estimates***

A survey of patterns of resource provision in Australia might be needed if resource provision in the direct randomised trials reflects patterns of resource use that are different from those used and likely to be replaced in Australia (e.g. if they reflect overseas health care systems or the requirements of the trial protocol) or were incompletely measured. This survey could be a cross-sectional study observing and recording patterns of resource provision in Australia. An alternative, but less preferred option could be a survey of Australian expert opinion on the likely patterns of resource provision, either describing overall Australian practice or advising on modifying overseas patterns that are more relevant to Australia (see Appendix 2).

Justify the application of these cross-sectional data into a longitudinal model and consider any possible implicit assumptions. For example, if response to the proposed therapeutic medical service involves returning to a less severe health state, the associated patterns of resource provision might not necessarily reflect those of an earlier health state (i.e. before the disease progression meant that the patient became eligible for the proposed therapeutic medical service). As an extreme example of this, applying patterns of resource provision for asymptomatic patients would obviously not be reasonable if those patterns ascertained for patients with watchful waiting at an early stage of an indolent disease were related to patients achieving full symptom control on analgesics at a terminal stage of the same disease.

If any patterns of resource provision from a trial are to be modified in a model (such as the exclusions of 'protocol-derived' resource provision), discuss the extent to which these resources might have affected the results of the trials in terms of health outcomes (e.g. high-intensity screening for deep vein thromboses in trials associated with lower rates of pulmonary embolism than in usual care). This might raise broader applicability issues in terms of changing the circumstances of use.

## **C3 Results of pre-modelling studies**

### **INFORMATION REQUESTS**

- Present the results of each pre-modelling study undertaken to address each translation issue specified in Sub-section C1 (and for which a plan is presented in Sub-section C2).
- Provide:
  - copies of all sources of data in an attachment or a technical document, cross-referenced from the main body of the assessment report; and
  - electronic copies of all computer-based analyses.

### **Results**

Where possible and appropriate, present the results of each analysis for which a plan is presented in Sub-section C2 and estimate the comparative treatment effect as results separately for:

- the proposed therapeutic medical service;
- its main comparator; and
- the increment with its 95% confidence interval.

Where a scenario-based valuation study has been used to transform the trial results or any other health state into utility valuations, present these as disaggregated results corresponding to each health state presented as a scenario (see Appendix 4). Also include an estimate of statistical uncertainty around each result.

Discuss the implications of each analysis on the conclusions from the results of the overall clinical evaluation in Sub-sections B6 and B8. Variations in the extent of comparative effectiveness are more likely than variations in the classification of the therapeutic medical service based on Table B8.1.

Where a cross-sectional study or expert opinion survey has been used to estimate patterns of resource provision, report that provision, where possible, on a per patient basis and on a per period of time basis.

Clear presentation of pre-modelling studies is expected to increase MSAC's confidence in the economic evaluations that rely on those translations. At all times in pre-modelling studies, it is important to maximise MSAC's confidence (in the primary inference) that substituting the proposed therapeutic medical service for the main comparator, as proposed alone, causes the differences in the subsequent streams of costs and outcomes. In practical terms, this means that if any stream of costs for a therapy is to be modified in a model, consideration should be given to any consequential impact on the corresponding stream of outcomes. Similarly, if any stream of outcomes for a therapy is to be modified in a model, consideration should be given to any impact on the corresponding stream of costs to ensure that the modification is plausible. Discuss these considerations whenever they are applicable to the results of a particular pre-modelling study.

Justify any results to be used in Section D of the assessment report where more than one option exists. Comment on any uncertainties in this selection, including how they will be assessed in the sensitivity analyses of the economic evaluation. Also comment on any combinations of the results of more than one analytical plan in constructing the economic evaluation and any uncertainties arising from those combinations, including how they will be assessed in the sensitivity analyses of the economic evaluation.

### **Original sources and electronic calculations**

Separately provide copies of the original sources of all data (beyond those already presented with Section B of the assessment report) and reports of studies commissioned for the assessment report in an attachment or technical document. Cross-reference the extraction of data from each source to the level of the page, table or figure number of the source document.

Also, to enable independent verification of each analysis, provide an electronic copy of any computer-based calculations of the analysis.

## **C4 Relationship of each pre-modelling study to the economic evaluation**

### **INFORMATION REQUESTS**

- Discuss the results of each pre-modelling study and explain how they will be used in the economic evaluation (assessment report Section D).
- Provide a summary table of results from Sub-section C3 and their uses in responses to Section D.

### **Uses of pre-modelling study results**

Each pre-modelling study has the objective of providing support for one or more inputs in the economic evaluation. There might be more than one pre-modelling study to support more than one translation step between the overall clinical evaluation and the economic evaluation. When this occurs, the combination of pre-modelling studies might compound the effect of uncertainty. This might need examination in the sensitivity analysis in Sub-section D6.

Section D provides more guidance on how to present the impacts on the economic evaluation of more than one translation step.

### **Summary table**

Table C4.1 provides a suggested format to summarise the main results of each pre-modelling study presented in Section C of the assessment report and their use in the economic evaluation presented in Section D of the assessment report, including in the sensitivity analyses presented in Sub-section D6. This will facilitate cross-referencing across the responses to information requests in the two sections and thus the transparency of the presentation of this information.

**Table C4.1      Summary of results of pre-modelling studies and their uses in the economic evaluation**

<b>Pre-modelling study</b>	<b>Results</b>	<b>Use in Section D</b>	<b>Cross-reference</b>	<b>Use in Sub-section D.6</b>	<b>Cross-reference</b>
Applicability pre-modelling studies					
Study 1					
Etc.					
Extrapolation pre-modelling studies					
Study 2					
Etc.					
Transformation pre-modelling studies					
Study 3					
Etc.					
Other translation pre-modelling studies					
Study 4					
Etc.					

## Section D

### Economic evaluation for the main indication

---

#### Introduction

The purpose of Section D of the assessment report is to present an economic evaluation of substituting the proposed medical service for the main comparator in the context of the listing requested. Requests are made for a full and transparent description of the economic evaluation, as well as for the presentation of sensitivity analyses to demonstrate the robustness of the economic valuation.

As already described in Part II, Section B and shown in Figure D1, the economic evaluation of the proposed medical service initially depends on whether the therapeutic conclusion shows:

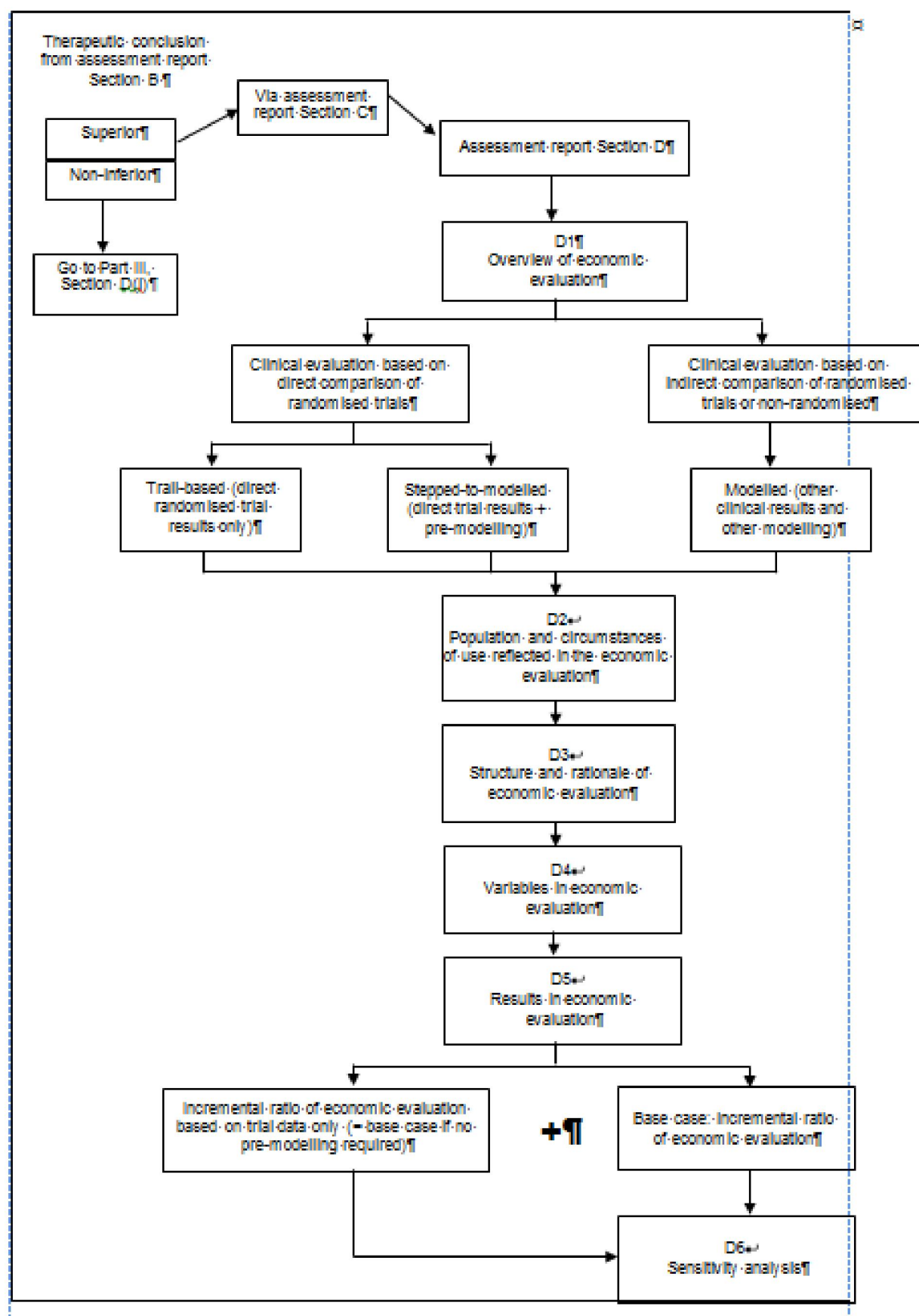
- the proposed medical service is therapeutically superior to the main comparator; or
- the proposed medical service is non-inferior (equivalent) to the main comparator; or
- the proposed medical service is inferior to, but significantly less expensive than, the main comparator.

This Section provides information requests for assessment reports for which there is a therapeutic conclusion of superiority. Information requests for economic evaluations based on a therapeutic conclusion of non-inferiority are provided in Part III, Section D(i).

Furthermore, the approach described in this Section mainly refers to assessment reports where the economic evaluation is either ‘trial based’ (i.e. based on results from direct randomised trials; see Part II, Section B) or ‘stepped-to-modelled’ (i.e. direct randomised trial results with pre-modelling; see Part II, Section C). Thus, it is intended to maximise MSAC’s confidence in an economic evaluation based on this most preferred means of detecting and estimating incremental treatment effects on health outcomes, resource use and cost effects relevant to the requested listing.

For economic evaluations that rely on incremental treatment effects based on results from either indirect comparisons of randomised trials (see Part III, Section B(i)) or comparisons based on non-randomised studies (see Part III, Section B(ii)), consider adapting the stepped approach described here to provide a ‘modelled’ evaluation to improve the transparency of the economic evaluation (see also Part III, Section C(i)).

**Figure D1** Key information requests for assessment report Section D of a standard assessment for MSAC





## D.1 Overview of the economic evaluation

### INFORMATION REQUESTS

- State whether the base case of the economic evaluation is generated by:
  - a trial-based economic evaluation (i.e. based on direct randomised trials presented in Section B of the assessment report)
  - a stepped economic evaluation (i.e. derived from direct randomised trials presented in Section B of the assessment report using variables reported in Section C of the assessment report)
  - a modelled economic evaluation based on an indirect comparison of randomised trials or non-randomised studies.
- State which type(s) of economic evaluation is presented.
- Provide copies of all the original sources of all data or opinion used, and cross-reference the extracted data to the source documents.

### Generation of the base-case economic evaluation

The three steps described below show the approach to an economic evaluation based on a therapeutic conclusion of superiority derived from direct randomised trials.

#### Step 1: Trial-based economic evaluation

The first step involves an economic evaluation based on the unmodified trial-based estimate of treatment effect on incremental provision of health care resources and incremental health outcomes (i.e. using the most internally valid evidence from the direct randomised trials presented in Section B of the assessment report. If the direct randomised trial(s) recruited patients directly representative of those for whom listing is sought, trialled the proposed medical service in the circumstances of use expected to apply to the requested proposed therapeutic medical service if MBS-funded in Australia, and directly measured and reported patient-relevant end points during an appropriate time horizon (i.e. if no pre-modelling studies are reported in Section C of the assessment report), the trial-based evaluation is sufficient to provide the base case of the economic evaluation, and steps 2 and 3 are not required.

#### Step 2: Applying treatment effects on health care resource use if MBS-funded in Australia

Frequently, the results of the direct randomised trials reported in Section B of the assessment report provide insufficient information on which to base a judgment about the full clinical and economic performance of the proposed medical service compared with its main comparator. In these instances, use a modelled economic evaluation to inform MSAC using the results of pre-modelling studies presented in Section C of the assessment report.

The first stage of the economic modelling is to examine the impact of applying the treatment effects on health care resources and health outcomes to the intended proposed medical service population and the circumstances of use identified by the requested restriction (as presented in Section C of the assessment report).

### **Step 3: Extrapolating and transforming health care resource use and health outcomes if MBS-funded in Australia**

The final stage is to examine the additional impact on the modified economic evaluation from step 2 of extrapolating the health care resource use and health outcomes to the time horizon of the economic evaluation and/or any transformation to final outcomes (also presented in Section C of the assessment report). This generates the stepped base case of the economic evaluation for assessment reports that present pre-modelling studies in Section C of the assessment report.

Justify any proposal to reverse the order of steps 2 and 3 (i.e. to extrapolate and/or transform the treatment effect before applying it). In this case, the final step would still generate the base case of the economic evaluation.

Examples of reasons for presentation of a stepped economic evaluation rather than just a study-based analysis include:

- the study population and setting might be different from the target population and setting;
- the outcomes measured in the studies might not be the final outcomes of interest for the proposed service;
- a range of outcomes are of interest;
- the time frame of outcomes measured in the studies might be inadequate; and
- resource-use patterns measured in the studies might not fully reflect those expected in practice (e.g. some resources might not be measured in the studies, and some 'protocol-driven' resources might be included that are not relevant to the proposed provision of the service).

Wherever relevant to information presented in response to requests in this Section, cross-reference to analyses summarised in Sub-section C4 to address the above issues.

#### **Type of economic evaluation**

To identify the most appropriate evaluation, the assessment report should first classify the proposed service using the grid provided in Table D1.1. This classification should be based on the differential effectiveness and safety of the service under consideration compared with the appropriate comparator(s) when used in the target population and setting (i.e. the information presented in Section B of the assessment report). In classifying the service, it might also be necessary to consider changes in the profile of risks associated with the proposed service, compared with the main comparator(s).

In classifying a service, the quality and strength of the available evidence should be taken into consideration. MSAC has a strong preference for making decisions on the basis of data from direct randomised trials and will be most influenced by the results of these types of trials as the most rigorous source of data. However, MSAC has considered and will continue to consider all levels of evidence.

Where there are trade-offs between incremental effectiveness and incremental safety; that is, where there is reduced effectiveness but improved safety (see Table D1.1) or improved effectiveness but reduced safety, consideration will be required as to whether there are net clinical benefits or net harms to patients, overall. This might involve a valuation of the different effects associated with a service and/or modelling of various outcomes.

Assumptions made in reaching the conclusion about whether a service has net clinical benefits should be stated explicitly.

**Table D1.1 Classification of a service under MSAC consideration/Classification of the effectiveness of the proposed medical service over its main comparator and guide to the suitable type of economic evaluation**

Comparative safety	Comparative effectiveness			
	Inferior	Uncertain <sup>a</sup>	Non-	Superior
<b>Inferior</b>	Health forgone: need other supportive factors	Health forgone possible: need other supportive factors	Health forgone: need other supportive factors	? Likely CUA
<b>Uncertain <sup>a</sup></b>	Health forgone possible: need other supportive factors	?	?	? Likely CEA/CUA
<b>Non-</b>	Health forgone: need other supportive factors	?	CMA	CEA/CUA
<b>Superior</b>	? Likely CUA	? Likely CUA	? Likely CEA/CUA	CEA/CUA

CEA = cost-effectiveness analysis; CMA = cost-minimisation analysis; CUA = cost-utility analysis

? = reflects uncertainties and any identified health trade-offs in the economic evaluation, as a minimum in a cost-consequences analysis

**a** 'Uncertainty' covers concepts such as inadequate minimisation of important sources of bias, lack of statistical significance in an underpowered trial, detecting clinically unimportant therapeutic differences, inconsistent results across trials, and trade-offs within the comparative effectiveness and/or the comparative safety considerations (e.g. where the safety profiles of the compared medical services differ, with some aspects worse for the proposed medical service and some aspects better for the proposed medical service).

**b** An adequate assessment on 'non-inferiority' is the preferred basis for demonstrating equivalence.

### Non-inferior (equivalent) service

If the proposed medical service has been shown to be non-inferior (equivalent) to the main comparator, a cost-minimisation analysis is appropriate (or cost analysis under limited circumstances where the proposed medical service is non-inferior to the main comparator, but has a superior safety profile that generates cost offsets from reduced use of health care resources to manage adverse reactions). Part III, Section D(i) provides the information requests associated with these evaluations.

A cost-minimisation analysis should only be presented when the proposed service has been indisputably demonstrated to be no worse than its main comparator(s) in terms of both effectiveness and safety, so the difference between the service and the appropriate comparator can be reduced to a comparison of costs. In most cases, there will be some uncertainty around such a conclusion (i.e. the conclusion is often not indisputable). Therefore, when an assessment report concludes that an intervention was no worse than a comparator, an assessment of the uncertainty around this conclusion should also be provided by the presentation of cost-consequences, cost-effectiveness and/or cost-utility analyses.

### Therapeutically superior service

If the proposed medical service has been shown to be therapeutically superior to the main comparator, there are four types of economic evaluation that might apply, depending on the outcome of the clinical evidence (see Table D1.1):

### ***Cost-utility analysis (generally preferred)***

A cost-utility analysis presents the health outcome in terms of the life-years gained from the start of the analysis, with each life-year adjusted by a utility weight that represents society's preferences for the health outcome experiences in that life-year relative to full health. The ultimate benefit of restored health is the restoration of health-related quality of life; for example, restoration of opportunities to undertake activities of daily living. Economists have attempted to identify the value placed by individuals on different health states. The basis for this valuation is that each increment in health-related quality of life gives satisfaction (measured as the strength of preference for the restored health over the pre-treatment state of health and termed 'utility' by economists), which is the ultimate outcome of life. The denominator in a cost-utility analysis is most commonly the incremental QALY gained, which is the difference between the two profiles following the use of the proposed medical service or its main comparator, each calculated as the times spent in successive varying health states, with each period of time weighted by the strength of preference for, or the utility weight of, its respective health state (see Appendix 5 for further guidance on valuing health outcomes in utility terms).

### ***Cost-effectiveness analysis***

A cost-effectiveness analysis measures the incremental cost per extra unit of health outcome achieved. It differs from a cost-utility analysis in that the health outcome is reported in its natural units. If the proposed medical service is demonstrated to offer more of a given health outcome than its main comparator (e.g. it achieves the desired health outcome in a higher proportion of patients), this goes beyond cost-minimisation. The outcomes reported from the clinical evaluation might need to be transformed in a modelled cost-effectiveness analysis; where this is done; the choice of outcome should be justified.

### ***Cost-benefit analysis (supplementary option)***

A cost-benefit analysis expresses all outcomes (health and non-health) valued in monetary rather than natural or utility units. This is in contrast to other forms of economic evaluation and requires a monetary valuation of these outcomes (see Section A6.2 of Appendix 6). Cost-benefit analysis can also include both health and non-health outcomes.

### ***Cost-consequences analysis (if disaggregation of outcomes would be helpful)***

A cost-consequences analysis compares the incremental costs of the proposed medical service over its main comparator with an array of outcomes measured in their natural units rather than a single representative outcome as presented in a cost-effectiveness analysis. It can be presented if the proposed medical service is demonstrated to have a different profile of effects that are not adequately captured by a single outcome measure; there might be trade-offs between the two therapeutic medical services in terms of the directions of the changes in effectiveness and safety (and within effectiveness and safety). As such, it is a form of disaggregated analysis of changes in patterns of health care resource provision and changes in health outcomes, and can be presented before presenting other types of aggregated economic evaluation, such as a cost-utility analysis (see footnotes to Table B8.1 and general guidance below).

Table D1.2 shows the type of economic evaluation that should be presented for each classification from Table D1.1.

**Table D1.2 Type of economic evaluation that should be presented for various classifications of a service under MSAC consideration**

Classification	Type of economic evaluation
The service is more effective than the appropriate comparator and is associated with improved safety.	Cost-consequences, cost-effectiveness, cost-utility, cost-benefit
The service is more effective than the appropriate comparator and is no worse than the comparator in terms of safety.	Cost-consequences, cost-effectiveness, cost-utility, cost-benefit
<p>The service is more effective than the appropriate comparator but is associated with reduced safety.</p> <p>(i) Overall, there are net benefits to patients as the benefits from improved effectiveness outweigh the harms from reduced safety and/or changed risk profile.</p> <p>(ii) Overall, the service is no worse than the comparator because the benefits from improved effectiveness at least offset the harms from reduced safety and/or changed risk profile.</p> <p>(iii) Overall, there are net harms to patients as the harms from reduced safety and/or changed risk profile outweigh the benefits from improved effectiveness.</p>	<p>Cost-consequences, cost-effectiveness, cost-utility, cost-benefit</p> <p>Cost-consequences, cost-effectiveness. This may be reducible to cost-minimisation (i.e. presentation of an incremental cost-effectiveness for the base case may be inappropriate if net clinical benefits are assumed to be zero)</p> <p>No economic evaluation needs to be presented; MSAC is unlikely to recommend Government subsidy of this service.</p>
The service is no worse than the comparator in terms of effectiveness but is associated with improved safety.	Cost-consequences, cost-effectiveness, cost-utility, cost-benefit
The service is indisputably demonstrated to be no worse than the comparator in terms of both effectiveness and safety.	Cost-minimisation. In the case where there is any uncertainty around the conclusion that the service is no worse than the comparator in terms of effectiveness and safety, cost-consequences, cost-effectiveness, and/or cost-utility analyses should be provided.
The service is no worse than the comparator in terms of effectiveness but is associated with reduced safety.	No economic evaluation needs to be presented; MSAC is unlikely to recommend Government subsidy of this service.
<p>The service is less effective than the comparator but is associated with improved safety.</p> <p>(i) Overall, there are net benefits to patients as the benefits from improved safety and/or changed risk profile outweigh the harms from reduced effectiveness.</p> <p>(ii) Overall, the proposed service is no worse than the comparator because the benefits from improved safety at least offset the harms from reduced effectiveness and/or changed risk profile.</p> <p>(iii) Overall, there are net harms to patients as the harms from reduced effectiveness outweigh the benefits from improved safety and/or changed risk profile</p>	<p>Cost-consequences, cost-effectiveness, cost-utility, cost-benefit</p> <p>Cost-consequences, cost-effectiveness (which may be reducible to cost-minimisation i.e. presentation of an incremental cost-effectiveness for the base case may be inappropriate if net clinical benefits are assumed to be zero)</p> <p>No economic evaluation needs to be presented; MSAC is unlikely to recommend Government subsidy of this service</p>
The proposed service is less effective than the comparator and is no worse than the comparator in terms of safety	No economic evaluation needs to be presented; MSAC is unlikely to recommend Government subsidy of this service
The proposed service is both less effective than the comparator and is associated with reduced safety compared with the comparator	No economic evaluation needs to be presented; MSAC is unlikely to recommend Government subsidy of this service

MSAC = Medical Services Advisory Committee

From Table D1.2, it can be seen that an economic evaluation should be presented in all assessment reports to be considered by MSAC except when a service is indisputably demonstrated to be associated with net clinical harms to patients (as it is unlikely that MSAC will recommend Government subsidy of the service).

The assessment report should state what type of economic evaluation is being presented. All analyses should explicitly consider all the advantages and disadvantages of the proposed service that are listed in the clinical balance sheet, compared with the comparator. However, there are some circumstances where simplified analyses will be appropriate and acceptable (see Sub-section D3 for further details).

An iterative approach to the classification and type of analysis might be required. For example, a valuation of the different effects associated with a service and/or modelling of various outcomes might be required before a service can be definitively classified according to Table D1.1. In these cases, the structure of the economic evaluation and the assumptions made in valuation of outcomes must be presented clearly. Adequate sensitivity analysis should also be provided to allow MSAC to gauge the robustness of the classification selected. Thus, although the service might ultimately be classified as being no worse than the comparator (e.g. where improved effectiveness is considered to offset reduced safety), such that a cost-minimisation analysis is considered appropriate, a cost-consequences and a cost-effectiveness analysis that explicitly shows the valuation of the various outcomes should also be presented. Sensitivity analyses should also be presented which examine the effect of varying assumptions in the valuation of outcomes.

Note that the various types of analyses should not be considered mutually exclusive. In many cases it will be appropriate for more than one type of analysis to be presented. As discussed in Sub-section D3, a stepped economic evaluation is requested. Such an analysis will typically start with a cost-consequences analysis and will progress, where appropriate, through various steps where various aspects of modelling are introduced such that, ultimately, a base-case cost-effectiveness or cost-utility analysis is presented. A trade-off between the most appealing outcome upon which to base the economic evaluation from a theoretical point of view and the degree of uncertainty in the estimate of incremental cost-effectiveness is often required. Extrapolation of outcomes beyond the evidence will introduce uncertainty in estimates of incremental cost-effectiveness. For example, the estimate of incremental cost-effectiveness generated by a study-based analysis (i.e. based directly on the outcome from a study) might be relatively robust. However, in moving to a cost-utility analysis (which is theoretically more appealing but where assumptions of utilities for various health states might be required), additional uncertainty might be introduced.

The common output of these evaluations is a comparison of changes in outcomes and changes in costs of achieving those outcomes across the proposed medical service and the main comparator. The objective is usually to justify a price advantage for the proposed medical service compared to its main comparator. A statistically significant improvement in effectiveness alone is not necessarily sufficient to support a conclusion of acceptable cost-effectiveness. Consideration is also given to whether the detected differences are clinically important overall and whether the extent of improvement is sufficient to justify any requested price advantage (after accounting for any justified cost offsets).

## **General guidance on preferred and supplementary types of economic evaluation**

The various types of economic evaluation are not necessarily mutually exclusive and it might be appropriate to present more than one type (e.g. both cost-effectiveness and cost-utility analyses). Depending on the circumstances, there might be a trade-off between the most appealing approach from a theoretical point of view and the degree of uncertainty in the estimate of incremental cost-effectiveness. For example, estimating the incremental cost-effectiveness based directly on the outcome from a trial might be relatively robust. However, additional sources of uncertainty might be introduced when moving to a cost-utility analysis. (A cost-utility analysis is theoretically easier to interpret and compare across assessment reports and medical conditions, but it might require assumptions of utility weights for various health states.) The three steps described in the beginning of this Sub-section to improve transparency for economic evaluations are designed to help make these trade-offs and their implications explicit.

Given these considerations, a cost-utility analysis is the preferred form of economic evaluation for either or both of the following situations:

- where there is a claim of incremental life-years gained in the economic evaluation — to assess the impact of quality adjusting that survival gain; and
- where relevant direct randomised trials report results using a MAUI.

However, for the reasons given above, the preference for a full cost-utility analysis is less clear in other situations, even where there is a claim of quality-of-life or disability improvements, or where there are differential quality-of-life impacts arising from the therapies being compared in an assessment report to derive a common outcome across assessment reports. Therefore, in the situation of an improvement in quality of life but not in quantity of life, an assessment report should present a cost-utility analysis or justify the decision to not transform the quantified health outcomes via a utility valuation.

Cost-benefit analysis is not preferred because it is not likely to be helpful to most MSAC deliberations (further reasons are given in Appendix 5). Thus, although monetary valuation of health outcomes is allowed, it is considered to be supplementary to utility valuation presented in a cost-utility analysis. If a cost-benefit analysis is presented in the absence of a cost-utility analysis, MSAC might not consider it to have the same weight.

Similarly, the base-case economic evaluation should be focused on material incremental changes in the provision of health care resources and on material incremental changes in health outcomes. Supplementary analyses can be used to present any material incremental changes in the provision of non-health care resources and/or in non-health outcomes.

### **Sources of information**

Separately provide copies of the original sources of all data (beyond those already presented in Sections B and C of the assessment report) or expert opinion used in the model in an attachment or a technical document. Cross-reference the extraction of data from each source to the level of the page, table or figure number of the source document.

## D2 Population and circumstances of use reflected in the economic evaluation

### INFORMATION REQUESTS

- Describe and justify the demographic and patient characteristics of the population included in the economic evaluation.
- Describe and justify the circumstances in which the proposed medical service and main comparator are used in the economic evaluation.
- Assess the consistency of the demographic and patient characteristics and of the specified circumstances of use across the study populations, the population in the economic evaluation and the population for whom listing is sought.

In this Section of an application or assessment report, analysts should provide information to allow MSAC to assess whether the evidence presented is applicable and generalisable to the population and circumstances of use for whom the service is proposed (see Table D2.1).

**Table D2.1** Definitions for populations and circumstance of use that should be taken into account in the evaluation

Term	Description
<b>Target</b> population and circumstances of use	Population and setting for which Government subsidy of the service is being requested
<b>Study</b> population and circumstances of use	Population and setting for which evidence of efficacy and safety has been presented in assessment reports Sections B and/or C
<b>Wider</b> population and circumstances of use	Broader population and setting in which the service is likely to be used if MBS-funded

### Population (demographic and patient characteristics)

Use summary statistics (where appropriate) to describe the demographic and clinical characteristics for the population entering the economic evaluation. Include information about the distribution around means where appropriate.

Examples of patient characteristics are provided in Section A.

Use cross-references, as appropriate, to Section A when justifying the definition of each characteristic of the population in the economic evaluation in relation to the population for whom listing is sought. Also highlight any difference in relation to the study populations for whom evidence of effectiveness and safety are presented (using cross-references, as appropriate, to Sub-section C4 if pre-modelling studies are presented to apply these results).

### Circumstances of use

Use cross-references, as appropriate, to Section A when describing and justifying the definition of each circumstance of use (setting) assumed in the economic evaluation in relation to the medical condition under which listing is sought. Also highlight any difference in relation to each circumstance for which evidence of effectiveness and safety is presented from the studies (using cross-references, as appropriate, to Sub-section C4 if pre-modelling studies are presented to apply these results).



The application or assessment report should describe the setting in which the service and its main comparator(s) are assumed to be used in the economic evaluation. Examples of elements of settings that could be detailed include:

- the position of the service in the overall algorithm for diagnosing, treating or managing the disease or condition (e.g. prevention, first-line treatment, second-line treatment);
- any limitations on the duration or frequency of delivery of the services; for example, in a 24-hour or in a 12 or 24-month period;
- any required co-delivered medical services or treatments (including any additional diagnostic tests required);
- any contra-indicated medical services or treatments;
- any unique characteristics of the referrer or provider (e.g. specific qualifications or training); and
- any specific requirements in terms of geography, facilities or location of delivery of service (e.g. limited to hospital setting or to approved laboratories; specification of any specific equipment or facilities that need to be available).

### **Consistency across characteristics**

Assess the degree of consistency of the demographic and patient characteristics and of the specified circumstances of use across:

- the study populations and circumstances of use described in Sub-section C4 if pre-modelling studies are presented to apply the results of these trials);
- the target population and circumstances of use, which should reflect the clinical management algorithms presented in Section A; and
- the wider population and circumstances.

The population for whom funding is being examined might be less well defined than the other two groups. However, its inclusion captures the potential for use of the proposed medical service in a broader population and/or broader circumstances than the target population and circumstances if the proposed medical service were MBS-funded in Australia. Including the population might also be useful for capturing any limitations of the economic evaluation in truly replicating the target population and circumstances. The importance of examining the incremental cost-effectiveness of the proposed medical service in this population increases with increasing risk of substantial use of the proposed medical service beyond the intention of the requested restriction (see also Sub-section D6).

Table D2.2 suggests a format that will summarise these characteristics and circumstances for which sensitivity analysis shows that the variable is important.

**Table D2.2 Comparison of characteristics of trial and requested populations and circumstances of use**

Population and circumstance <sup>a</sup>	As defined in trial(s) using ITT population	As defined by the requested restriction	If use beyond the requested restriction might arise
Medical condition of the population			
Age of the population			
Restriction criteria (including any limitations on disease severity, preconditions or previous treatments, or continuation rules)			
Limitations on response or surgical experience considerations of use of proposed medical service			
Repeat for each other variable that varies across these populations and circumstances, and for which sensitivity analysis shows the variable is important			

ITT = intention to treat

<sup>a</sup> For each identified population characteristic and circumstance of use, provide a footnote explaining any differences between these populations and relate this to any pre-modelling study presented in Section C to apply the evidence from the overview of the trial(s) to the requested restriction.

## Justifying restrictions

In the case where it is proposed that eligibility for a service be restricted to a subgroup of patients with a clinical condition, the proposed restriction should be justified as follows:

- The intention of the requested restriction should be indicated in the assessment report.
- To help minimise usage beyond the intention of the requested restriction, for each population or setting element included in the wording of a restriction, the assessment report should:
  - identify and define the element unambiguously; for example:
    - risk factors associated with the medical condition;
    - markers of severity or progression of the medical condition; and
    - name of service and duration criteria for previous medical services, as appropriate.
  - specify objective criteria in preference to subjective criteria in identifying the element;
  - justify any thresholds within these criteria (these thresholds and justifications should be consistent with study eligibility criteria and subgroup stratification criteria as appropriate); and
  - resolve copyright issues about any proposed medical service before proposing its use as part of a restriction.

Please Note: The assessment report should present a discussion addressing the trade-offs between the clinical preference for simple, unambiguous listings versus increasingly complex restrictions designed to limit new services to those relatively few patients for whom the proposed service might be justified as being acceptably cost-effective at the price requested.

The further the eligibility criteria specified in a restriction shift practice away from otherwise uninfluenced practice, the more incentive there is for referrers/providers and patients to seek subsidy despite the restriction. The approach listed above (identifying and justifying any restrictions) is intended to help justify the choice of restriction from the alternative options that might apply. This approach becomes more important as the restriction becomes more complex or more expensive for Government to administer.

If the proposal is for eligibility for a service to be restricted to a subgroup of patients with a clinical condition, the potential for use of the service in a wider population or setting than the target population and setting, if Government subsidy of the service is recommended, this should also be assessed.

### **Presenting the information**

Table D2.3 shows a hypothetical example where it is proposed that a new treatment be made available as a second-line agent for the management of adults with hyperthyroidism to provide a suggested format for presentation of information about the target, study and wider populations and settings.

Where there are differences, or potential differences, between any of the groups, economic analyses should be presented for each of the scenarios.

When presenting economic evaluations for different populations, the assessment report should consider whether changes in the population have implications for the cost associated with the proposed service (e.g. if economies of scale might be captured by using a service in a wider population). Further advice is provided in Section D4.

**Table D2.3**      **Example of a comparison of the characteristics of target, study and wider populations and settings**

<b>Population</b>	<b>Target</b>	<b>Study</b>	<b>Wider</b>
<b>Clinical condition</b>	<b>Hyperthyroidism due to any cause</b>	<b>Hyperthyroidism due to Graves' disease</b>	<b>Hyperthyroidism due to any cause</b>
Comment	Only patients with hyperthyroidism due to Graves' disease were recruited to the only direct randomised trial comparing service A with service B (Jones et al 2000), but subsidy is requested for all patients with hyperthyroidism, regardless of aetiology. Smaller, non-comparative studies (Brown et al 1995, Smith et al 1997) have examined the efficacy and safety of Service A in patients with hyperthyroidism due to other causes. The effect size observed in these studies was similar to that observed in Jones et al, 2000; however, it is acknowledged that <etc..>.		
<b>Age</b>	<b>Adults</b>	<b>18–75 years</b>	<b>Adults</b>
Comment	Although only patients aged up to 75 years were eligible for entry to the direct randomised trial comparing service A with service B (Jones et al 2000), service A has been used in patients over the age of 75 with similar effects as in other adult populations (Smith et al 1990) <etc..>.		
<b>Gender</b>	<b>70% females 30% males</b>	<b>50% females 50% males</b>	<b>70% females 30% males</b>
Comment	Although the proportion of females with condition X recruited to the trial reported by Jones et al 2000 was lower than the proportion of females with hyperthyroidism in the Australia, a test for interaction did not demonstrate gender to be a treatment effect modifier <etc..>.		
<b>Initiation criteria</b>	<b>Serum TSH &lt; 70% x Serum T<sub>3</sub> &gt; 120% y</b>	<b>Serum TSH &lt; x Serum T<sub>3</sub> &gt; y</b>	<b>Serum TSH &lt; 85% x Serum T<sub>3</sub> &gt; 110% y</b>
Comment	Subsidy of service A is requested for a more severely affected population than recruited to the trial reported by Jones et al (2000). Subgroup analysis demonstrates serum TSH and T <sub>3</sub> levels at baseline to be a treatment effect modifier, with a greater relative response rate to service A in patients with levels of serum TSH below 70% x and levels of serum T <sub>3</sub> greater than 120% y. It is acknowledged that there might be some use beyond the population for whom subsidy of service A is sought. Thus, sensitivity analyses are presented examining the effect on incremental cost-effectiveness and financial implications of use of the service beyond the population for whom subsidy is sought.		
<b>Position in management algorithm</b>	<b>Second line</b>	<b>Second line</b>	<b>Second line but some first-line use</b>
Comment	Consistent with the direct randomised trial (Jones et al 2000) comparing service A with service B, subsidy is proposed for use of service A only in patients failing to respond to service C. However, it is acknowledged that there might be some use of service A in the first-line management of hyperthyroidism (i.e. as a substitute for service C instead of service B). Thus, cost-effectiveness analysis is also presented versus service C.		
<b>Limitations on frequency of use</b>	<b>Patients will be permitted to receive service A as a subsidised service on two separate occasions</b>	<b>Patients were permitted to receive service A on two separate occasions</b>	<b>Patients will be permitted to receive service A as a subsidised service on two separate occasions</b>
Comment	The number of times the service might be delivered to the patients on a subsidised basis is consistent with the number of times the patients were able to receive the service in the clinical trial reported by Jones et al (2000).		

## **D3 Structure and rationale of the economic evaluation**

### **INFORMATION REQUESTS**

- Review the relevant economic literature and present the results.
- Specify any software used to conduct the economic evaluation.
- Ensure that all variables in the electronic copy of the economic evaluation can be changed independently during the evaluation, including allowing the base case of the economic evaluation to be completely respecified and allowing a new set of sensitivity analyses to be conducted with each respecified base case.
- Describe the structure of the economic evaluation.
- Justify the appropriateness of the structure in reflecting the context of use of the compared alternatives and the outcomes of their use.
- Define and justify the time horizon and nature of the outcomes used in the economic evaluation.
- Describe the methods used to calculate the results of the economic evaluation (e.g. cohort expected value analysis, Monte Carlo simulation).
- Provide copies of identified papers in an appropriately labelled attachment separate from the main body of the assessment report.

By definition, the economic evaluation is intended to inform a decision. Therefore, the structure of the evaluation allows the comparison of the streams of outcomes and resources following the use of either the proposed medical service or its main comparator to calculate incremental outcomes and costs of these streams. MSAC has a preference for a decision-analytical framework that clarifies the comparison of these streams of outcomes and resources.

### **Literature review**

Applicants should search the literature for published cost-effectiveness analyses of the proposed service. A list of all of the published reports that are retrieved by the search should be provided in the application or assessment report.

The economic analyses that are directly relevant to MSAC's considerations (i.e. economic evaluations performed for the same population and setting in which the service will be used) should be identified using a tight set of inclusion and exclusion criteria, which should be detailed in the application or assessment report. The application or assessment report should also provide a critical review of the included studies.

An independent economic evaluation might not be required if there is already a high-quality economic evaluation in the public domain that provides an estimate of incremental cost-effectiveness for the proposed service in a population and setting that is similar to the proposed Australian population and setting. Such an evaluation needs to be based on the appropriate:

- therapeutic and management setting;
- patient population; and
- input variables.

In these circumstances, an assessment of the most appropriate publicly available evaluation should be presented in the report for MSAC according to the requirements of these Guidelines. That is, the evaluation available in the public domain should be assessed according to this Sub-section, and Sub-sections D4 and D5. All details requested in these Sections should be provided in the assessment report.

Where a model in the public domain is considered to have an appropriate structure, but is populated with values for variables that do not correspond to the values that would apply in the Australian population and setting or proposed by the PICO Confirmation (see Section A), it might be appropriate to use the model, but to update values for the variables to values that would apply in the Australian context. Again, the model should be assessed according to this Sub-section, and Sub-sections D4 and D5. All of these Guidelines and all details requested in these Sections should be provided in the assessment report.

If a search of the literature fails to identify any directly relevant economic evaluations, an independent economic evaluation should be conducted. This Sub-section, and Sub-sections D4 and D5 of these Guidelines describe the information required and how the economic evaluation should be presented.

Present the results of a search of the literature for reports of economic evaluations of similar decision analyses (in terms of similarity to the treatment algorithm and/or the proposed and similar medical services). Where the assessment report's model is different from the literature-sourced models, explain the basis for the selection of the assessment report's approach.

### **Software package**

Specify the name and version of any software package used to conduct the economic evaluation. Software packages that support decision analyses and can be readily critiqued currently consist of:

- TreeAge Pro Suite®;
- Excel 2010-2013®, including @RISK®, but not necessarily including all advanced features and plug-ins (e.g. Crystal Ball® and customised macros developed using Visual Basic); and
- STATA.

Economic evaluations constructed using any of these may be submitted without earlier arrangement with the HTA Team. For further information, please refer to Section 5 of these Guidelines.

### **Fully accessible electronic copy of the economic evaluation**

Ensure that all variables in the electronic copy of the economic evaluation can be changed independently, including allowing the base case of the economic evaluation to be completely respecified and allowing a new set of sensitivity analyses to be conducted with each respecified base case.

## Structure of the economic evaluation

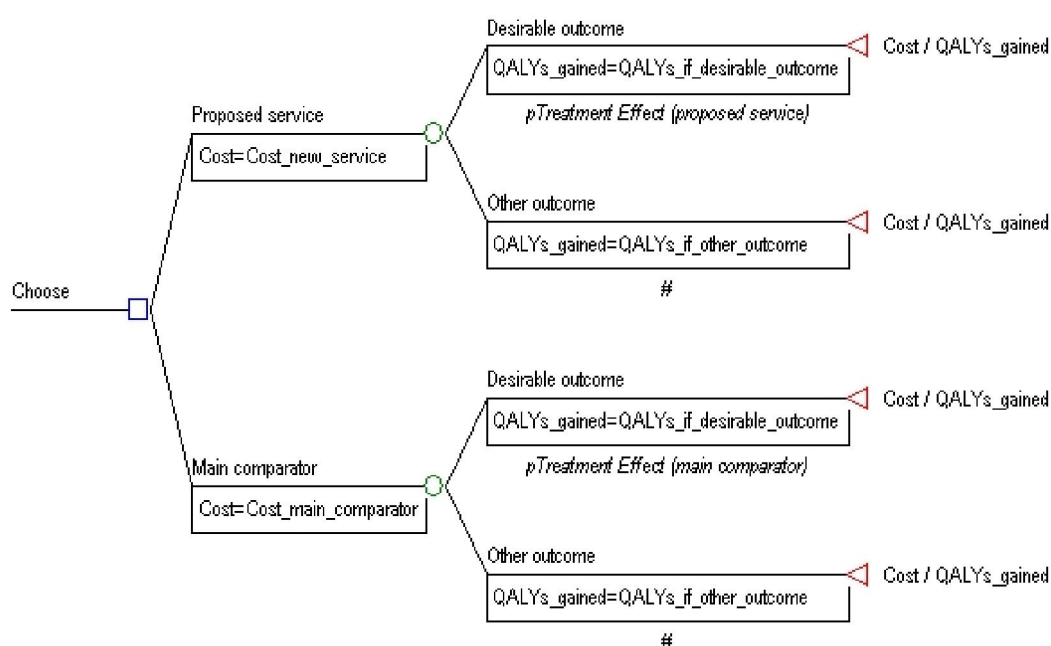
The description of the economic evaluation should include:

- a statement defining in detail the therapy options for which costs and outcomes are estimated in the economic evaluation;
- a description of each of the types of event and health states possible in the economic evaluation, together with a justification of the selection of each health state for inclusion in the evaluation and a justification for those that were considered potentially suitable but that were excluded to avoid excessive complexity;
- a description of the relationships and interactions between the various events and health states possible in the economic evaluation (including, where relevant for a state transition model, a detailed description of all possible transitions between the health states; see below);
- a description of all assumptions made in the construction of the economic evaluation; and
- a decision tree diagram summarising the structure of the economic evaluation.

## Study-based evaluation

If the study population and setting are the same as the target population for the proposed service, and outcomes have been reported for all patient-relevant endpoints, it might be appropriate to present a simple economic evaluation based directly on the results of the included studies. The structure of a basic economic evaluation is shown in Figure D3.1.

Figure D3.1 Example of the structure of a basic economic evaluation



## **Stepped economic evaluation**

Frequently, the results of the available studies provide insufficient information on which to base a judgement about the clinical and economic performance of the proposed service relative to that of the comparator. In these circumstances (which are a matter of judgement), a stepped economic evaluation (which introduces the various aspects of modelling in separate steps) will be useful to MSAC. Examples of reasons for presentation of a stepped economic evaluation rather than just a study-based analysis include:

- the study population and setting might be different to the target population and setting;
- the outcomes measured in the studies might not be the final outcomes of interest for the proposed service;
- a range of outcomes are of interest;
- the time frame of outcomes measured in the studies might be inadequate; and
- resource-use patterns measured in the studies might not fully reflect those expected in practice (e.g. some resources might not be measured in the studies, and some 'protocol-driven' resources might be included that are not relevant to the proposed provision of the service).

### ***Presenting a stepped evaluation***

To ensure that the manner in which available information is incorporated into the economic evaluation is transparent, MSAC requires the presentation of a stepped economic evaluation that starts with a study-based cost-consequences analysis and progresses through various steps of the modelling in turn (population and setting, outcome, time horizon, resource use, etc.). These steps might require the presentation of additional evidence. Guidance for the presentation of this evidence is provided in Section C.

MSAC recognises that the conduct of a complex economic evaluation for a service might be associated with costs that could exceed the costs of actually providing that service. Therefore, a simple economic evaluation, such as a study-based economic evaluation or a simplified model, is acceptable if the following criteria are both met:

- the service is likely to be used by small numbers of patients; and
- the total Government expenditure on the service is likely to be small.

To ensure consistency across economic analyses considered by MSAC, the preferred elements of a base-case economic evaluation are summarised in Table D3.1.



**Table D3.1 Key elements of the base-case economic evaluation**

Element of economic evaluation	MSAC's preference for the base-case analysis	Section providing further details
Perspective	Societal perspective. However, costs and benefits should be presented aggregated to the following three levels: <ul style="list-style-type: none"> <li>• taking an MBS (or other relevant Government program) perspective (i.e. including costs and benefits incurred by the MBS);</li> <li>• taking a health care perspective (i.e. including only costs related to provision of health care resources regardless of who incurs them, and including only health outcomes); and</li> <li>• taking a societal perspective (i.e. including all costs and benefits).</li> </ul>	Section D
Comparator	Currently available service that is most likely to be replaced by the new service	Section A
Type of economic evaluation	Cost-effectiveness analysis	Section D
Source of evidence	Systematic review	Section C
Values of parameters	Unbiased, plausible estimates. Where there is room for judgement and considerable uncertainty around the value of a parameter, a conservative approach to the valuation of that parameter should be adopted.	Section D
Outcome on which evaluation should be based	The outcome measure that most closely and validly estimates the final health outcome from a patient perspective. Health-related QALYs should be used where feasible.	Section D
Discount rate	An annual rate of 5% for both costs and benefits	Section D

A description of the structure of each step of the economic evaluation should be provided, and include:

- an explicit statement of the options for which costs and benefits are being estimated in the economic evaluation, and the justification for the selection of options included in the evaluation;
- a description of each of the events and health states possible in the economic evaluation;
- justification of the selection of health states for inclusion in the economic evaluation (and those excluded to avoid excessive complexity);
- a description of the relationships and interactions between the various events and health states possible in the economic evaluation (including detail of the transitions possible between the health states);
- a description of assumptions (both implicit and explicit) made in the construct of the economic evaluation; and
- a decision-tree diagram summarising the structure of the economic evaluation.

The application or assessment report should present a justification for the overall structure of the base-case economic evaluation, particularly in relation to:

- the natural history of the condition being managed, prevented or diagnosed;
- the management algorithm that applies currently and the management algorithm that will apply should the service be MBS-listed;
- the management algorithm that applied in the studies used as evidence to demonstrate the safety and effectiveness of the proposed service; and
- the structure of other relevant models reported in the public domain.

The report should also identify and consider assumptions built into the structure of the economic evaluation and comment as appropriate.

### ***Defining and justifying the time horizon***

The time horizon over which costs and benefits of a service and its comparator are measured in each step of the evaluation should be defined and justified. The assessment report should define and justify the time points at which events are assumed to occur and the duration of time spent in health states (include details of cycle length for Markov models). The appropriate time horizon for follow-up will relate to the natural history of the disease, the treatment pattern and the time period over which outcomes from the service or main comparator could be expected to occur. For example, the time horizon over which costs and health benefits of a diagnostic test for an acute event (e.g. a nonlife-threatening infection) might be relatively short, whereas the appropriate time horizon to consider for a treatment for a chronic illness will be longer.

### ***Discounting***

Where costs and benefits of a service and/or its comparator are presumed to be borne over more than one year, the present value of future costs and benefits should be used in the economic evaluation. This means that discounting should be applied to both costs and benefits sustained in the period beyond the first year. Costs and benefits should be discounted at an annual rate of 5%. As discussed in Sub-section D6, a sensitivity analysis examining the impact of discounting should be performed.

### ***Describing the methods used***

The methods used to generate results of the economic evaluation should be described; for example:

- expected value analysis (or cohort analysis);
- Monte Carlo simulation (the assessment report should specify whether first-order and/or second-order distributions are sampled); and
- Markov models (the assessment report should specify whether a half-cycle correction has been included or justify its exclusion).

## ***Dealing with uncertainty***

The value of information from a complex economic evaluation diminishes as greater uncertainties are introduced through the process of modelling. The application or assessment report should consider the extent to which the value of more extensive analysis will be limited by the quality of the underlying data and the extent to which uncertainties in the clinical evidence will be amplified by modelling. Progression through modelling steps should continue only as long as the results generated are likely to be of value and informative to MSAC.

The type of presentation that is likely to be of greatest value to MSAC might vary with the level of evidence available. For example, in some circumstances, the evidence base might be extremely weak (e.g. where a claim that a service is safe and ‘promising’ in terms of effectiveness is based on low-level evidence, such that the claim cannot yet be considered proven). In such cases, a threshold analysis that examines incremental cost-effectiveness over a range of possible benefits, and that essentially seeks to determine the minimum extent of benefit that would be required for the service to be considered acceptably cost-effective, might be more informative than reporting of an incremental cost-effectiveness ratio based on a single point-estimate of incremental effectiveness.

The objective of cost-effectiveness analysis should be to provide an unbiased, plausible estimate of the incremental cost-effectiveness of the medical services being compared. Where an element of judgement is required, and where there is considerable uncertainty around the value of a parameter, a conservative approach to the assignment of a value to that parameter should be adopted for inclusion in the base case.

## ***Justification of the structure***

Justify the overall structure of the economic evaluation in relation to the current and proposed clinical management algorithms (and the requested restriction, as appropriate) presented in Section A of the assessment report, and the treatment algorithms used in the studies presented (using cross-references, as appropriate, to Sections B and C of the assessment report). When justifying the overall structure of the economic evaluation in relation to the current and proposed clinical management algorithms, discuss the consistency across:

- the alternative therapy options examined in the economic evaluation and those considered appropriate in response to Sub-section A5;
- the clinical management algorithms assumed in the structure of the economic evaluation before and after the implementation of the requested listing and the algorithms presented in response to Sub-section A5; and
- the clinical management algorithms assumed in the structure of the economic evaluation and the clinical management algorithms for which clinical evidence is presented in Sections B and C of the assessment report.

Identify and consider implicit assumptions built into the structure of the economic evaluation and comment as appropriate.

## **Time horizon and outcomes used in the evaluation**

### ***Time horizon***

Define and justify the time horizon over which the costs and outcomes of the proposed medical service and its main comparator are estimated in the economic evaluation. The appropriate time horizon for follow-up relates to the natural history of the medical condition, the treatment patterns, and an estimation of the time period(s) over which outcomes from the two therapies would be expected to occur. For example, a relatively short time horizon could apply when treating an acute event, whereas a longer time horizon would be required for a chronic illness.

### ***Outcomes***

Indicate whether the outcomes generated by the economic evaluation represent the final outcomes of treatment. Where the economic modelling structure is used (rather than a separate pre-modelling study, see Section C) to transform a quantified treatment effect measured on a surrogate outcome in the trials to predict a subsequent quantified treatment effect on the intended final outcome, explain and justify the method of this transformation, including a justification for how the relationship might vary over time. Use a pre-modelling study to show that a systematic approach has been taken to select and justify the modelling approach taken to estimate the final outcomes.

### **Methods used to generate the results**

Describe the methods used to calculate the results of the economic evaluation (e.g. directly trial-based, cohort expected value analysis, Monte Carlo simulation).

If the economic evaluation is directly based on individual patient data on costs and outcomes from a relevant, direct randomised trial, indicate whether a probabilistic sensitivity analysis has also been conducted. If so, indicate whether the sensitivity analysis has been calculated parametrically (e.g. Fiellers method) or non-parametrically (e.g. bootstrapping), and justify the choice of method.

Where quantified estimates of outcomes are generated over time, explain the underlying assumptions and rationale, for instance, in sufferers of Chronic Obstructive Pulmonary Disease, the incidence of severe exacerbation events requiring therapy becomes more frequent and severe as the disease progresses. In other medical conditions, assuming a linear relationship between outcomes and time might be clinically plausible, identify and consider inferential assumptions built into the structure of the economic evaluation and comment as appropriate. Show that a systematic approach has been taken to select and justify the assumptions made to quantify the outcomes over time; for example, reference the literature search for similar economic evaluations and/or using a pre-modelling study to present the search for studies of the natural history of the condition.

### **State transition models**

For models involving more than one time period (e.g. state transition models), present the transition diagram (or matrix). This complements the decision-tree diagram by identifying the health states possible in the economic evaluation, indicating the presence and direction of transitional paths between health states, and defining the type of each health state as appropriate (e.g. temporary, absorbing).

Describe the model mechanics: define and justify the cycle length and the follow-up time, and comment as necessary. Define and justify the time points at which events are assumed to occur and the duration of time spent in health states. For a Markov model, specify whether a half-cycle correction has been included or justify its exclusion.

Clearly link each patient-relevant outcome and resource item in the model to its relevant health state(s).

Comment as appropriate on the impact of implicit assumptions inherent in the method chosen. For example, for an economic evaluation that includes Markov components, it is relevant to check the following assumptions:

- Is the memorylessness assumption of the model valid in this case (i.e. is it correct to assume no memory for previous states, such that transition probabilities are independent of previous states)?
- Are there constant or non-constant transition probabilities? If the transition probabilities are constant or homogenous across cycles in the model, they are assumed to be independent of time and thus independent of time-related probabilities, such as ageing of the population and variation in competing risks of the population over time. Allowing for ageing and variation in competing risks of the population over time requires transition probabilities that can vary (i.e. are non-homogenous) across time (number of cycles) in the model.

Describe how the model is calculated (e.g. hypothetical cohort or Monte Carlo simulation). If a Monte Carlo simulation is used, then also:

- specify the number of iterations used per simulation and justify this selection in terms of whether it samples the distribution(s) adequately;
- specify the number of simulations per analysis and justify this selection; and
- indicate whether second-order (or parameter) uncertainty has been simulated and hence whether probabilistic sensitivity analysis is enabled.

## Sources of information

Papers identified from the literature review are a useful resource for assumptions relating to the structure and variables in the economic evaluation. Provide copies of all identified papers used in the evaluation in an appropriately labelled attachment separate from the main body of the assessment report.

## D4 Variables in the economic evaluation

### INFORMATION REQUESTS

- Present, as a minimum, the following information for each variable used in the economic evaluation:
  - name (and definition, as necessary);
  - quantity in natural units (as appropriate; for example, this is not applicable for unit costs); and
  - source.
- Identify and list the direct health care resource items for which there would be a change in use associated with substituting the proposed medical service for the main comparator and define each in terms of natural units.

- Estimate the present value of direct health care resource costs and health outcomes.
- Discuss the implications for the economic evaluation of any important deficiencies in the available evidence base.
- Summarise this information in a table for each type of variable and provide further details of calculations, as necessary.

### **Variables used in the evaluation**

Variables used in the economic evaluation might include:

- health care resource items provided (unit costs should be presented and sourced, quantities should be provided as appropriate);
- outcomes (presented in such a way as to allow the three steps to increase transparency to be distinguished);
- probabilities within each branch of a decision analysis (including transition probabilities or rates in a state transition decision analysis); and
- the discount rate applied to costs and outcomes (discount costs and outcomes incurred beyond the first year at a rate of 5% per year).

The names and definitions of variables should be sufficiently precise to permit verification and replication of the economic evaluation. For example, referring to the Australian Refined Diagnosis Related Group (AR-DRG) classification system is more precise than an episode of hospitalisation. For each source, provide full citation details, including page number as appropriate. It might be necessary to cite more than one source for some variables (e.g. the quantity and unit cost of a resource item).

Each economic evaluation should consider explicitly all material differential effects between the proposed medical service and its main comparator (i.e. include all advantages and disadvantages in the analysis). To help demonstrate this, Sub-section D5 requests the presentation of the results of the economic evaluation first in disaggregated form (i.e. as an array of all material costs and consequences; see the definition of a cost-consequences analysis in Sub-section D1).

For the results of trials and pre-modelling studies conducted to provide variables for the economic evaluation, cross-refer to the responses to Sections B6 and C4 as appropriate.

Justify and assess the impact of any change in the source of information for a variable used in the evaluation from that given or recommended elsewhere.. For some variables where there is no recommended source and several different options are available (e.g. rates of progression of a chronic medical condition), it might be important to show that a systematic approach has been taken to select and justify the option used in the economic evaluation (e.g. using a pre-modelling study). The judgment of this importance should be influenced by the sensitivity of the results of the economic evaluation to substituting the different options for the selected option.

Discuss the implications for the economic evaluation of any important deficiencies in the available evidence base. For example, some variables might be estimated imprecisely, or evidence might have been gathered in different populations and circumstances of use or in other health care systems (which is arguably more important for costs). In such cases, explain the limitations of the data and provide details of any attempts to overcome those limitations. Assess the implications using a sensitivity analyses (see also Sub-section D6).

## **Adverse reactions**

Including information on adverse reactions in an economic evaluation can be difficult. Adverse reactions have two main impacts on an economic evaluation: they affect the health outcomes of proposed medical service treatment, and they contribute to the total cost of therapy. Avoidance of an adverse reaction typically associated with the use of the main comparator might be an important and intended outcome of therapy with the proposed medical service. Adverse reactions might affect quality of life, particularly if they have to be tolerated over long periods. Adverse reactions might also lead to discontinuation of the medical service and subsequent substitution of another medical service. A comparative analysis of time-to-treatment cessation of the proposed medical service and the main comparator on the basis of ITT is useful in this situation. Adverse reactions can contribute to costs through unintended hospitalisations, and additional procedures and investigations. Deal appropriately with these impacts to avoid double-counting in the economic evaluation. Generally, the preferred approach is to include them in a full economic evaluation. However, in some circumstances, presenting a cost analysis might suffice (see Part III, Section D(i)).

## **Direct health care resources**

The health care resource items for which there would be a change in use associated with substituting the proposed medical service for the main comparator need to be identified.

The following should be considered where appropriate:

- proposed therapeutic medical services (direct costs of treatment and proposed therapeutic medical services used to treat adverse reactions);
- medicines, including pharmaceutical benefits;
- hospital services;
- diagnostic and investigational services;
- community-based services; and
- any other direct medical costs.

Define the natural units, such as number of general practitioner consultations or admissions per diagnosis-related group, used to measure the change in the amount of each resource item.

## **Present value of direct health care resource costs**

For each type of health care resource, quantify the number of natural units provided for each alternative (e.g. number of general practitioner consultations, allied health practitioners, surgery assistants, anaesthetists, number of episodes of hospital admissions). The relevant economic measure is the amount of resource provided, rather than the amount of resource consumed.

Describe and justify the basis for these estimates, specifying the source of the information. The pattern of provision of resources might be measured prospectively in the course of a clinical study by retrospective review of relevant records, by administration of a questionnaire or survey, or through the use of diaries. Distinguish between data on resource use that are directly derived from the primary evidence, and extrapolations or modelling of resource use beyond that available from the primary evidence. Justify any

choice to use data that are not consistent with data from the primary evidence, particularly where this has an important impact on incremental costs as revealed in the sensitivity analyses.

Section D adopts a broad perspective for the valuation of health care resources, so all contributions to the costs of health care resources - including those paid for by patients, Governments, health insurance agencies and any other part of society - should be considered for inclusion in the economic evaluation. In contrast, Section E primarily considers contributions to resources paid for if publicly funded in Australia only and by Government health budgets only .

It might be reasonable to exclude types of resources that have such a small impact on incremental costs that they would not have a material influence on the conclusion of the economic evaluation.

The unit prices should be as current as possible at the date of the assessment report. If there are particularly pressing reasons to use different unit prices, justify each and supply its source or describe its generation. Ensure that any different unit price is consistent with the broad perspective of including all contributions to the costs of health care resources, in keeping with the rest of this document. To permit MSAC to gauge the effect of using the alternative unit costs, present the results of the economic evaluation using first the unit costs recommended by the manual and then the alternative unit costs.

A format for summarising the minimum dataset of resource items and their associated unit costs relevant to the economic evaluation is suggested in Table D4.1. Some rows have been completed to clarify the suggested format. These are samples for each identified category, which are consistent with the manual, but are not comprehensive of all types of health care resource items, natural units of measurement, or sources of unit costs.



**Table D4.1 List of health care resource items and unit costs included in the economic evaluation**

Resource item	Unit of measurement	Unit cost	Bearer of cost	Source of unit cost
<b>Medical services</b>				
GP attendance	Visit	\$x	\$y– MBS	MBS item 23 <sup>a</sup>
Initial specialist attendance	Visit	\$x	\$y– MBS	MBS item 104 <sup>a</sup>
Subsequent specialist attendance	Visit	\$x	\$y– MBS	MBS item 105 <sup>a</sup>
<b>Hospital services</b>				
Hospitalisation for retinal procedure	Hospital stay	\$x	\$x– Government	Average cost per DRG according to AR-DRG Public Sector Estimated Cost Weights Round 7 — Item C03Z
<b>Diagnostic and investigational services</b>				
Ultrasound of orbital contents	Visit	\$x	\$y– MBS	MBS item 55030 <sup>a</sup>
<b>Medicines</b>				
Tobramycin eye drops	Bottle (5 mL)	\$x	\$y– PBS \$z– patient	PBS item 2328 M — average co-payment estimated assuming a percentage of patients are general and remainder are concessional
Tobramycin eye ointment	Tube (3.5 mL)	\$x	\$y– PBS \$z– patient	PBS item 2329N — average co-payment estimated assuming a percentage of patients are general and remainder are concessional

<sup>a</sup> Costs can be obtained from MBS Online

All steps taken to calculate costs in the economic evaluation should be presented in a way that allows independent verification of the calculations. If a complete presentation is likely to make the main body of the assessment report too bulky, the calculations should be presented in a technical document (see Sub-section 4.3 in Part I). Provide clear cross-references between the calculations and the main body of the assessment report. Include an electronic version of the detailed calculations.

Value future costs at current prices. This is consistent with using constant prices in the economic evaluation. Accordingly, no allowance for future inflation should be included in the calculations.

The present value of future costs should also be estimated. This means that where costs extend over a number of time periods (beyond one year), they should be discounted. Discounting of future costs and benefits is a standard feature of economic evaluation. Costs or benefits are discounted at an annual rate of 5%. If discounting is important in an economic evaluation, this can be examined in sensitivity analyses using different discount rates (see Sub-section D6).

## **Present value of health outcomes**

Nominate and justify the outcome that is considered to best reflect the comparative clinical management algorithm performance of the medical services being compared. This should generally be based on the outcome measure that most closely and validly estimates the final health outcome from a patient perspective. The outcome on which the economic evaluation is based might need to reflect more than one type of intermediate outcome (e.g. where desired and adverse outcomes need to be considered). Justify the choice of any other outcome measure included in the economic evaluation.

For each relevant outcome, quantify the effect of the proposed medical service on the course of the medical condition being managed, either in terms of direct increments, or as streams of effects for the proposed medical service and main comparator in separate arms of the decision analysis, with the increments determined across the arms. Where possible and appropriate, quantify this effect in terms of the patient's health-related quality of life, distributed across different health states over time. Where utility weights were not elicited via a MAUI in the direct randomised trials, this might form a basis for valuing these effects in a manner that reflects the preferences of the general population (see Section C and Appendix 5). Describe and justify the basis for these estimates, specifying the source of the information, including by reference to the data presented in Sections B or C of the assessment report. Distinguish between data on outcomes that are directly derived from the primary evidence, and extrapolations or modelling of outcomes beyond that available from the primary evidence. For example, refer to any analysis presented in Section C of the assessment report to transform an outcome as measured in the direct randomised trials into an outcome presented in the economic evaluation. This includes transforming a modelled final outcome from a measured extent of treatment effect in the trials.

List and document all variables influencing the estimate of outcomes in a table. In the table, highlight the variables that generate the incremental treatment effect on the final outcome estimated in the economic evaluation. These variables include the health states representing the patient-relevant outcomes and the probabilities in each branch of the decision analysis that together simulate a treatment effect by differing between the two arms (representing the proposed medical service and its main comparator) of the economic evaluation. Explain the mechanics of this simulation, because it is usually an important driver of an economic evaluation, and assess the resulting estimate of incremental treatment effect in the context of the analyses presented in Sections B or C of the assessment report.

The present value of future health outcomes measured from the trials or estimated from the model should also be calculated using the approach described above for costs.

If health-related quality of life is not measured directly in the direct randomised trials using a MAUI, which allows direct translation to utility weights via the associated preference-based scoring algorithm, the economic evaluation might include scenario-based utility weights to transform the outcomes measured in those trials into a cost-utility analysis (see Sub-section D1 and Appendix 4).

Transition variables can affect both the streams of costs and outcomes. It is usually easier to discuss them alongside the outcome variables.

## State transition models

Present the transition probabilities of the model, preferably in a matrix. Provide the source of each transition probability and justify the estimate used. Pay particular attention to the transition probabilities that simulate a treatment effect by differing between the proposed medical service and its main comparator. For each transition probability, and for any other time or age-dependent variable, indicate whether it is assumed to be constant or to vary over time, and justify the assumption. If a transition probability is modelled as varying according to time or age, describe how this is achieved in the model.

Where probabilistic cost-effectiveness modelling is presented, list the probability distribution around each variable and justify the selection of each type. For example, gamma or log-normal distributions (i.e. non-negative) could be used for cost parameters, beta distributions for transition probabilities in a control arm, and log-normal distributions for relative risks. For a modelled estimate of incremental effectiveness derived from direct randomised trial evidence, explain how the assumed distribution of the variable reflects the 95% confidence interval around the estimate reported in the trial(s). For each other variable, explain and justify how the selected distribution reflects the extent of statistical imprecision associated with the variable. Also explain and justify each assumed correlation (or lack of correlation) of distributions across the variables.

## Time-to-event data (extrapolated)

Present the calculations of the integrals between the two Kaplan–Meier curves from within the horizon of the median duration of follow-up in the trial(s), with appropriate discounting of any patient-relevant events occurring beyond 12 months of starting the therapy. Similarly, but separately, present the corresponding calculations based on the methods justified in response to Sub-section C2 to extrapolate beyond the horizon of the median duration of follow-up in the direct randomised trial(s).

Where patients transit uni-directionally in a modelled economic evaluation from one mutually exclusive health state to the next, more than one time-to-event analysis can be applied in the same economic evaluation (‘partitioned survival’). A particular application of this in economic evaluations of late-stage cancer treatment has involved the quality-adjusted time without symptoms of the disease or toxicity (Q-TWiST) health state. Time with toxicity is measured using mean time-to-treatment cessation for each arm of the trial; time in the Q-TWiST health state is measured as the difference between mean time-to-disease progression and mean time-to-treatment cessation for each arm of the trial; and time with symptoms of the disease is measured as the difference between mean time-to-death and mean time-to-disease progression. These health states are assigned utilities to then calculate QALYs gained.

## Additional considerations relating to necessary diagnostic criteria

A number of issues arise when an economic evaluation needs to reflect the impact of requesting that diagnostic tests and/or criteria be specifically used to determine eligibility to start or continue MBS-funded therapy (see Sub-section A5 for advice on identifying and specifying tests and criteria).

Ensure that the costs of conducting tests and/or implementing criteria are included in the economic evaluation and are generated for the population tested, not just the population with positive results. The costs should include assessments that demonstrate that certain individuals do not meet the eligibility criteria and for repeat assessments of these individuals.

Also examine the overall impact of false positive and false negative results on the identification of eligible patients, and/or treatment response on the application of the trial results for the economic evaluation, particularly if the latter are used in any proposed continuation criteria in the requested restriction. This examination of predictive value typically requires a separate presentation of additional information on the reliability, sensitivity and specificity of the relevant tests and/or criteria, both across all trials presented and in regular Australian practice. Because predictive value also varies by varying prevalence, evidence of varying prevalence should also be provided. False positives and false negatives both tend to diminish the ability of the tests and/or criteria to make the incremental cost-effectiveness ratio more favourable than an analysis that does not include the tests and/or criteria that the costs of the diagnostic work-up alone make the ratio less favourable.

When considering the impacts of diagnostic tests, distinguish between health outcomes and non-health outcomes. Affected health outcomes include a risk of harm to individuals examined for the diagnostic test, or a risk of harm that arises from changes in treatment that result from the diagnostic test. Include health outcomes only in the base-case analysis. Consider including any non-health-related impacts in a supplementary analysis.

## **D5 Results of the economic evaluation**

### **INFORMATION REQUESTS**

- Present the cost per course of treatment if the proposed medical service is for acute or self-limited therapy, or the cost per year if the proposed medical service is for chronic or continuing therapy.
- Present the remaining results of the economic evaluation first in a disaggregated form, then in increasingly aggregated forms. Use discounting as appropriate.
- Present the appropriately aggregated and discounted results separately for outcomes and costs, and separately for the proposed medical service and its main comparator.
- Present separate estimates of the incremental cost and the incremental effectiveness of substituting the proposed medical service for the main comparator.
- For cost-effectiveness and cost-utility analyses, present the incremental cost-effectiveness ratio as the incremental cost of achieving each extra unit of outcome with the proposed medical service substituted for the main comparator (the base case of the economic evaluation).
- Draw a conclusion from the base-case economic evaluation that reflects the degree of uncertainty around the presented incremental cost-effectiveness ratios.

### **ADDITIONAL INFORMATION REQUESTS IF THE EVALUATION INCLUDES VARIABLES REPORTED IN SECTION C**

- Present the results of the three steps described in Sub-section D1 to derive a stepped base-case economic evaluation.
- Identify components of the evaluation that have more important impacts on the incremental cost-effectiveness ratio.

- Assess the strength of the evidence that supports the components with the more important impacts and as the basis for identifying matters for the sensitivity analyses.

The presentation of disaggregated results depends on the methods used to generate the results of the economic evaluation. For example, where possible, present the quantities of each type of resource provided in its natural units, as well as its cost valued in dollar terms; and/or present the costs and outcomes associated with each branch in the tree of the decision analysis; and/or each health state where the economic evaluation involves a state transition model.

### **Health care resource costs**

Present the estimated health care resource costs in disaggregated form (i.e. separately for each type of resource provided). The nature of this disaggregation is likely to vary across types of economic evaluations.

For a decision analysis that does not calculate costs and outcomes over multiple intermediary time periods (e.g. a decision analysis that is not a state transition model), estimate and present the number of each type of resource item provided in its natural units at each stage in each branch of each arm of the economic evaluation. Then sum the numbers of each type of resource item in each arm before multiplying by the appropriate unit cost for the resource item. In this circumstance, it is helpful to present a table similar to Table D5.1.

For a comparison across state transition models that calculate costs and outcomes over multiple intermediary time periods (e.g. Markov models), two tables (see Tables D5.2 and D5.3) are needed to summarise this type of information.

First, present in a table the number of each type of resource item provided in their natural units for each health state of the models calculated over the duration of one cycle (this should be constant over any cycle in each model each time the health state is entered). Then multiply by the appropriate unit cost for the resource item before summing to estimate the costs for the health state (see Table D5.2).

Second, present a table that partitions the costs according to their health states across all cycles of the models (see Table D5.3).

**Table D5.1 List of health care resource items and summary of cost impacts in the economic evaluation**

Type of resource item	Cost for proposed therapeutic medical service	Cost for main comparator	Incremental cost	% of total incremental cost
<i>Medicines</i>				
PBS drug form and strength	A\$x	A\$y	\$x – \$y	z%
Non-PBS drug form and strength	A\$x	A\$y	\$x – \$y	z%
<i>Medical services</i>				
Type of medical practitioner attendance	A\$x	A\$y	\$x – \$y	z%
<i>Hospital services</i>				
Hospitalisation admission	A\$x	A\$y	\$x – \$y	z%
Outpatient clinic	A\$x	A\$y	\$x – \$y	z%
Emergency department	A\$x	A\$y	\$x – \$y	z%
<i>Diagnostic and investigational services</i>				
Type of service	A\$x	A\$y	\$x – \$y	z%
<i>Allied health care services</i>				
Type of allied health consultation	A\$x	A\$y	\$x – \$y	z%
<b>Total</b>	<b>A\$x</b>	<b>A\$y</b>	<b>\$x – \$y</b>	<b>100%</b>

**Note:** For a decision analysis that does not calculate costs and outcomes over multiple intermediary time periods.

**Table D5.2 List of health care resource items and summary of cost impacts for each health state in a state transition model**

Type of resource item	Number of items in natural unit of measurement	Unit cost	Total cost
<b>Health state 1</b>			
Resource type 1		A\$x	A\$x
Resource type 2		A\$x	A\$x
Etc.		A\$x	A\$x
Total for health state 1			A\$x
<b>Health state 2</b>			
Etc.		A\$x	A\$x

**Table D5.3 List of health states and summary of cost impacts included in the economic evaluation**

Health state in model	Cost for proposed medical service	Cost for main comparator	Incremental cost	% of total incremental cost
Health state 1	\$x <sub>1</sub>	\$y <sub>1</sub>	\$x <sub>1</sub> – \$y <sub>1</sub>	z <sub>1</sub> %
Health state 2	\$x <sub>2</sub>	\$y <sub>2</sub>	\$x <sub>2</sub> – \$y <sub>2</sub>	z <sub>2</sub> %
Etc.	\$x <sub>etc.</sub>	\$y <sub>etc.</sub>	\$x <sub>etc.</sub> – \$y <sub>etc.</sub>	z <sub>etc.</sub> %
<b>Total</b>	<b>\$x</b>	<b>\$y</b>	<b>\$x – \$y</b>	<b>100%</b>

Calculate and present the present value of the direct health care resource costs for each therapy (i.e. separately for the proposed medical service and its main comparator).

Calculate and present the incremental direct health care resource costs by subtracting the present value of direct health care resource costs of the main comparator from those of the proposed medical service. The incremental costs are therefore the costs of any increase in resource provision minus offsets resulting from any improvement in outcome.

## Health outcomes

Present the estimated present value of the health outcomes in disaggregated form (i.e. separately for the proposed medical service and its main comparator).

Calculate and present the incremental health outcomes by subtracting the present value of the health outcomes of the main comparator from those of the proposed medical service.

For a comparison across state transition models that calculate costs and outcomes over multiple intermediary time periods (e.g. Markov models), also present a table that partitions the outcomes in the models according to their health states (see Table D5.4).

**Table D5.4 List of health states and summary of health outcomes included in the economic evaluation**

Health state in model	Outcome for proposed medical service	Outcome for main comparator	Incremental outcome	% of total incremental outcome
Health state 1	$x_1$	$y_1$	$x_1 - y_1$	$z_1\%$
Health state 2	$x_2$	$y_2$	$x_2 - y_2$	$z_2\%$
Etc.	$x_{etc.}$	$y_{etc.}$	$x_{etc.} - y_{etc.}$	$z_{etc.}\%$
<b>Total</b>	<b>x</b>	<b>y</b>	<b>x - y</b>	<b>100%</b>

## Additional disaggregations of state transition models

Where the economic evaluation involves a state transition model, present model traces (e.g. Markov traces) that plot key outputs on a graph with time on the x-axis against the changing outputs on the y-axis in tabulated or graphical form, or, preferably, both forms. For some state transition models, such as those calculated by Monte Carlo simulations, tracker variables could be used to record the information necessary to construct the model traces. Comment on whether each of the model traces makes sense.

For each arm (i.e. for the proposed medical service and its main comparator) and after each cycle, present model traces that:

- identify the proportions of the cohorts in each health state (both for the increment of each cycle over the previous cycle and as cumulative results);
- correspond to observed data (e.g. a model of a medical service used in oncology that generates life-years gained from disease-free survival can be compared with a Kaplan–Meier curve of overall survival, or a model of a medical condition that generates clinical events can be compared with observed data on the natural history of the medical condition, or a genetic test leading to reduced costs in the next generation); and
- sum the outcomes (e.g. QALYs) and the costs (both for the increment of each cycle over the previous cycle and as cumulative results), discounted as appropriate.

For the increment of the proposed medical service over its main comparator after each cycle, present model traces that calculate the incremental costs, incremental outcomes and incremental cost-effectiveness, each discounted as appropriate. For each of these, present model traces both for the increment of each cycle over the previous cycle and as cumulative results.

Where possible, compare those model traces that correspond with observed or empirical data (e.g. overall survival or partitioned survival) as a means of validating the model. Comment on and explain any differences indicated by this comparison to help validate the model (see below).

### **Incremental costs and effectiveness**

Present the base-case incremental cost-effectiveness ratio calculated as the incremental costs divided by the incremental health outcomes.

If the outcome in the denominator of the incremental cost-effectiveness ratio does not include time as part of the units of measurement (e.g. the outcome is expressed on a per-patient or on a per-event basis rather than a per life-year gained basis or a per QALY gained basis), then also specify the duration of the economic evaluation when presenting these results (e.g. 'per extra responder at six months'). This helps in the interpretation of the ratio, because — except when limited to a defined course of therapy — the cost of therapy per patient usually increases over time.

Reflect the degree of uncertainty (see Sub-section D6) around the incremental cost-effectiveness ratios from the presented results when drawing conclusions from the economic evaluation. Avoid terms such as 'dominant' and 'dominated' except in situations where one alternative both costs less and is more effective than the other under a wide range of plausible assumptions.

Where probabilistic cost-effectiveness modelling is undertaken or a probabilistic cost-effectiveness analysis is based directly on a direct randomised trial, present the distribution of overall results both in a scatter plot on the cost-effectiveness plane and in a tabulated format, including the percentages of the distribution of the results in each quadrant of the cost-effectiveness plane. Also present cost-effectiveness acceptability curves. Avoid over interpreting these results. For example, unless the data contributing to this analysis are derived directly from individual patient data collected in the context of a direct randomised trial, important sources of non-statistical uncertainty also need to be examined separately from this analysis.

If the incremental cost-effectiveness ratio is based on a disease-specific outcome (i.e. other than extra life-years gained or extra QALYs gained), consider whether this ratio can be compared to a similar ratio known to the applicant that might be related to one or more previous MSAC decisions. Such previous decisions might provide a narrower benchmark or frame of reference than the more widely conceptualised 'league table' based on the two more widely comparable outcomes above. The precedence value is not necessarily determinative because it is indirect at best and might not capture all elements of an overall comparative cost-effectiveness assessment, let alone the influence of other relevant factors (such as disease severity; see Section F for an opportunity to identify and comment on these). However, a proposed medical service with a less favourable incremental cost-effectiveness ratio in a particular restriction than another comparable medical service and restriction previously rejected is unlikely to be recommended.



On the other hand, a proposed medical service with a more favourable incremental cost-effectiveness ratio in a particular restriction than another comparable medical service and restriction previously recommended is likely to be recommended.

If a claim is made for a change in non-health care resource costs or a change in non-health outcomes such as production changes, present a supplementary analysis with these included (see Appendix 6 for a rationale).

### **Validating the incremental cost-effectiveness ratio**

Consider developing and presenting any approaches to validate the results of a modelled economic evaluation. The comparison of model traces with observed or empirical data (see above) is one such approach where the economic evaluation involves a state transition model. Comment on and explain any differences indicated by this comparison to help validate the model.

Related approaches might compare the output of the model assuming no medical service, with any epidemiological data on the natural history of the medical condition being modelled. Related approaches might also compare the output of the model assuming a particular medical service, with any available long-term longitudinal observational data on that medical service.

Where a model relies on one estimate of treatment effect (e.g. a treatment effect used to transform a surrogate outcome to a final outcome, or a treatment effect on one component of a composite outcome) and there is a comparable estimate of treatment effect on another outcome generated by the model (e.g. the final outcome or another component in the composite outcome), consider using this as a basis to validate the results of the model.

### **Stepped economic evaluation (requested if the evaluation includes variables derived from Section C)**

As explained in Sub-section D1, if pre-modelling studies are presented in Section C, a stepped approach is requested to help MSAC gauge the impact of making these modifications on an unmodified trial-based economic evaluation. See Tables D5.5 and D5.6 for further advice on presenting this analysis.

The preferred order of considering the translation of the trial-based economic evaluation (Step 1) is to consider next the impact of applying the treatment effect (Step 2), where applicable. To facilitate this consideration, the structure of Table D5.5 is aligned to the structure of Table D2.1. More flexibility is warranted in considering the impact of extrapolating and transforming the treatment effect (Step 3). Table D5.6 therefore suggests three alternative next steps to combine the results of Step 2 with either an extrapolation step or a transformation step (Step 3a). Each of these represents the incorporation of a possible pre-modelling study; an assessment report need only report the option for Step 3a that is relevant to its economic evaluation. The final row of Table D5.6 incorporates all pre-modelling studies to complete the impacts of translation (application, extrapolation and transformation) of the trial-based economic evaluation into a modelled economic evaluation. The incremental cost-effectiveness ratio should therefore correspond to the base case of a stepped economic evaluation presented in an assessment report.

If it would further clarify the impacts of translation of the clinical evaluation to the economic evaluation, present more steps and/or more detail of each step (e.g. costs for the proposed medical service and the main comparator, as well as the incremental costs).

The three steps also help identify assumptions and approaches to be examined in more detail in the sensitivity analyses. For example, if the main impact is achieved by extrapolating the final outcome over time, discuss the rationale for the important underlying assumptions for the extrapolation, such as an assumption about the duration of treatment effect (continued divergence of survival curves) or an assumption that a difference generated by one point in time is maintained (at which point the survival curves remain parallel), rather than the more biologically plausible assumption of eventual convergence of survival curves. In this example, it is therefore important that the biological plausibility and validity of the extrapolations are considered (e.g. an assumption of a linear relationship between outcomes and time might not be clinically plausible for many medical conditions).

Consider also the compounding impact on uncertainty of combining these steps to estimate the overall treatment effect on the final outcome in the economic evaluation.

**Table D5.5 Assessment of the implications for the economic evaluation of applying the clinical evaluation (Step 1 then Step 2)**

Population and circumstances of use	As defined in trial(s) using ITT population	As defined by the requested restriction <sup>a</sup>
<b>Costs</b>		
Costs of therapy involving the proposed medical service	(Trial-based)	(Trial-based) <sup>b</sup>
Costs of therapy involving the main comparator	(Trial-based)	(Trial-based) <sup>b</sup>
Incremental costs	(Trial-based)	(Trial-based) <sup>b</sup>
<b>For each trial-based outcome relied on in the economic evaluation before any extrapolation and/or transformation</b>		
Extent of outcomes with the proposed medical service		
Extent of outcomes with the main comparator		
Incremental effectiveness (with 95% CI)	(From Section B)	(From Section C4)
ICER	<b>(Step 1)</b>	<b>(Step 2)</b>
Sensitivity analysis of ICER substituting the upper 95% confidence limit of the difference in outcomes achieved		
Sensitivity analysis of ICER substituting the lower 95% confidence limit of the difference in outcomes achieved		

CI = confidence interval; ICER = incremental cost-effectiveness ratio; ITT = intention to treat

**a** If there is no need to apply the results of the clinical evaluation, the data in this column should be identical to the data in the adjacent column reporting the incremental impacts using the results for the ITT population.

**b** Justify any variation in estimate of incremental costs from the trial-based costing.

**Table D5.6 Assessment of the implications for the economic evaluation of extrapolating and transforming the clinical evaluation (Step 3)**

	<b>Incremental costs</b>	<b>Incremental effectiveness</b>	<b>Incremental cost-effectiveness <sup>a</sup></b>
For each trial-based outcome relied on in the economic evaluation without further modification	(From corresponding row of Step 2 in Table D5.5)	(From corresponding row of Step 2 in Table D5.5)	(From corresponding row of Step 2 in Table D5.5)
For any trial-based outcome relied on in the economic evaluation <i>with any extrapolation from the time horizon of the trial(s) only</i> <sup>b</sup>	(Based on corresponding extrapolation of duration of treatment, if any)	(From Sub-section C4 if extrapolation is required)	<b>(Alternative Step 3a)</b>
For any important outcome generated for or by the economic evaluation from the trial-based outcome(s) ( <i>'transformation of nature of outcome' only</i> ) <sup>c</sup>	(Include here any modelled increases in the provision of some resources and any modelled offsetting decreases of others)	(From Sub-section C4 if possible, or if this approach is used, explain why a presentation here is not possible)	<b>(Alternative Step 3a)</b>
For the final outcome relied on in the economic evaluation generated as a valuation of the trial-based outcome(s) ( <i>'value transformation' only</i> )	(Should not change from Step 2 because nature of outcome does not change)	(From Sub-section C4 if possible, or if this approach is used, explain why a presentation here is not possible)	<b>(Alternative Step 3a)</b>
For the final outcome relied on in the economic evaluation <i>combining any extrapolation from the time horizon of the trial(s) with any transformation of the trial-based outcome(s)</i>			<b>(Completed Step 3 and expected base case) <sup>d</sup></b>

**a** With sensitivity analyses substituting the upper and lower 95% confidence limits of the difference in outcomes achieved.

**b** Justify and explain the methods of the approach taken to align the changes in the incremental costs (or incremental effectiveness) to correspond to the changes in incremental effectiveness (or incremental costs) reported by any pre-modelling study summarised in Sub-section C4 to extrapolate the evidence from the trial(s) to the time horizon of the economic evaluation.

**c** Where the approach to transforming the nature of the outcome also involves extending the time horizon of the analysis, justify and explain the methods of the approach taken to align the changes in the incremental costs to correspond to the changes in incremental effectiveness reported by any pre-modelling study summarised in Sub-section C4.

**d** Justify if claiming a different base-case analysis from that defined above.

## **D6 Sensitivity analyses**

### **INFORMATION REQUESTS**

- Present univariate (one-way) sensitivity analyses on all variables using plausible extremes of values, and justify the selection of those extreme values.
- Tabulate all univariate sensitivity analyses alongside the base case.
- Present multivariate sensitivity analyses combining variables shown to be sensitive in the univariate analyses.
- Examine and present the sensitivity of the results of the economic analysis to any changes in assumptions concerning the structure of the modelled economic evaluation that are important but uncertain.

The purpose of a sensitivity analysis is to examine the effect of uncertainty around estimates and assumptions included in the economic evaluation on the results of the base-case economic evaluation. Statistical (probabilistic) uncertainty involves random error and can be reduced by increasing sample size. The many other sources of uncertainty involve systematic error, are harder to identify and cannot be reduced by increasing sample size. For example, they arise in the selection and measurement of information, the specification of the structure of a model, and the plausibility of the implicit and explicit assumptions relied on for the model, particularly in aggregating across the various sources of information.

### **Univariate sensitivity analyses**

The univariate (one-way) sensitivity analyses on all variables should use plausible extremes of values. Justify the selection of the plausible extreme values of each variable. For example, the upper and lower 95% confidence limits of the relevant incremental treatment effect variables reported in direct randomised trials, the considerations summarised in Table C4.1 or the range of estimates from the available studies of the natural history of a medical condition.

Tabulate all univariate sensitivity analyses alongside the base case. A tornado diagram with incremental cost-effectiveness on the x-axis can be used, where possible, as an efficient and informative way of summarising the results of the univariate sensitivity analyses.

Use the univariate sensitivity analyses to highlight the variables that are important drivers of the economic evaluation. Consider providing a matrix with the effects of variables on various outcomes that differ across the two arms (e.g. in terms of health outcomes, mortality and utility).

The three steps to improve the transparency of the economic evaluation are intended to help identify the basis of plausible extreme values of variables for further examination. For example, when curves have been fitted to time-to-event data to extrapolate the results beyond the duration of observed follow-up, the sensitivity analysis should examine both the uncertainty in fitting the curves for the extrapolation, and the upper and lower 95% confidence limits of the time-to-event results measured within the direct randomised trials.

### **Multivariate sensitivity analyses**

The multivariate sensitivity analyses should combine variables shown to be sensitive in the univariate analyses. Explain the selection of these variables and their combination; for example, varying more than one of the steps to improve transparency at the same time. Present the analyses in tabular and graphical format.

Where a probabilistic sensitivity analysis is provided, also examine the sensitivity of base case estimates of incremental cost, incremental effect and incremental cost-effectiveness to changes in one variable at a time as univariate sensitivity analyses conducted on each variable using plausible distributions.

## Sensitivity of the results to changes in the modelled economic evaluation

Examine assumptions concerning the structure of the modelled economic evaluation that are uncertain to assess their importance by the extent to which they affect the results of the evaluation. The three steps to improve the transparency of the economic evaluation might help identify structural issues for further examination.

Similarly, if there is a risk of substantial usage beyond the intended population and circumstances of use defined in the requested restriction, examine the sensitivity of the results to the assumption of usage within these intentions. As discussed in Sub-section D2, this wider population and circumstances would be expected to have demographic and patient characteristics and circumstances that differ from the target population and circumstances. If the intention of the restriction is to limit usage to the population for which the proposed medical service is most cost-effective, these sensitivity analyses should examine the extent to which the incremental cost-effectiveness ratio would become less favourable with increasing usage beyond the restriction. Table D6.1 gives advice on presenting this analysis in a format that is comparable to Tables D2.1 and D5.5.

**Table D6.1**      **Analyses of the implications for the economic evaluation of usage beyond the requested populations and circumstances of use**

<b>Population and circumstances of use</b>	<b>As defined by the requested restriction</b>	<b>If use beyond the requested restriction might arise</b>
Incremental costs		
Incremental effectiveness		
Incremental cost-effectiveness ratio		

If a cost-utility analysis is presented, also present the results of the economic evaluation with the utility in all health states set to one to generate the incremental cost per extra life-year gained. This helps identify the contribution of any life extension component to the incremental effectiveness claim.

If discounting has been necessary, the robustness of the results to different discount rates (including a zero discount rate on nonmonetary outcomes alone and on both costs and outcomes) should be tested.

## Section E

### Estimated utilisation and financial implications

---

#### Introduction

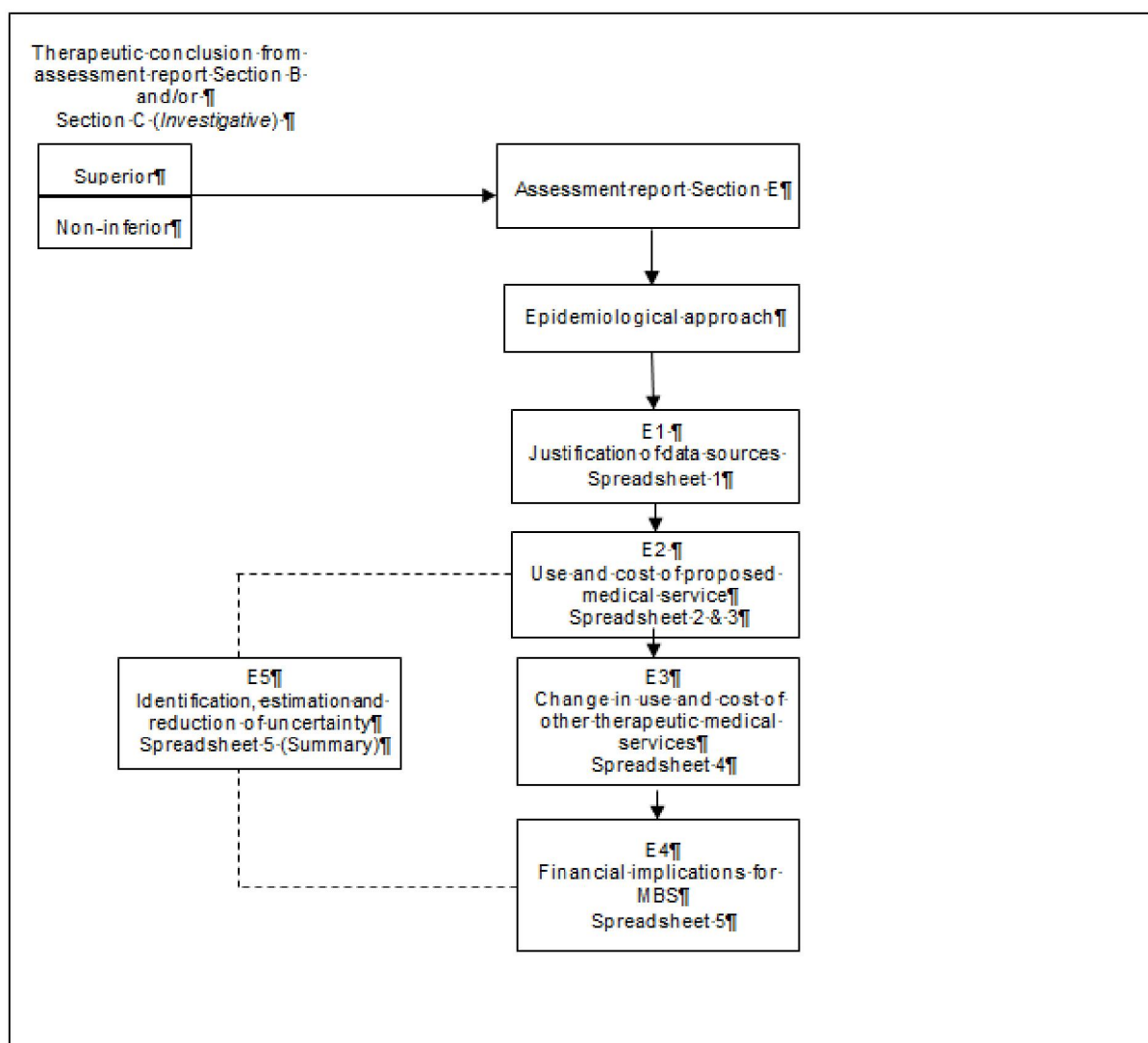
The purpose of this Section is to generate the most likely utilisation and financial estimates by requesting a set of budget impact analyses. These analyses are relevant to both the MSAC and the Australian Government. In the event of a positive recommendation by MSAC, the Australian Government needs utilisation and financial estimates to help provide the necessary funds.

Figure E1 shows the epidemiological approach for developing utilisation and financial estimates for a medical service. A market share approach, as used in some cases for applications to PBAC is not applicable for medical services. As the flowchart shows, these are not mutually exclusive. It also helps explain the logic behind the steps that build on the epidemiological basis and that support the preferred format of calculating and presenting these estimates using the utilisation and cost model spreadsheets supplied alongside these Guidelines, based on a standardised Excel 2010- or STATA or Triage or Reference Manager / RevMan workbook. Together with this Section, this preferred workbook format is primarily designed to present the necessary calculations using the epidemiological basis consistently across assessment reports.

An epidemiological base is usually preferred for generating utilisation and financial estimates if in the prepared assessment report it concludes that, overall, the proposed medical service has an advantage over its main comparator(s). This decision parallels the cost-effectiveness approach that would be taken in Section D of the assessment report. The epidemiological approach first estimates the number of people with the medical condition and then uses several steps to estimate the use of the proposed medical service (see Sub-section E2) and of other medical services in the context of the main indication (see Sub-section E3).

Section E of these Guidelines focuses on the presentation of estimates adopting an epidemiological basis. This approach is informative for some assessment reports prepared — for example, where there is uncertainty in the investigative conclusion or where there is large uncertainty in the expected utilisation. (see Sub-section E5).

**Figure E1** Key information requests for assessment report Section E of a standard assessment for MSAC



Sub-sections E2–E4 request financial analyses relevant to the funding program (e.g. MBS-listing budgets) by only considering health care resources subsidised through those programs. In contrast to the economic evaluation presented in Section D of the prepared assessment report, these financial analyses exclude health outcomes, scale up estimates to assess the impact for the program overall, do not use discounting, and exclude any resource item.

The following Sections lay out a preferred stepwise process to generate utilisation and financial estimates. Whenever it is thought appropriate to include an approach that is not requested below, justify the approach in the main body of the assessment report. Whenever it is thought appropriate not to take an approach that is requested below, a particularly strong justification should be provided and, where possible, the alternative approach should be presented separately and in addition to the requested approach.

Where an assessment report seeks listing for more than one indication, present a separate standardised Excel 2003 workbook for each indication (refer to Section 5 of the Guidelines). As a final step in each of Sub-sections E4 and E5, these results can be aggregated across the indications.

## E1 Justification of the selection of sources of data

### INFORMATION REQUESTS

- Where data are available (published or unpublished) from one or more types of data sources:
  - summarise the methods used to obtain the data;
  - present the relevant main results;
  - interpret the findings; and
  - discuss the limitations (including the representativeness of the results) and biases of the method adopted.
- Where data are obtained via one or more studies commissioned for the assessment report:
  - describe the gap in the information to be addressed by the commissioned analysis;
  - summarise the methods used to obtain and analyse the data;
  - present the relevant main results;
  - interpret the findings; and
  - discuss the limitations (including the representativeness of the results) and biases of the method adopted.
- Use Spreadsheet 1 of the standardised Excel workbook to summarise all the background information, primary (non-calculated) variables and assumptions essential to the calculation of results presented in this Section.
- Provide a copy of the data from each published and commissioned study with the attachments to the assessment report. Include the correspondence that requested the data for a commissioned study.

### Published data sources

Data sources suitable to the approach taken should be stated and discussed in the assessment report. Data availability for prevalence and incidence is variable, but the best available data should be justified and used where possible. Data sources fall under the broad headings listed in Table E1.1, however, there might be other suitable data sources.<sup>2</sup> In each case, the methods used should be summarised and the results presented and interpreted, including a discussion of the limitations and biases of the method used.

Sources include data from Australia or overseas, such as MBS or Casemix data, for equivalent medical services that are already listed, and overseas data on the use (in markets similar to Australia) of a proposed medical service that has no comparator that is publically funded in Australia. Where there are multiple sources of data, assess the validity and applicability of both the source and the data in relation to their use in the assessment report's calculations. The demonstration of concordance across multiple data sources of similar validity and applicability is encouraged to reduce uncertainty. Present sensitivity analyses reflecting the variation in the estimates from the available data.

---

<sup>2</sup> See *Sources of Epidemiological Data for Use in Generating Utilisation Estimates* for suggested sources of data that might be suitable for the medical condition, relevant to the assessment report.  
<http://www.pbs.gov.au/info/industry/useful-resources/sources>



**Table E1.1 Categories of data sources**

<b>Disease epidemiological data (provide estimates of prevalence or incidence in the population)</b>
<ul style="list-style-type: none"> <li>• Australian case or mortality registers estimate the incidence or prevalence of a disease</li> <li>• Large, well-designed Australian studies estimate the incidence or prevalence of a disease</li> <li>• Australian national health surveys estimate the prevalence of a disease</li> </ul>
<b>Treatment epidemiological data (provide estimates of treated prevalence)</b>
<ul style="list-style-type: none"> <li>• Surveys of the treated prevalence of the disease in Australia</li> <li>• Studies using utilisation databases, including MBS data</li> </ul>

Studies commissioned for the assessment report may include data requests to disease registries, established epidemiological studies or ongoing utilisation studies seeking specific analyses. In each case, the information gap to be filled should be clearly described, and the results presented and interpreted, including a discussion of the limitations and biases of the method used.

In the absence of Australian observed data, a range of observed data from overseas sources could be used. When presenting these data, also discuss the applicability of the estimates from an overseas source to the Australian population. In the case of prevalence data, this discussion should further assess the impact of any variations in the subsidy arrangements between overseas health care systems and those in Australia.

Where multiple sources of data are available to address a single assumption or estimate, compare the results, assess their concordance or lack of concordance, and justify the selection of the base-case estimate and the estimates used in the sensitivity analyses. Present a summary table where multiple sources or multiple variables are being compared.

In the absence of observed data, expert opinion might be required (see Appendix 2). A commissioned evaluation of recent usage practice has many similarities with a survey of expert opinion; a distinguishing characteristic might be that a usage evaluation measures what was done, whereas experts are asked to report what they would do now or in the future.

Each time an assumption is required in the absence of data, state the assumption concisely and explain its basis. Describe the nature and likely magnitude of uncertainty for each assumption (see Sub-section E5). Present an examination of the impact of each assumption by altering it in sensitivity analyses.

### **Spreadsheet 1 ('Background and assumptions')**

When using Spreadsheet 1 of the standardised Excel workbook to summarise the data sources, background information, primary (non-calculated) variables and assumptions, it might be helpful, if the analyses are complex, to add one or more other supporting spreadsheets in the workbook to provide more detail, such as identifying the sources of variables relied on and supporting the assumptions made. The remaining spreadsheets, which calculate the estimates (see below), should be fully integrated so that changes to any variable for the purposes of sensitivity analyses flow on appropriately through succeeding calculations to all results.

## **Copies of data**

To allow independent assessment of the data, include copies of the data used (published, unpublished and commissioned) in an attachment to the assessment report. Ensure that the responses in Section E of the assessment report and Spreadsheet 1 provide adequate cross-references of the extraction of all data used to generate the estimates in these analyses from each attached data source (to the level of the page, table or figure number of each source document).

## **E2 Estimation of use and costs of the proposed medical service**

### **INFORMATION REQUESTS**

- Estimate the number of patients with the medical condition targeted by the proposed medical service, the number who would be eligible for the requested restriction and the number of patients likely to use the proposed medical service.
- Use Spreadsheet 2 of the standardised Excel workbook to calculate the results presented in this part of the Section.
- Estimate the number of times the proposed medical service is delivered in each year over five years (disaggregated into proportions for MBS-listing, and by beneficiary type).
- Estimate the costs for each form of the proposed medical service in each year over five years, multiplying by the relevant unit costs.
- Aggregate these cost calculations for the proposed medical service overall in each year over five years.
- Use Spreadsheet 3 of the standardised Excel workbook to calculate the results presented in this part of the Section.

### **Numbers of patients**

#### ***Use of incidence or prevalence data***

The choice between using incidence and prevalence data is important in estimating the likely number of patients eligible for the proposed medical service in any one year. This choice depends on the nature of the medical condition and its treatment.

In general, an incidence-based approach is preferred for a treatment of short duration, with 12 months being a suggested upper limit, because estimates should be presented in periods of one year (see below). Examples include an acute self-limiting medical condition, each episode of which is treated with a single course of treatment, and a medical condition that is managed by a single course of treatment given once in a lifetime. Incidence should be estimated on a 12-month basis.

In general, a prevalence-based approach is preferred for a treatment that is to be used for long periods, with 12 months being a suggested lower limit; for example, chronic medical conditions for which treatment is delivered regularly (i.e. without breaks in the standard treatment regimen).

For some treatments, a combination of incidence and prevalence bases might be informative. Examples include intermittent treatment of a series of acute episodes of a chronic medical condition, treatment for which is restricted to each episode and in which the proposed medical service is expected to prolong the duration of disease, including by an extension of expected overall survival.

The first example (regular treatment for chronic medical conditions) is complex, because although the number of patients who have the condition might be determined using an epidemiological approach, the number of presentations for treatment can be more difficult to determine. In the second example (intermittent treatment), allowance for an increase in prevalence might be necessary. If disease duration or life expectancy is expected to increase from fewer than five years in the current situation before the listing of the proposed medical service, it would generally be appropriate to increase the initial prevalence pool estimate on an annual basis by the difference in the 12-month incidence of new patients and the 12-month incidence of cured patients or of deaths. This should be continued either until a new steady state is achieved, with constant rather than increasing prevalence, or until the five-year horizon of the analyses is reached.

Expert epidemiological advice should be sought when estimating prevalence from incidence data or estimating incidence from prevalence data, particularly where there is doubt that the duration of disease has not remained constant over time or where it is not expected to remain constant after the listing of the proposed medical service.

### ***Estimate the number of patients with the medical condition***

Estimate the likely number of patients in the current year and in the first year of listing using one of the bases above (incidence or prevalence). These estimates should also incorporate the most probable estimates of patients who are misdiagnosed (i.e. where there might be pressure to diagnose the patient as having the medical condition to be eligible for the proposed medical service and where the differential diagnosis is unclear). Then project the numbers of patients on an annual basis for a total of five years, accounting for population growth and expected changes in prevalence and/or incidence of the condition. If appropriate, more frequent periods (e.g. monthly or three-monthly) could be calculated in the supporting spreadsheets. If so, summarise the presentation of these aggregated data as annual aliquots for a total of five years from listing (Year 1, Year 2, Year 3, Year 4 and Year 5).

### ***Estimate the number of patients eligible for the proposed medical service***

Using these annual numbers of patients with the medical condition for Years 1–5, estimate the proportions that would be expected to be eligible to receive the proposed medical service. These estimates should also include the most probable estimate of patients who are misclassified.

### ***Estimate the number of patients likely to use the proposed medical service***

Using these annual numbers of eligible patients, estimate the proportions likely to use the proposed medical service in each of the five years. The resulting estimates should reflect the likely share of the proposed medical service compared with the other treatment options currently used for eligible patients.

## **Spreadsheet 2 – Epidemiology of the disease and patient numbers'**

Calculate the above three sets of estimates of patient numbers in Spreadsheet 2 ('Epidemiology of the disease and patient numbers') of the standardised Excel workbook.

### ***Number of times the proposed medical service is delivered***

Three elements are involved in translating the numbers of patients likely to be treated to the number of times the proposed medical service is delivered. There is no basis to suggest a preferred order in which they should contribute to the calculations.

The first element is the rate of uptake of the proposed medical service across the five years from listing. If appropriate, shorter periods (e.g. monthly or three-monthly) could be calculated in the supporting spreadsheets. If so, summarise the presentation of these aggregated data as annual aliquots for a total of five years from listing.

The second element is the frequency and duration of treatment involving the proposed medical service. Duration of treatment might be affected by adherence to treatment and rates of discontinuation (e.g. due to poor tolerance or disease progression). Consistent with the information requests in Section D, the estimates should be in terms of the quantities of treatment actually delivered rather than planned. In determining the impact of this element, the variation in duration of treatment between the context of the available randomised trials and probable use of the proposed medical service once listed for MBS funding should be considered. Aspects of this include patient preferences, physician's preferences, switching of proposed medical service, comorbidity in the patients and co-administration of other treatments. Determining estimates of treatment use for the MBS context is therefore based on a number of assumptions and uncertainties that are difficult to quantify; therefore, they should be justified and subjected to sensitivity analyses.

The third element is the mix of forms of the proposed medical service. Where more than one form is specified in response to Section A, there will be more than one product or item listed for MBS funding in Australia to distinguish between these forms, strengths and quantities. The estimates should be disaggregated to the level of the proportions of use of each of these types of the proposed medical service.

Estimate the number of times the proposed medical service is delivered each year over five years by applying these three elements to the patient number estimates from Spreadsheet 2.

### ***Aggregated cost calculations***

Estimate the costs to the MBS of the proposed medical service in each year over five years by applying these breakdowns and unit costs and then aggregating each set of cost estimates.

## **Spreadsheet 3 – Cost of the proposed medical service to the MBS**

Calculate the above sets of estimates of administrations of the proposed medical service and costs in Spreadsheet 3 ('Cost of the proposed medical service to the MBS') of the standardised Excel workbook.

## **E3 Estimation of changes in use and cost of other medical services**

### **INFORMATION REQUESTS**

- Identify the other MBS-listed medical services that are likely to be affected by listing the proposed medical service.
- For each proposed medical service, estimate the extent of change in the number of times the proposed medical service is delivered each year over five years (disaggregated into proportions for the MBS and by beneficiary type).
- Aggregate both these cost calculations for the other affected medical services in each year over five years.
- Use Spreadsheet 4 of the standardised Excel workbook to calculate the results presented in this Section.

#### **Medical services likely to be affected by the listing of the proposed medical service**

MBS-listed medical services likely to be affected by the listing of the proposed medical service include:

- MBS-listed medical services substituted by the proposed medical service;
- other MBS-listed medical services with decreased usage; and/or
- other MBS-listed medical services with increased usage.

As an initial step, identify and list all MBS-listed medical services that fall into each of these three categories. The list should include those MBS-listed medical services identified in Section A.

Of the three categories, substituted medical services usually have the largest impact on the financial implications of listing the proposed medical service. There would be no substituted medical services if the proposed medical service has no comparators or if it is designed to replace a medical procedure. Where all substituted MBS-listed medical services come from a single group of medical services listed on a cost-minimisation basis, the cost differential of each against the proposed medical service should be similar. However, where the cost differential is expected to vary to an important extent across the substituted medical service, also estimate the breakdown of the proportions of the overall substitution to capture the cost implications of the variation.

Proposed medical services that are listed for MBS funding in Australia, with expected decreased usage after being listed, include those that are co-delivered with substituted medical services, those used to treat adverse outcomes to substituted medical services, and those used to treat the clinical end points that might be reduced after treatment involving the proposed medical service.

Medical services that are expected to have an increased usage after being listed for MBS funding in Australia include those that are co-delivered with the proposed medical service, and those used to treat adverse effects caused by, or outcomes of, the proposed medical service.

The impact of adverse outcomes might have less weight if the information provided in Sub-section B7 shows that they are of insufficient clinical importance to require management with MBS-listed medical services, or if they are similar for the proposed medical service and its major comparators. If there is insufficient information available to include the impact of adverse reactions on MBS expenditure, this should be noted.

### **Number of times the proposed medical service is delivered**

Justify the approach adopted for estimating the extent of change for the forms of each affected medical service, where the approach and calculations involve uncertainty. Use the information provided in Section A and Sub-section E2. Identify and justify any inconsistency between Sections D and E of the assessment report in the identification of MBS-listed medical services that would change as a result of listing the proposed medical service, and the extent of change per patient in the first five years of listing.

### **Disaggregation of estimates**

Disaggregation into proportions for the MBS and by beneficiary type should usually be based on the most recent 12 months of usage data from Medicare Australia. An exception could be where the expected substitution is for a distinctive subgroup of current use of the substituted medical service(s), in which case the disaggregation should be based on the subgroup.

### **Costs over five years**

Estimate the costs in each year over five years of each of the forms of each of these medical service substituted, decreased and increased on the basis of each of the estimated utilisation changes. For these calculations, use constant prices, make no allowance for inflation and use a zero discount rate.

### **Aggregated cost calculations**

Estimate the cost offsets to the MBS of the other affected medical services in each year over five years by applying these breakdowns and unit costs, and then aggregating each set of estimates by subtracting the costs of substituted medical services and the costs of medical services with decreased usage from the costs of medical services with increased usage.

### **Spreadsheet 4 – Cost implications to the MBS from substitutions and other increases and decreases**

Calculate the above sets of estimates of number of deliveries and costs in Spreadsheet 4 ('Cost implications to the MBS from substitutions and other increases and decreases') of the standardised Excel workbook.

## **E4 Estimated financial implications for the MBS**

### **INFORMATION REQUESTS**

- Estimate the net financial implications for the MBS in each year over five years by subtracting the net cost offsets for both the aggregated estimates calculated in Sub-section E3 from the corresponding estimates calculated in Sub-section E2.
- Use Spreadsheet 5 of the standardised Excel 2003 workbook to calculate the results presented in this Section.

## Spreadsheet 5 – Net cost of the proposed medical service to the MBS

Calculate the two sets of net financial implications in Spreadsheet 5 ('Net cost of the proposed medical service to the MBS') of the standardised Excel workbook.

### E5 Identification, estimation and reduction of uncertainty

#### INFORMATION REQUESTS

- In each step of the calculations, assess the sources of uncertainty and distinguish the type and degree of uncertainty in utilisation and financial estimates.
- Where possible, explain the nature of each uncertainty and its impact on the overall estimates.
- Estimate the level of the uncertainty and propose ways to reduce it.
- Provide a separate workbook to generate the results of any calculations (e.g. sensitivity analyses and scenario analyses) to examine the impact of uncertainty. Summarise these in Spreadsheet 5 of the standardised Excel 2003 workbook.

#### Nature of uncertainty

When presenting the most likely utilisation and financial estimates, consider the degree of uncertainty of those estimates. Two types of uncertainty should be distinguished:

- 1) *usage that differs from expectations* — generally arises from uncertainty within and across particular variables in the analysis. Sensitivity analyses should be presented to examine the impact of this source of uncertainty; and
- 2) *usage that extends beyond the restriction* (sometimes called 'leakage') — generally arises from uncertainty as to whether the requested restriction would achieve its intended objective in limiting use. Usage beyond the requested restriction raises doubts about the overall cost-effectiveness of the proposed medical service where the intention of the restriction is to exclude its subsidised use in patients for whom that use would not be acceptably cost-effective. Scenario analyses might be relevant to examine the impact of this uncertainty.

#### Sources of uncertainty

The following lists summarise the factors that could be considered when assessing uncertainties in predicted utilisation patterns and financial implications resulting from the listing of a proposed medical service as requested. The lists are not intended to be prescriptive, but generally reflect factors that have been considered previously by MSAC and may arise from epidemiological data, treatment prevalence data, expert opinion and assumptions used in generating the quantified predictions. Any of these factors might provide information that will increase understanding of the uncertainties present in utilisation estimates. It might be useful to consider these factors explicitly, but not all the factors will apply to all assessment reports. Thus, it might not be necessary to address any or all of these questions for each assessment report, as the uncertainties outlined might be very small or of little importance to the overall cost to the MBS. Therefore, consideration should be given to how relevant each of the factors might be for a particular assessment report.

### ***Factors that could affect the extent of usage within the requested restriction***

Consideration of the following factors might provide relevant information on uncertainties within the requested restriction. Some factors might not be relevant in all assessment reports or might have a negligible impact on the overall estimates:

- Promotion might result in greater identification of the proposed medical service, resulting in more health care practitioners considering patients for treatment.
- Indirect media exposure to consumers might result in some consumers being more aware of the proposed medical service and seeking treatment with it. These patients might not be identified if a treated prevalence approach has been used.
- Outcomes of related research might have an impact on uptake of the proposed medical service. This could be positive or negative, and could emerge at the time the assessment report is lodged or be expected to occur within five years of listing.
- More health care practitioners and patients might seek treatment if the proposed medical service treats a medical condition for which the alternatives are considered to be substantially inferior to the proposed medical service (e.g. in terms of effectiveness, tolerability, or patient acceptability and convenience).
- Limited access to designated types of health care practitioners or to designated diagnostic procedures in a requested restriction might limit uptake and utilisation.
- The duration of treatment might be longer than expected, compared to the time frame of the randomised trials, particularly when trials are truncated.
- Patients might be treated more often than expected, particularly in the case of medical conditions with episodic manifestations.
- There might be a likelihood of treatment increasing over time.

### ***Factors that could affect the likelihood of usage beyond the requested restriction***

Some of the factors listed above might also affect the likelihood of usage beyond the requested restriction. More detailed guidance is given in Section A about ways of designing a restriction to minimise usage beyond its intention, however, the following factors might be considered:

- The requested restriction is for a subset of the types of patients who are eligible according to the TGA-approved indication(s).
- The requested restriction is for a subset of the types of patients who were eligible for the randomised trial(s) published for the proposed medical service, or there are randomised trials demonstrating evidence in other medical conditions.
- The requested restriction is for a subset of the types of patients who have been subsidised by the applicant before lodgement of the assessment report (e.g. on compassionate grounds or as part of clinical studies).
- The requested restriction is for a subset of the types of patients for whom the applicant plans to promote use of the proposed medical service before or after the listing for MBS funding is implemented.



- The requested restriction is for a subset of the types of patients who have the underlying medical condition, in this case identify whether:
  - there are any likely difficulties for health care practitioners in determining eligibility for the proposed medical service (e.g. a difficult differential diagnosis, ambiguity in the wording of the restriction, or poor precision or accuracy in a diagnostic test) that might result in misclassifications of eligible patients from the population with the underlying condition; and /or
  - patient advocacy groups are likely to have an influence on determination of eligibility by health care practitioners .

## **Estimating and reducing uncertainty**

The following three aspects should be addressed in any consideration of uncertainty:

- the direction of impact on the estimate (underestimate or overestimate);
- the impact on the magnitude of the estimate (small or large); and
- the likelihood that another estimate should replace the base-case estimate (probable or improbable).

Although quantitative estimates of uncertainty are preferred, semi-quantitative assessments may need to be given in many instances. Where the effects of some uncertainties are difficult to quantify, this should be noted. As a general principle, the more sensitive the overall financial implications are to a particular source of uncertainty, the more important it is to minimise that uncertainty.

One way to reduce uncertainty is to use data from multiple sources, where available. Where estimates derived from different sources are concordant, there might be more confidence and therefore less uncertainty in the resulting estimates. Where this is not the case, the disparity between the estimates might contribute to the estimate of uncertainty. This can be referred to as ‘triangulation’ (the use of multiple sources of data or multiple approaches to determine the consistency or otherwise of the conclusions from those sources or approaches).

## **Summary of calculations**

Summarise the results of any calculations (e.g. sensitivity analyses and scenario analyses) to examine quantitatively the impact of uncertainty in Spreadsheet 5 (‘Net cost of medical service to the MBS’) of the standardised Excel workbook. Do not include the supporting calculations in that workbook. If additional calculations need to be explained, a separate workbook should be provided for any analysis other than the base-case analysis (most likely). Spreadsheet 1 (‘Background and assumptions’) of the separate workbook should highlight the differences from the base-case workbook.

## Section F

### Options to present additional relevant information

---

#### Introduction

Over time, a number of issues have arisen that are important for some assessment reports, but are not necessary for all assessment reports. These include equity principles, ‘rule of rescue’ and other relevant factors that can affect MSAC’s assessment of proposed therapeutic medical services.

This Section is intended to assist the consideration of such issues in relation to an assessment report. It does not cover all possible issues. Ultimately, an applicant may include in an assessment report any information that is relevant to MSAC’s decision.

#### F1 Other relevant factors

##### INFORMATION REQUESTS

- If the assessment report raises any issue relating to equity principles, discuss it in descriptive terms.
- If the assessment report raises any equity assumption that particularly affects consideration of the cost-effectiveness of the proposed medical service, describe the implications, where appropriate, with reference to a sensitivity analysis.
- If the assessment report makes any claim that the ‘rule of rescue’ is applicable, set out the basis for that claim.
- If the assessment report identifies any other relevant factor not requested elsewhere, discuss it in response to this section.

##### Equity principles

From a general policy viewpoint, the MBS promotes fairness in its subsidy arrangements by promoting affordable access to safe, effective and cost-effective medical services. Thus, any listing that is likely to promote particularly, or hinder, these or any other general equity principles should be discussed. For example, if the requested listing of the proposed medical service would raise particular patient affordability considerations, their implications should be discussed.

##### Equity assumptions

From a technical viewpoint, many elements of an economic evaluation contain embedded equity assumptions (e.g. see utility valuation in Appendix 4). In the rare cases in which such underlying assumptions might be important enough to influence a particular MSAC decision, a description of how the issue affects consideration of the cost-effectiveness of the medical service, and preferably an examination of its impact in a sensitivity analysis, should be sufficient.

## Guidance on the ‘rule of rescue’

Four factors, which apply in exceptional circumstances, are particularly influential in favour of listing. When all four factors apply concurrently, this is called the ‘rule of rescue’. The four factors are as follows:

- No alternative exists in Australia to treat patients with the specific circumstances of the medical condition meeting the criteria of the restriction. This means that there are no suitable medical services for these patients.
- The medical condition suffered by the target patient population is severe, progressive and expected to lead to premature death. The more severe the condition, the younger the age at which a person with the condition might die or the closer a person with the condition is to death, the more influential the rule of rescue might be in the consideration by MSAC.
- The medical condition defined by the requested restriction applies to only a very small number of patients. Again, the fewer the patients, the more influential the rule of rescue might be in the consideration by MSAC. However, MSAC is also mindful that the MBS is a community-based scheme and cannot cater for individual circumstances.
- The proposed medical service provides a worthwhile clinical improvement sufficient to qualify as a rescue from the medical condition. The greater the rescue, the more influential the rule of rescue might be in the consideration by MSAC.

As with other relevant factors, the rule of rescue supplements, rather than substitutes for, the evidence-based consideration of comparative cost-effectiveness. A decision on whether the rule of rescue is relevant is only necessary if MSAC would be inclined to reject an assessment report because of its consideration of comparative cost-effectiveness (and any other relevant factors). If MSAC concludes that the rule of rescue is relevant (such as for last-line therapy for terminally ill patients), it will consider whether there is a strong enough case for listing for MSAC to reverse a decision not to recommend listing in the absence of the rule of rescue.

This guidance on the rule of rescue is kept deliberately narrow. Although there are relevant arguments for broadening the guidance, MSAC is concerned that doing this would reduce the relative influence of the rule of rescue when it is applied to a broader set of eligible assessment reports. In other words, the greater the proportion of assessment reports that the rule of rescue is applied to, the smaller its average impact in favour of listing across the identified assessment reports.

One issue that has arisen concerning the rule of rescue is that a second medical service to treat the medical condition considered to meet the requirements of the rule is not suitable for this consideration. This is because, by definition, the second therapeutic medical service does not meet the essential first factor of the four factors (i.e. that there is no currently alternative intervention). This causes a difficulty if listing of the second therapeutic medical service is sought on a cost-minimisation basis.

Another difficulty is that indiscriminate application of arguments such as the rule of rescue can lead to overall inefficiencies, unless MSAC compensates when considering medical services that clearly fall outside the rule.

### **Discuss any other relevant factor**

If any other relevant factor is thought to be worth emphasising and is not already requested elsewhere for inclusion in the assessment report, discuss it in the response to this Section.

## **Part III**

### **Alternative clinical evidence for proposed medical services to be considered by MSAC**

## **Section B(i)**

### **Clinical evaluation for the main indication: indirect comparison of randomised trials**

---

#### **Introduction**

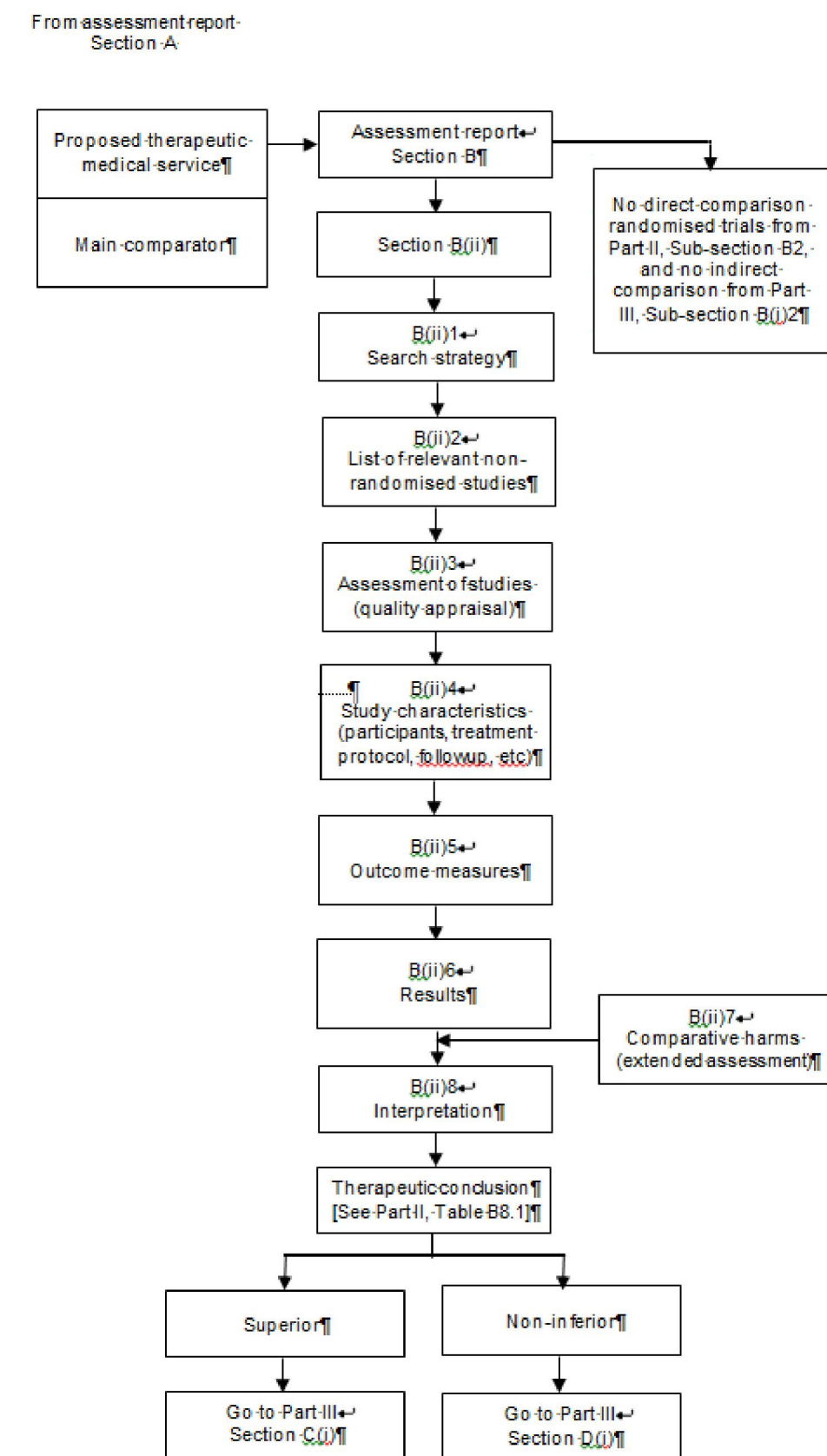
Where relevant, if direct randomised trials comparing the proposed therapeutic medical service directly with the main comparator are available, their analysis and presentation are preferred as the basis of the clinical evaluation (see Part II, Section B). However, in the absence of any such direct randomised trials, the second step in the hierarchy is to determine whether it is possible to present an indirect comparison based on two or more sets of randomised trials involving one or more common reference. Such an analysis indirectly compares the proposed therapeutic medical service with its main comparator by comparing one set of trials, in which participants were randomised to the proposed therapeutic medical service or to a common reference, with another set of trials, in which participants were randomised to the main comparator or to the common reference.

If an indirect comparison (as defined above) is also not possible, the third step in the hierarchy is to present a comparison across non-randomised studies, including comparisons across single arms extracted from randomised trials that do not involve a common reference arm (see Section B(ii)).

The common reference might involve a placebo, such as a sham intervention, but might be another active intervention. There might be more than one common reference (e.g. the proposed therapeutic medical service can be compared with the main comparator via common reference A and via common reference B). In these circumstances, all possible indirect comparisons should be presented and the conclusions compared. The indirect comparison might also involve more than one step (e.g. the proposed therapeutic medical service can be compared with common reference A in one set of randomised trials, common reference A can be compared separately with common reference B in another set of randomised trials, and common reference B can be compared with the main comparator in a third set of randomised trials). In this circumstance of a multistep indirect comparison, there is limited basis for giving guidance on presenting the analysis. The greater the number of steps, the greater the uncertainty associated with the comparison.

This Section gives guidance on presenting a clinical evaluation based on an indirect comparison. The information requests are arranged in the same order, with the same issues for assessing the evidence as those for presenting direct randomised trials. For clarity, assessment reports should adopt the suggested Section headings in the order presented here. A summary of this approach is shown in Figure B(i)1.

**Figure B(i)1** Key information requests for assessment report Section B of a standard assessment for MSAC with clinical data from an indirect comparison of randomised trials



## **B(i)1 Description of search strategies**

### **INFORMATION REQUESTS**

- After demonstrating that no relevant direct randomised trials exist, broaden the literature search criteria to identify all randomised trials relevant for an indirect comparison of the proposed therapeutic medical service and the main comparator.
- Describe the search strategies and characteristics used to locate reports of potentially relevant trials from the published literature, registers of randomised trials and unpublished sources held by the applicant, as described in Part II, Sub-section B1.

If no relevant direct randomised trials have been retrieved in response to the systematic searches requested in Part II, Sub-section B1, broaden the search criteria to identify all randomised trials of the proposed therapeutic medical service and of the main comparator.

This involves relaxing the inclusion criteria to identify all randomised trials involving possible common references (i.e. therapies that are compared with the proposed therapeutic medical service or with the main comparator in separate trials). For the proposed therapeutic medical service, this includes a search internal to the applicant of all trials conducted by, or on behalf of, the applicant, its head office, its subsidiaries elsewhere and any co-licensing applicant.

The search should follow the same methods as described in Part II, Sub-section B1, including provision of a detailed description of the search and printouts of the searches. As it is not possible to pre-specify the common reference(s), these searches should identify, for the proposed therapeutic medical service and for the main comparator, all randomised trials that recruited participants with characteristics that overlap those of patients who would be eligible for the main indication.

## **B(i)2 Listing all randomised trials considered for inclusion in indirect comparisons**

### **INFORMATION REQUESTS**

- Present tables listing all citations of randomised trials for the proposed therapeutic medical service and the main comparator that included a common reference and that recruited participants with characteristics that overlap those of patients who would be eligible for the main indication as identified from the expanded searches of the published literature, marketing dossier and other sources. Show the inclusion and exclusion criteria, and state which trials have been published.
- On the hard copy of each of the search printouts supplied as technical documents with the assessment report, annotate each citation to indicate excluded citations with the reason for the exclusion.
- Collate all reports of each randomised trial included in the indirect comparison to create a master list, arranging the randomised trials into sets for the proposed therapeutic medical service and the main comparator according to each identified common reference. Indicate the preferred ID for each trial to be used throughout the assessment report for consistency.



- Before comparing the proposed therapeutic medical service with the main comparator, establish the comparability of the randomised trials, both within each set and across the two or more compared sets. Justify the exclusion of each randomised trial deemed non-comparable within each set.
- In the absence of any relevant randomised trials to form an indirect comparison, include a 'nil return' in the assessment report.

## **Search results**

Assess all citations retrieved by the expanded searches to extract all trials that meet the following inclusion criteria for randomised trials to support one or more indirect comparisons involving the identified common reference(s):

- the trial included a randomisation procedure in its design;
- the trial compared the proposed therapeutic medical service or the main comparator against an identified common reference in separate arms; and
- the trial recruited participants with characteristics that overlap those of patients who would be eligible for the main indication.

Adapt the guidance given in Part II, Sub-section B2 to present the results of the searches, and to list and provide details of all the randomised trials that meet the inclusion criteria separately for the proposed therapeutic medical service and the main comparator. In addition to the two tables presented to establish that there are no direct randomised trials, replicate the format of those tables to present the expanded searches for all randomised trials of the proposed therapeutic medical service. A fifth table is needed to present the literature searches for all randomised trials of the main comparator (the sixth table might not be needed, because it is unlikely that the applicant would have access to any unpublished randomised trials of the main comparator).

## **Search printouts**

Present annotated search printouts as described in Part II, Sub-section B2.

## **Master list of trials**

From the two tables reporting the results of the expanded searches for the proposed therapeutic medical service, list all identified relevant citations of randomised trials for the proposed therapeutic medical service. Similarly, list all identified relevant citations of randomised trials for the main comparator. Table B(i)2.1 provides a suggested format for presentation of a master list of all the relevant randomised trials identified in the search for the indirect comparison.

**Table B(i)2.1 Trials (and associated reports) presented in the assessment report**

<b>Trial</b>	<b>Description</b>	<b>Reports</b>	<b>Comparable?</b>
<b>Common reference A</b>			
<i>Proposed therapeutic medical service</i>			
Identification(ID) of trial used in remainder of assessment report	Brief description of trial	Internal study report title. Date. Author(s). Title. Journal Year; Vol(No):pages Author(s). Title. Journal Year; Vol(No):pages	Yes/No
<i>Main comparator</i>			
ID of trial used in remainder of assessment report	Brief description of trial	Internal study report title. Date. Author(s). Title. Journal Year; Vol(No):pages Author(s). Title. Journal Year; Vol(No):pages	Yes/No
<b>Common reference B</b>			
Etc.			

If there is no basis for an indirect comparison as defined above, see Figure B1 for the next step in the clinical evaluation. The third step in the hierarchy is to present a comparison across non-randomised studies, including comparisons across single arms extracted from randomised trials that do not involve a common reference arm. To do so, follow Section B(ii) in place of the remainder of this Section B.

### **Presentation of non-inferiority (equivalence) trials**

If an indirect comparison is provided to support a therapeutic conclusion of non-inferiority or equivalence in Sub-section B(i)8, see Appendix 3 for additional guidance on the presentation of the information.

### **Assess comparability of identified randomised trials to justify any exclusions**

Given that there is no randomisation step in the comparison of the proposed therapeutic medical service and the main comparator, it is appropriate, when establishing the comparability of the compared sets of randomised trials, to consider justifying the exclusion of randomised trials from those included in the list above to select similar trials for inclusion in the indirect comparison. The grounds for exclusion might include any aspect reported in Sub-sections B(i)3–B(i)5 (i.e. the quality of the trials, the patient characteristics and circumstances of use, and the outcomes reported in the trials; see examples below). Observable differences across the randomised trials should be minimised, or their contribution to heterogeneity across the trials examined and adjusted where possible. By definition, non-observable differences cannot be minimised or adjusted, and this contributes to the residual uncertainty inevitably associated with indirect comparisons.

### **Aspects that might justify the exclusion of trials from an indirect comparison include:**

- important differences in the quality of the trials (e.g. inadequate follow-up in one of the trials);

- important differences in baseline patient characteristics (e.g. the treatment effects detected in a trial of patients with severe disease might not be comparable with those detected in a trial of patients with mild disease);
- differences in outcomes reported (e.g. a trial might report outcomes that are not assessed in any other trial); and
- differences in the ‘common’ reference — this might not be identical across the trials; for example, an active common reference might have different methods of delivery across the trials (an important aspect because the indirect comparison relies to a large extent on the consistency of the common reference). Specific examples include different levels of experience for a surgical procedure or a different dose of radiation for radiotherapy.

In addition, it might be reasonable to exclude a trial because changes in medical practice and patient characteristics might also mean that nominally similar therapies might not be comparable when the randomised trials have been conducted at different times or in different geographical regions.

It is not possible to give unequivocal guidance on the exclusion of randomised trials from an indirect comparison at this stage. The justification to exclude a randomised trial should anticipate whether this would raise issues of selection bias, while the justification to include a randomised trial should anticipate whether this would raise issues of comparability. If a decision to exclude or include one or more randomised trials is likely to be controversial, it is usually wiser to also present a sensitivity analysis examining whether the decision makes a difference to the conclusions from the overall clinical evaluation.

If one or more trials are to be excluded from an indirect comparison, identify the aspect(s) of each trial that form(s) the reasons for the proposed exclusion (see Table B(i)2.2). Indicate whether each reason relates to the quality of the trials, the patient characteristics and circumstances of use, and/or the outcomes reported in the trials. Present more detail of each aspect (as a minimum, to the extent requested in the relevant text adapted from Part II, Section B). If there is more than one type of reason for exclusion, arrange the trials for exclusion in Table B(i)2.2 by the reason for exclusion.

**Table B(i)2.2 Reasons to exclude each trial from the indirect comparison**

<b>Trial ID</b>	<b>Ground(s) for seeking exclusion</b>	<b>Details <sup>a</sup></b>
Quality of the trial		
Trial 1		
Patient characteristics and circumstances of use in the trial		
Trial 2		
Outcomes reported in the trial		
Etc.		

ID = identification

<sup>a</sup> Cross-reference each set of details to the source of information (specifying the trial report with page, table, figure number)

Table B(i)2.3 shows a suggested format for presenting included trials that are used to indirectly compare the proposed therapeutic medical service and its main comparator. This presentation is useful because it also provides details of the common reference(s) and summarises the comparative strategy adopted for the assessment report.

**Table B(i)2.3 Summary of randomised trials used to conduct the indirect comparison**

Trial ID	Proposed therapeutic medical service	Common references				Main comparator
		Placebo	Active intervention A	Active intervention B	Active intervention C	
Trial 1	X	X	–	–	–	–
Trial 2	X	X	–	–	–	–
Trial #	–	X	–	–	–	X
Trial #	–	X	X	–	–	–
Trial #	–	–	X	X	–	–
Trial #	–	–	–	X	X	–
Trial #	–	–	–	–	X	X

ID = identification; X = treatment included in trial ; – = not tested

### **Trial details**

Present the included comparable trials in the main body of the assessment report and attach a report of each to the main body of the assessment report. Provide a report of each included, but incomparable, trial in a separate volume of the assessment report. Provide clear cross-references between the presentation of the trials and the reports.

## **B(i)3 Assessment of the measures taken by investigators to minimise bias**

### **INFORMATION REQUESTS**

- Adapt the guidance given in Part II, Sub-section B3, including the suggested tables, to describe the minimisation of bias within each included indirect comparison of randomised trials.
- Compare and assess the minimisation of bias in the trials across each set of trials forming the indirect comparison.

It is not possible to minimise bias across the indirect comparison beyond the assessment of comparability and selection bias discussed in the Section above. For trials deemed comparable for the assessment report, identify any differences that might exist in the quality of the trials across the indirect comparison.

## **B(i)4 Characteristics of the trials**

### **INFORMATION REQUESTS**

- Adapt the guidance given in Part II, Sub-section B4, including the suggested tables, to describe the characteristics of each included comparable randomised trial.
- Indicate when and where each included comparable randomised trial was conducted.
- Compare these aspects of the trials across each set of randomised trials forming the indirect comparison and assess any important differences.

The indirect comparison of randomised trials does not include a direct randomisation of patients to the proposed medical service and main comparator, which would allow the characteristics of the patients to differ only due to the play of chance. The description of the characteristics of each randomised trial should facilitate their comparison across the compared sets of trials. For trials deemed comparable for the assessment report, it is particularly important to assess the baseline risk of the patients recruited into the randomised trials and the methods of delivery of the active intervention used for the common reference.

Similarly, assess how far apart in time and place the trials were conducted. This is necessary because changes in medical practice and patient characteristics might mean that nominally similar therapies might not be comparable when the randomised trials have been conducted at different times or in different geographical regions. Such changes might confound the indirect comparison.

## **B(i)5 Outcome measures**

### **INFORMATION REQUESTS**

- Adapt the guidance given in Part II, Sub-section B5, including the suggested tables, to present definitions of the patient-relevant outcomes measured, their natural units of measurement and the duration of follow-up when the outcomes were assessed in each included indirect comparison of randomised trials.
- Compare and assess any important differences in the outcomes measured across each set of randomised trials forming the indirect comparison

The methods of measurement of the same outcome might differ across the trials. The description of the patient-relevant outcomes should facilitate a comparison both within each set of trials and across the compared sets of trials. The distinctions between primary and secondary outcomes, and between primary and secondary analyses are less important in an indirect comparison.

## **B(i)6 Results**

### **INFORMATION REQUESTS**

- Assess the results for each common reference for any important differences across the sets of randomised trials.
- Present the results as follows:
  - for dichotomous outcomes, present the results of each individual randomised trial as the relative risk, with its 95% confidence interval, between the common reference and the proposed therapeutic medical service and the main comparator
  - for time-to-event outcomes, present the results of each individual randomised trial as the hazard ratio, with its 95% confidence interval, between the common reference and the proposed therapeutic medical service and the main comparator
  - where there is more than one randomised trial in a set, separately pool the treatment effect between the common reference and the proposed therapeutic medical service, and between the common reference and the main comparator results as the relative risk (or hazard ratio), with its 95% confidence interval, using the random effects model
  - calculate the indirect estimate of effect as the ratio of relative risks (or the ratio of hazard ratios), with its 95% confidence interval.
- Clearly document and reference any additional or other methods used to quantify the results of the indirect comparison in terms of magnitude of effect and its 95% confidence interval

Assess the results for each common reference across the sets of randomised trials for any important differences. This serves as a check of the comparability of the trials — ideally, the results should be similar for similar outcomes measured in similar patients given the same common reference.

When presenting the results for each randomised trial and for the pooled analysis for each set of trials, calculate relative treatment effects. For the indirect treatment effect across the sets of trials, calculate the ratio of relative treatment effects, with its 95% confidence interval. Using relative treatment effects might help to adjust for any differences in the results of the common reference, and relies on a usual finding that the relative treatment effect varies to a lesser extent across populations than the absolute treatment effect (including different durations of follow-up; see Part II, Sub-section C1 for further explanation of this finding). A suggested manner of presentation is illustrated in Table B(i)6.1.

**Table B(i)6.1 Summary of results of the indirect comparison of randomised trials**

Trial ID	Trial(s) of proposed therapeutic medical service			Trial(s) of main comparator			Indirect estimate of effect <sup>c</sup> Indirect RR (95% CI)
	Treatment effect <sup>a</sup> RR (95% CI)	Proposed therapeutic medical service <i>n</i> with event/ <i>N</i> (%)	Common reference <i>n</i> with event/ <i>N</i> (%)	Common reference <i>n</i> with event/ <i>N</i> (%)	Comparator <i>n</i> with event/ <i>N</i> (%)	Treatment effect <sup>b</sup> RR (95% CI)	
Trial 1							—
Trial 2							—
Etc.							—
Pooled <sup>d</sup>		—	—	—	—		

— = not applicable; CI = confidence interval; ID = identification; *n* = number with event; *N* = number in group; RR = relative risk

<sup>a</sup> proposed therapeutic medical service over common reference

<sup>b</sup> main comparator over common reference

<sup>c</sup> inferred as proposed therapeutic medical service over main comparator

<sup>d</sup> pooled using the random effects model

When documenting and referencing any additional or other methods used to quantify the results of the indirect comparison, ensure that the methods are reproducible and able to be independently verified. For example, if there are enough randomised trials, meta-regression might also be used to analyse and present indirect treatment comparisons.

Where appropriate, assess the implications for the conclusions of the indirect comparison of excluded trials considered to be less comparable (e.g. in terms of trial populations or treatments). Alternatively, justify, describe and present any other adjustment of the indirect comparison.

Where possible, assess whether there is statistical support for the underlying assumption that there is little variation in the relative treatment effect (see Part II, Sub-section C2 for guidance on assessing heterogeneity).

## **B(i)7 Extended assessment of comparative harms**

The presentation of a wider basis of comparative harms is relevant beyond the context of indirect comparisons of randomised trials, as well as beyond that of direct randomised trials (see Part II, Sub-section B7).

## **B(i)8 Interpretation of the clinical evidence**

### **INFORMATION REQUESTS**

- Discuss the results and the interpretation of the indirect comparison of randomised trials cautiously.
- Based on the results of the clinical evaluation, state the category from Part II, Sub-section B8 that best describes the proposed therapeutic medical service.

Discuss the results and the interpretation of the indirect comparison of randomised trials cautiously, due to the inability to minimise important biases, such as selection bias across the indirect comparison.

Refer to Sub-section B8 to determine the therapeutic conclusion from the trials.

## **Section B(ii)**

### **Clinical evaluation for the main indication: Presenting non-randomised studies**

---

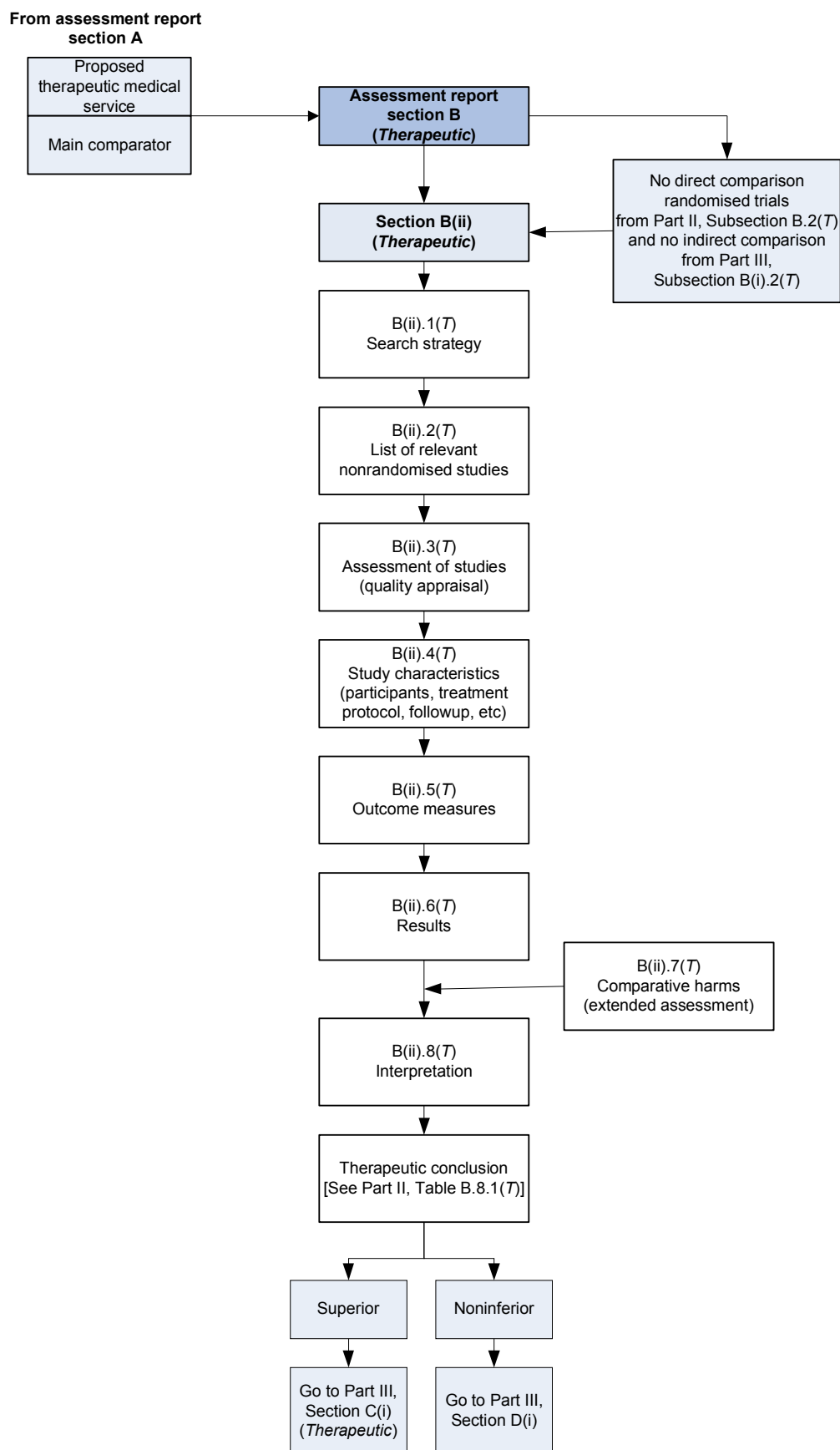
#### **Introduction**

Where relevant, direct randomised trials (as defined in Part II, Sub-section B2) comparing the proposed therapeutic medical service directly with the main comparator are available, their analysis and presentation are preferred as the basis of the clinical evaluation (see Part II, Section B). However, in the absence of any such direct randomised trials, the second step in the hierarchy is to determine whether it is possible to present an indirect comparison of randomised trials as defined in Section B(i). If this is also not possible, the third step in the hierarchy is to present a comparison across non-randomised studies, including comparisons across single arms extracted from randomised trials that do not involve a common reference arm.

This Section provides guidance on presenting a clinical evaluation based on non-randomised studies. The information requests are arranged in the same order, with the same issues for assessment of the evidence, as those for the presentation of direct randomised trials. For clarity, assessment reports should adopt the suggested section headings in the order presented here. A summary of this approach is shown in Figure B(ii)1.



**Figure B(ii)1      Key information requests for assessment report Section B of a standard assessment for MSAC with clinical data from non-randomised studies**



## **B(ii)1 Description of search strategies**

### **INFORMATION REQUESTS**

- Broaden the search criteria to identify all other randomised trials and all non-randomised studies of the proposed therapeutic medical service that recruited participants whose characteristics overlap with those of patients who would be eligible for the main indication.
- Identify all other randomised trials and all non-randomised studies of the main comparator that recruited participants whose characteristics overlap with those of patients who would be eligible for the main indication.

If neither direct randomised trials nor relevant randomised trials to construct an indirect comparison have been retrieved in response to the systematic searches requested in Part II, Sub-section B1 and Part III, Sub-section B(i)1, broaden the search criteria to identify:

- all non-randomised studies of the proposed therapeutic medical service that recruited participants whose characteristics overlap with those of patients who would be eligible for the main indication (conducted by, or on behalf of, the applicant, its head office, its subsidiaries elsewhere and any co-licensing applicant); and
- all non-randomised studies of the main comparator that recruited participants whose characteristics overlap with those of patients who would be eligible for the main indication.

Adapt the guidance provided in Part II, Sub-section 1 to describe the search.

## **B(ii)2 Listing all non-randomised studies**

### **INFORMATION REQUESTS**

- Present tables listing all citations of randomised trials and non-randomised studies that recruited participants with characteristics that overlap those of patients who would be eligible for the main indication as identified from the expanded searches of the published literature, marketing dossier and other sources. Show the inclusion and exclusion criteria, and state which trials have been published.
- Collate all reports of each randomised trial and non-randomised study included to create a master list, arranging the studies for the proposed therapeutic medical service and the main comparator. Indicate the preferred ID for each trial to be used throughout the assessment report for consistency.
- Before comparing the proposed therapeutic medical service with the main comparator, establish the comparability of the studies, especially for the comparison across studies for the proposed therapeutic medical service and studies for the main comparator. Justify exclusion of any study
- Include copies (or sufficient details) of the included comparable studies as attachments in the main body of the assessment report.

### **Search results**

Assess all citations retrieved by the expanded searches to extract all randomised trials and non-randomised studies that recruited participants with characteristics that overlap those of patients who would be eligible for the main indication. Adapt the guidance given in Part II, Sub-section B2 to present the results of the searches, and to list and provide details of all randomised and non-randomised trials. In addition to the tables presented to establish that there are no direct randomised trials and no basis to construct an indirect comparison, replicate the format of those tables to present the expanded searches for all non-randomised studies of the proposed therapeutic medical service. A separate table is

needed to present the literature searches for all non-randomised studies of the main comparator (only one table might be needed, because it is unlikely that the applicant would have access to any unpublished non-randomised studies of the main comparator).

## Master list of studies

From the tables reporting the results of the expanded searches for the proposed therapeutic medical service, list all identified relevant citations of randomised trials and non-randomised studies for the proposed therapeutic medical service. Similarly, list all identified relevant citations of randomised trials and non-randomised studies for the main comparator. Table B(ii)2.1 provides a suggested format for presentation of a master list of all the relevant studies identified in the search.

## Presentation of non-inferiority (equivalence) studies

If non-randomised studies are provided to support a therapeutic conclusion of non-inferiority or equivalence in Section B(ii)8, adapt the additional guidance in Appendix 3 to identify the preferred approach for the presentation of the studies.

**Table B(ii)2.1 Studies (and associated reports) presented in the assessment report**

Study	Description	Reports	Comparable?
<b>Proposed therapeutic medical service</b>			
<i>Single arms of randomised trials</i>			
Unique identification (ID) of study used in remainder of assessment report	Brief description of study	Internal study report title. Date. Author(s). Title. <i>Journal</i> Year; Vol(No):pages Author(s). Title. <i>Journal</i> Year; Vol(No):pages	Yes/No
ID of study used in remainder of assessment report	Brief description of study	Internal study report title. Date. Author(s). Title. <i>Journal</i> Year; Vol(No):pages Author(s). Title. <i>Journal</i> Year; Vol(No):pages	Yes/No
<i>Non-randomised studies</i>			
ID of study used in remainder of assessment report	Brief description of study	Internal study report title. Date. Author(s). Title. <i>Journal</i> Year; Vol(No):pages Author(s). Title. <i>Journal</i> Year; Vol(No):pages	Yes/No
<b>Main comparator</b>			
Etc.			

## Assess comparability of identified studies to justify any exclusions

Before comparing the proposed therapeutic medical service with the main comparator, establish the comparability of the compared sets of non-randomised studies, including single arms extracted from randomised trials. Given that there is no randomisation step across the comparison of the proposed therapeutic medical service and the main comparator, it is appropriate to consider justifying the exclusion of studies from those included in the list above to select similar studies for inclusion in the non-randomised comparison. Possible grounds for exclusion are provided in Sub-section B(i)2.

It is not possible to give unequivocal guidance on the exclusion of studies at this stage. The justification to exclude a study should anticipate whether this would raise issues of

selection bias; the justification to include a study should anticipate whether this would raise issues of comparability. If a decision to exclude or include one or more studies is likely to be controversial, it is usually wiser to present an additional sensitivity analysis examining whether the decision makes a difference to the conclusions from the overall clinical evaluation.

If one or more studies are to be excluded, identify the aspect(s) of each study that form the reasons for the proposed exclusion (see Table B(ii)2.2). Indicate whether each reason relates to the quality of the studies, the patient characteristics and circumstances of use, the outcomes reported in the trials and/or any other reason. Present greater detail of each aspect (as a minimum, to the extent requested in the relevant text adapted from Section B or Appendix 7). If there is more than one type of reason for exclusion, arrange the studies for exclusion in Table B(ii)2.2 by reason for exclusion.

**Table B(ii)2.2 Reasons to exclude each study**

Study ID	Ground(s) for seeking exclusion	Details <sup>a</sup>
Quality of the study (see Appendix 7)		
Study 1		
Patient characteristics and circumstances of use in the study		
Study 2		
Outcomes reported in the study		
Study 3		
Other reasons		
Etc.		

ID = identification

<sup>a</sup> Cross-reference each set of details to the source of information (specifying the trial report with page, table, figure number)

### Study details

Present the included comparable studies in the main body of the assessment report and attach a report of each to the main body of the assessment report. Provide a report of each included but / incomparable study in a separate volume of the assessment report. Provide clear cross-references between the presentation of the studies and the reports.

## B(ii)3 Assessment of the measures taken by investigators to minimise bias

### INFORMATION REQUESTS

- For each included comparable non-randomised study:
  - categorise into the study type(s) defined below; and
  - assess the quality of the study.
- If the assessment report includes a number of studies of the same type, tabulate the responses.

As for the assessment of randomised trials, the purpose of the assessments in this Section is to provide the applicant and MSAC with a clear idea of which studies are of greater scientific rigour by assessing the measures taken by the investigators to minimise bias. There is no minimum standard, but MSAC is most likely to be persuaded by the data of the highest scientific rigour.

There might be other aspects of particular non-randomised studies that might affect the results of such studies and their comparability with different studies of the same type. If these aspects are likely to be important, they should also be identified.

## **Categorise studies**

Non-randomised studies include:

- classical observational designs such as:
  - cohort studies (with concurrent controls); and
  - case-control studies;
- quasi-experimental designs such as:
  - ‘before-and-after’ studies;
  - case series with historical controls; and
  - a comparison of the results of two or more single-arm studies.

Single-arm studies might be extracted from randomised trials when there is no common reference on which to construct an indirect comparison.

See Appendix 7 for definitions of each type of study.

## **Assess quality of studies**

Classical community-based epidemiological designs, such as controlled cohort and case-control studies, can be used to estimate the comparative clinical performance of therapy if randomised trials are not available. However, it has been repeatedly shown that such studies are subject to a range of biases that frequently lead to overestimation of the true benefit of the treatment given to the intervention group. Consequently, claims about comparative clinical performance that are based solely on data from such sources will be treated with some scepticism.

Data from the other types of quasi-experimental non-randomised designs (e.g. before and after studies, case series with historical controls, comparisons of results of two or more single-arm studies) are subject to major and (often) non-quantifiable biases. Consequently, claims about comparative clinical performance that are based solely on data from these types of analyses will be treated with scepticism.

Some criteria that should be used to assess the quality of non-randomised studies are provided in Appendix 7. However, these are for general guidance only and might have to be adapted to particular situations. The interpretation of the results of such studies is difficult, and expert epidemiological guidance will be helpful if data of this type are central to the assessment report.

## **Tabulate responses**

Where there is more than one study of the same type, it is more efficient to present the assessments in a table.

## **B(ii)4 Characteristics of the studies**

### **INFORMATION REQUESTS**

- Adapt the guidance given in Part II, Sub-section B4, including the suggested tables, to describe the characteristics of each included comparable non-randomised study.
- Indicate when and where each included comparable non-randomised study was conducted.
- Compare these aspects of the studies and assess any important differences.

The description of the characteristics of the non-randomised studies should facilitate their comparison across the studies. For studies deemed comparable for the assessment report, it is particularly important to assess the comparability of the patients included in the studies and the treatment regimens used for the proposed therapeutic medical service and, as relevant, for the main comparator.

Similarly, assess how far apart in time and place the studies were conducted. This is necessary because changes in medical practice and patient characteristics might mean that nominally similar therapies might not be comparable when the studies have been conducted at different times or in different geographical regions.

## **B(ii)5 Outcome measures**

### **INFORMATION REQUESTS**

- Adapt the guidance given in Part II, Sub-section B5, including the suggested tables, to present definitions of the patient-relevant outcomes measures, their natural units of measurement and the duration of follow-up when the outcomes were assessed in each included comparable non-randomised study.
- Compare and assess any important differences in the outcomes measured across the non-randomised studies.

When presenting definitions of the study outcomes, the distinctions between primary and secondary outcomes and between primary and secondary analyses, are less important in a comparison involving non-randomised studies.

## **B(ii)6 Results of the comparison involving non-randomised studies**

### **INFORMATION REQUESTS**

- Present the results of all patient-relevant outcomes measured, together with their respective 95% confidence intervals.

In general, the results will be in the form of a proportion, a difference in proportions, an odds ratio or a relative risk. Occasionally, the results will be in the form of a difference in some other response variable (e.g. forced expiratory volume).

## **B(ii)7 Extended assessment of comparative harms**

The presentation of a wider basis of comparative harms is relevant for a comparison involving non-randomised studies (see Part II, Sub-section B7).

## **B(ii)8 Interpretation of the clinical evidence**

### **INFORMATION REQUESTS**

- Discuss the results and the interpretation of the comparison involving non-randomised studies cautiously.
- Based on the results of the clinical evaluation, state the category from Part II, Sub-section B8 that best describes the proposed therapeutic medical service.

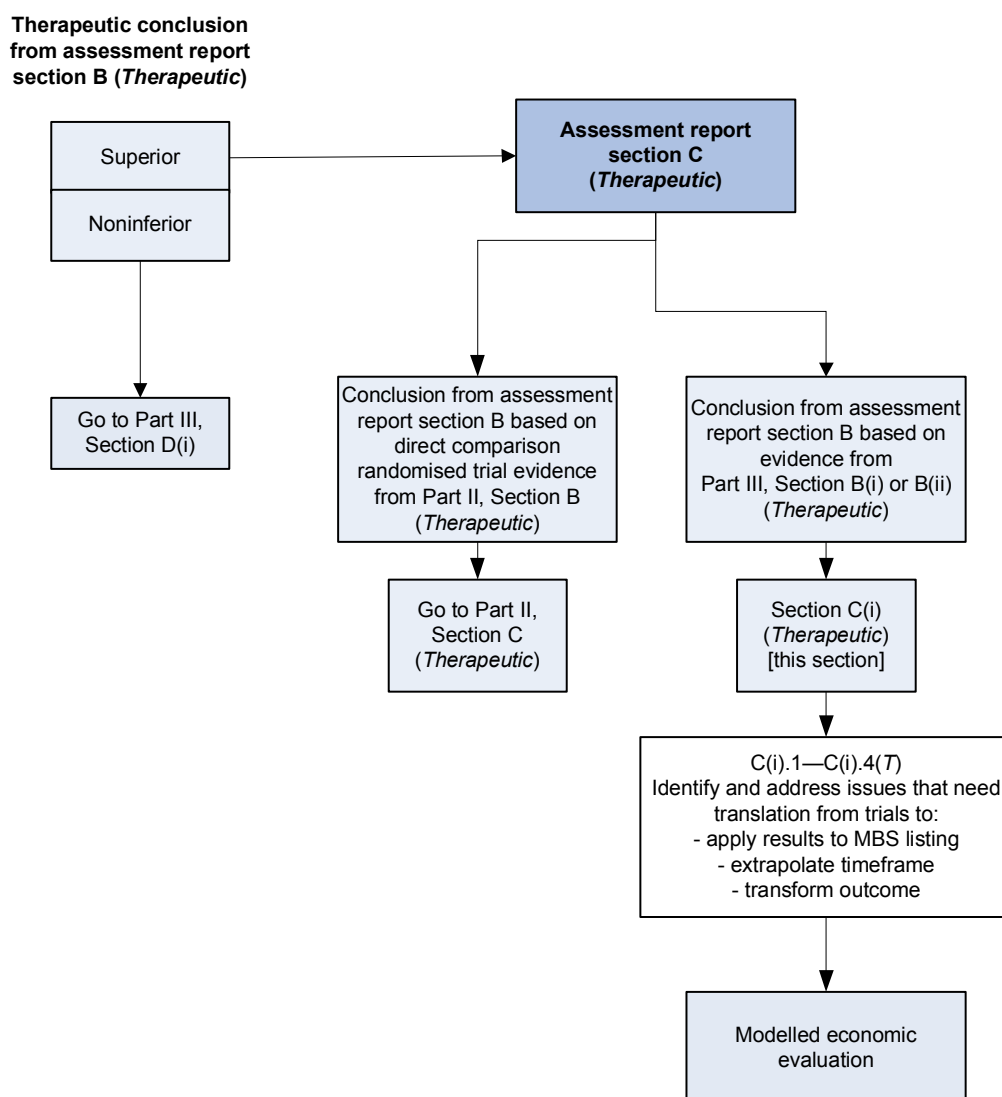
Discuss the results and interpretation of the comparison involving non-randomised studies cautiously because of the inability to minimise important biases such as selection bias. If providing a narrative conclusion in relation to whether the proposed service is either superior or non-inferior to the main comparator (based on non-randomised evidence), provide justification why a statistical conclusion in this regard cannot be reached or is not feasible.

## Section C(i) Translating the clinical evaluation: indirect comparisons of randomised trials or non- randomised studies

---

This Section provides guidance on any pre-modelling studies that might be useful for assessment reports relying on either indirect comparisons of randomised trials (see Section B(i)) or non-randomised studies (see Section B(ii)). Figure C(i)1 shows the relationship between this Section, Sections B(i) and B(ii), and D(i), Part II, Sections B and C.

**Figure C(i)1** Key information requests for assessment report Section C of a standard assessment for MSAC with clinical data from an indirect comparison of randomised trials (Section B(i)) or from non-randomised studies (Section B(ii))





## **C(i)1 Relevant pre-modelling studies**

### **INFORMATION REQUESTS**

- Following the format requested in Part II, Section C, present any relevant pre-modelling studies used to translate the clinical evaluation to the economic evaluation reflecting the population and circumstances of use for the proposed therapeutic medical service.
- Provide copies of all sources of data in an attachment or a technical document (cross-referenced from the main body of the assessment report) and electronic copies of all computer-based analyses.

Although not all the guidance given in Part II, Section C can be applied to clinical evaluations based on indirect comparisons of randomised trials or non-randomised studies, the intention of presenting additional pre-modelling studies to translate the results to an economic evaluation, presented in Section C of the assessment report, might still be relevant. That is:

- identify the issue (see Part II, Sub-section C1);
- present a focused analytical plan (see Part II, Sub-section C2);
- present the results of the pre-modelling study (see Part II, Sub-section C3); and
- identify the relationship between the pre-modelling study and the economic evaluation (see Part II, Sub-section C4).

If any pre-modelling studies are performed in this way, the three steps to improve transparency requested in Sub-sections D1 and D5 might also apply to the economic evaluation for these types of studies.

However, the guidance given in Sub-section C2 on translating the clinical evaluation to the economic evaluation does not usually apply for these types of studies. For example, in these circumstances, there is no basis to guide an assessment of variation in the comparative treatment effect, an extrapolation of comparative time-to-event data or a transformation of a comparative treatment effect on surrogate outcomes to a comparative treatment effect on final outcomes. The resulting economic evaluations are generally either modelled without reference to supporting pre-modelling studies or presented as cost-minimisation analyses.

In presenting the economic evaluation for indirect comparisons of randomised trials in Sub-section D5, also provide the separate incremental cost-effectiveness ratios of the proposed therapeutic medical service against the common reference and of the main comparator against the common reference, together with the results (applying the 95% confidence interval of the respective incremental treatment effects). This enables MSAC to confirm that, if the main comparator is therapeutically superior to the common reference and of acceptable cost-effectiveness, an assessment can be made as to whether a similar conclusion can be drawn for the proposed therapeutic medical service.

## **Sources of information and electronic calculations**

Separately provide copies of the original sources of all data (beyond those already presented in Section B of the assessment report) in an attachment or technical document. Cross-reference the extraction of data from each source to the level of the page, table or figure number of the source document.

Also, to enable independent verification of each analysis, provide an electronic copy of any computer-based calculations of the analysis (see Part I, Section 5).

## **Section D(i)**

### **Economic evaluation for the main indication: presenting a cost-minimisation approach**

---

#### **Introduction**

The purpose of this Section is to present an economic evaluation of substituting the proposed medical service for the main comparator in the context of the listing requested. As already described in Sub-section B8 and shown in Figures D1 and D(i)1, the economic evaluation of the proposed medical service initially depends on whether the therapeutic conclusion shows:

- the proposed medical service is superior to the main comparator; or
- the proposed medical service is non-inferior (equivalent) to the main comparator.

If the proposed medical service has been shown to be superior to the main comparator, presentation of the economic evaluation according to Part II, Section D is appropriate. However, if the proposed medical service has been shown to be non-inferior (equivalent) to the main comparator, cost-minimisation analysis is appropriate (or cost analysis under limited circumstances where the proposed medical service is non-inferior to the main comparator, but has a superior safety profile that generates cost offsets from reduced use of health care resources to manage adverse reactions).

#### **D(i)1 Cost-minimisation analysis**

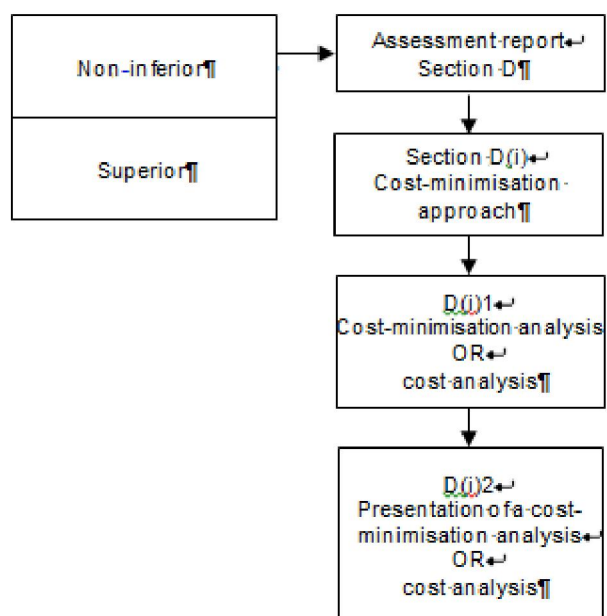
A cost-minimisation analysis applies when the proposed medical service is demonstrated to be no worse (non-inferior) than other medical services at the same or a lower price. Assuming MSAC accepts the alternative therapies as providing acceptable outcomes in terms of both effectiveness and safety for their cost, a new treatment that offers these outcomes at a lower cost is preferable.

##### **Cost analysis**

A cost analysis compares costs only and so is strictly defined as a partial rather than a full economic evaluation, because it does not quantitatively assess comparative costs in a ratio over comparative effectiveness. Although less preferred than a full economic evaluation, cost analyses have sometimes been presented and found to be acceptable if the proposed medical service is demonstrated to be no worse in terms of effectiveness but to have a superior safety profile compared with the main comparator.

**Figure D(i)1      Key information requests for assessment report Section D of a standard assessment for MSAC in which the therapeutic conclusion from assessment report Section B is non-inferior**

Therapeutic conclusion from assessment report Section B



## **D(i)2 Presentation of a cost-minimisation analysis or a cost analysis**

### **INFORMATION REQUESTS**

- Present a cost-minimisation analysis OR a cost analysis.
- Provide copies of all sources of data in an attachment or a technical document (cross-referenced from the main body of the assessment report) and electronic copies of all computer-based analyses.

#### **Cost-minimisation analysis**

When the proposed medical service is regarded as non-inferior to its main comparator in terms of both effectiveness and safety, the appropriate type of economic evaluation is a cost-minimisation analysis. That is, the difference between the proposed medical service and the main comparator is reduced to a comparison of costs.

Such an assessment report need only present an abbreviated Section D, except where there are differences in the costs of delivering the two alternatives. Take particular care to justify any decision to model a difference due to a factor that is excluded in the trials. Only rarely has a model been accepted that contradicts a conclusion from the evidence of direct comparison randomised trials that fail to detect a statistically significant advantage when designed to do so.

If the conclusion of non-inferiority is not also supported by clinical data, the assessment report will be difficult to evaluate.

#### **Cost consequences related to the provision of resources**

Listing a non-inferior medical service might have cost consequences related to its differing mode of administration. These have sometimes arisen if the proposed medical service and its main comparator are available in different forms. If this applies in an assessment report, identify the types of other resources affected, estimate the extent to which the quantity of each type of resource provided would change (in its natural units of measurement) following a listing, and multiply by the relevant unit costs. Aggregate this with the medical service cost impact to estimate the net cost impact within the cost-minimisation analysis.

#### **Cost analysis to reflect cost consequences related to management of adverse reactions**

If the proposed medical service is demonstrated to be no worse in terms of effectiveness but to have a superior safety profile compared with the main comparator, the generally preferred approach would be to compare also the improved health outcomes due to this safety advantage with the associated incremental costs in a cost-consequence, cost-effectiveness or cost-utility analysis (see Part II, Sub-section D1). However, cost analyses have sometimes been presented and found to be acceptable in these circumstances. The cost analysis could be presented to quantify a claim that the costs offsets from the reduction in resources provided to treat the adverse events avoided are sufficient to reduce the incremental cost to zero or a negative value. In a cost analysis, the extent of the health impact would not be assessed other than to estimate the extent to which the provision of the identified types of other resources is reduced i.e. the economic claim could be that, at the MBS fee requested, the overall cost of treatment with the proposed medical service is the same or less than the overall cost of treatment with the main comparator.

### **Sources of information and electronic calculations**

Separately provide copies of the original sources of all data (beyond those already presented in Section B of the assessment report) or expert opinion used in the model in an attachment or technical document. Cross-reference the extraction of data from each source to the level of the page, table or figure number of the source document.

Also, to enable independent verification of each analysis, provide an electronic copy of any computer-based calculations of the analysis (see Part I, Section 5).

## **Appendices**

# Appendix 1 Relevant factors influencing provision of advice by MSAC

---

This appendix provides lists of quantitative and qualitative factors that are relevant to the provision of advice by MSAC

**Table A1.1 Factors that are more readily quantified**

Relevant factor	Description
Comparative safety	Presented as safety of the service under consideration compared with the appropriate comparator(s) when used in the target population and setting
Comparative health gain	Presented as effectiveness This is assessed in terms of both magnitude of effect and clinical importance of effect.
Comparative cost-effectiveness	Presented as cost-minimisation analysis or incremental cost-effectiveness ratios (including incremental cost-utility ratios). Includes a consideration of comparative costs, including the full spectrum of cost offsets
Patient affordability in the absence of MBS subsidy	Presented as cost/patient/course for acute or self-limited treatment, or cost/patient/year for chronic or continuing treatment Calculations for episodic treatment are more difficult.
Financial implications for the MBS	Presented as the projected annual net cost to the MBS
Financial implications for government health budgets	Presented as the projected annual net cost/year



**Table A1.2 Examples of factors that are less readily quantified**

Relevant factor	Description
Uncertainty	<p>The extent and nature of assumptions compared with the extent and nature of data-sourced evidence are important considerations.</p> <p>The presence of uncertainty increases the hesitation involved in making the decision, increasing the likelihood that a risk-averse decision will be made from the perspective of the MBS.</p> <p>Issues which may impact the decision of MSAC include (but are not limited to) uncertainty related to:</p> <ul style="list-style-type: none"> <li>• the direct comparison randomised trial evidence;</li> <li>• an indirect comparison of two or more sets of randomised trials involving one or more common references;</li> <li>• the non-randomised study evidence;</li> <li>• translating the direct comparison randomised trials to the listing requested;</li> <li>• translating an indirect comparison of randomised trials or non-randomised studies to the listing requested;</li> <li>• the economic evaluation;</li> <li>• cost minimisation;</li> <li>• the utilisation and financial estimates; or</li> <li>• the plausibility of the valuation of health outcomes.</li> </ul>
Equity	<p>Affordable access is a central policy principle of the MBS (Part II, Section F) and is considered alongside the economic evaluation.</p> <p>There are many implicit equity and ethical assumptions in the use of quality-adjusted life-years gained; for example, age and socioeconomic and geographical status (Part II, Section D). This means that these assumptions might also need to be reconsidered alongside the economic evaluation on a case-by-case basis.</p>
Presence of effective alternatives	<p>This distinguishes between:</p> <ul style="list-style-type: none"> <li>• an active comparator or placebo for add-on treatment; and</li> <li>• a placebo for no active intervention.</li> </ul> <p>It also helps to define the clinical need for the proposed medical service.</p>
Severity of medical condition treated	<p>This depends on any restriction requested in Part II, Sub-section A2.</p> <p>The emphasis here is only on the nature and extent of disease as it is currently managed (as described in Part II, Sub-section A2).</p>
Ability to target therapy with the proposed medical service precisely and effectively to patients likely to benefit most	<p>If the proposed medical service appears not to be acceptably cost-effective across the broader population, it might become acceptably cost-effective in patients likely to benefit more than the average (assuming costs of the treatment do not increase proportionally).</p> <p>This aspect is usually discussed in Part II, Sub-section A2 (and can influence the choice of comparator in Part II, Sub-section A5).</p> <p>Claims of benefits greater than the average result from the ITT analysis should be supported by appropriate trial evidence (see Part II, Sub-section C1).</p>

## Appendix 2 Expert opinion

---

This appendix outlines the situations in which expert opinion can be used, and explains how expert opinion should be collated and presented in an assessment report.

Expert opinion, where sought, will be considered in conjunction with advice provided to MSAC by the PASC and the ESC during the assessment report assessment and evaluation stages, respectively.

### A2.1 Uses of expert opinion

Expert opinion is not a substitute for sound scientific evidence. Therefore, expert opinion is only considered where there are no observed data available, or where such data addressing the matter for which expert opinion has been sought are unlikely to become available in the near future. Observed data might come from randomised trials or non-randomised studies, including from cross-sectional studies or case studies. Expert opinion can also supplement observed data; for example, to review the likely representativeness to the national level of a cross-sectional study conducted in a single locality or in another country. Such supplementation will help the interpretation of observed data, and therefore reduce its uncertainty.

Expert opinion can be useful in several aspects of preparing assessment reports for MSAC; for example, to help:

- define the clinical need for the proposed medical service and thus the context of its use by defining the proposed medical service's place in treatment in terms of the main indication(s) based on what should be recommended (see Part II, Section A), and the main comparator(s) and clinical management algorithms based on what is likely to change (see Part II, Section A);
- interpret the clinical importance and patient relevance of the outcome measures reported in the trials (see Part II, Section B);
- modify the patterns of resource use and, very rarely, the clinical outcomes measured in randomised trials conducted in different settings, such as in other countries (see Part II, Section B);
- predict which resources would be used and how often each would be used to manage outcomes reported in the randomised trials, but not followed up (see Part II, Section C);
- identify the proportion of patients with the medical condition who would meet the eligibility criteria established by the requested restriction (see Part II, Section E2);
- predict the proportion of patients within this eligible population who would take the proposed medical service (see Part II, Section E2);
- predict the rates of uptake of the proposed medical service (see Part II, Section E2); and
- predict the extents of substitutions, increases and decreases of other medical services that are MBS-funded (see Part II, Sub-section E3).

## A2.2 Presenting expert opinion

### INFORMATION REQUESTS

- Present expert opinion as a technical document or an attachment to the assessment report, with clear cross-references to the relevant Sections of the main body of the assessment report.
- Justify the need for expert opinion.

If expert opinion is included, its use should be justified in the introduction of the Section involved. Include a clear rationale for, and the aims of, eliciting the expert opinion. Where expert opinion is used to fill in a gap in information, describe the nature of this gap clearly and indicate the steps that have been taken to address the gap, such as a literature search.

## A2.3 Describing the collection and collation of expert opinion

### INFORMATION REQUESTS

- Describe and justify the approach chosen to elicit expert opinion.
- Describe the methods used to obtain and collate the opinions, and summarise the opinions together with the extent of any variability in the opinions (see Table A2.1).
- Indicate how the opinions have been used in the main body of the assessment report and justify the approach used in the sensitivity analysis (see Part II, Sections D6 and E6) to reflect any variability in the opinions obtained.

Using a well-designed methodology to elicit expert opinion helps to reduce uncertainty. The methods used might vary from large, published questionnaires and surveys with statistical analysis to a summary of interviews with a panel of clinical experts. Expert opinion might be presented as qualitative or quasi-quantitative information.

There are many approaches to addressing information gaps. The choice of the preferred approach might be influenced by the availability of existing surveys, small numbers of prescribers with appropriate expertise and resource limitations (e.g. time). Options for primary collection of opinions include interviews, focus groups, self-administered questionnaires and telephone surveys. If the survey is to determine what changes a prescriber might make to their prescribing behaviour, ensure that the hypothetical future scenario is clearly detailed.

When summarising the opinions and their variability, interpret the findings and discuss the limitations and biases of the method chosen. Indicate how the opinions have been used in the main body of the assessment report.

Where multiple sources of expert opinion are available to address a single assumption or estimate, compare the results and assess their concordance or lack of it. Where expert opinion is used to modify estimates from randomised trials or non-randomised studies, particularly estimates reported in Part II, Sections B6 or C2 or any other input into the economic evaluation in Part II, Section D4, compare the results and justify the modification. Present a summary table that compares multiple sources or multiple variables. Table A2.1 provides guidance on the details that should be included.

**Table A2.1 Methods to collect and collate expert opinion**

Information to be provided	Notes
The criteria for selecting the experts	Prefer: <ul style="list-style-type: none"> <li>• a random or comprehensive set of medical providers likely to deliver the proposed therapeutic medical service, OR</li> <li>• the appropriate medical specialty group.</li> </ul>
The number of experts approached <sup>a</sup>	
The number of experts who participated <sup>a</sup>	Assess whether the extent and characteristics of the non-responders are likely to diminish the representativeness of the opinions provided, compared with the intended sample approached.
Declaration of potential conflict(s) of interest from each expert or medical specialty group whose opinion was sought	Provide a signed statement from each expert and specialty group specifying any potential conflict of interest and stating the nature of any contractual arrangement, including how much payment was offered and accepted. Where the collection of expert opinion has been contracted out, the contractor should provide this statement, reporting on both the arrangements made between the applicant and the contractor, and the arrangements made between the contractor and those whose opinions were sought.
The background information provided and its consistency with the totality of the evidence provided in the assessment report	Include a copy of any background information provided in the technical document or attachment. If background information has been provided, it might help to ask the experts to define the comparative clinical place of the proposed medical service and the main comparator based on this background information. Including the experts' definitions in the technical document or attachment would allow an assessment of the consistency of the background information with the evidence provided in the assessment report.
The method used to collect the opinions	For example, were the experts approached individually or was a meeting convened? Was any incentive used to maximise responses?
The medium used to collect the opinions	For example, was information gathered by direct interview, telephone interview or self-administered questionnaire?
The questions asked <sup>b</sup>	Explain the design of the tool (quantitative or qualitative). Describe its development. Indicate whether it was pilot-tested and, if so, provide the results of that testing and explain how the results were used to improve the questions. On a question-by-question basis, assess: <ul style="list-style-type: none"> <li>• the extent to which each question is neutral or biased</li> <li>• the extent to which each question is open or closed.</li> </ul> To allow an independent assessment to be made, include in the technical document (or as an attached copy) the questionnaire or an outline of the interview questions.
Whether iteration was used in the collation of opinions and, if so, how it was used	The Delphi technique, for example, uses an iterative approach.

Information to be provided	Notes
The number of responses received for each question <sup>a</sup>	Assess whether the extent of any nonresponse is likely to diminish the representativeness of the opinions provided to particular questions, compared with the intended sample approached.
Whether all experts agreed with each response, and, if not: (i) the approach used to finalise the estimates  And (ii) the approach used to present the variability in the opinions.	For example, the majority opinion or a Delphi technique could be applied; for quantitative results, point estimates (such as the mean, median or the mode) could be presented.  For example, present the range of opinions including common and outlying views expressed; for quantitative results, measures of variance (such as confidence intervals, range, centiles) could be presented.

**a** Tabulate these information items

**b** The way the questions are asked is an important source of potential bias in obtaining expert opinion. A particularly influential extension question extends the respondent beyond 'what' the opinion is (e.g. what would be done, what extent of benefit would be clinically important) to also ask the reason 'why' (e.g. explain why would you do this, explain why is this important). Conveying these reasons alongside expert opinion-based estimates might help improve their acceptability, particularly if a small group of experts has been approached. Including these explanations in the technical document or attachment would allow the opinions to be assessed on the basis of the underlying reasoning, rather than only depending on the authority of the experts.

## Appendix 3 Assessment of non-inferiority

---

### A3.1 Introduction

Non-inferiority means that, in terms of effectiveness, the proposed medical service is no worse than its main comparator. It is used to support a claim of equivalence, because it is not adequate to demonstrate the *absence* of a statistically significant difference between the treatments to claim equivalence; such a lack of a significant difference might occur when the trials are too small to demonstrate a real difference in the effects of the compared services. The appropriate comparison to present is the point estimate of the difference with its 95% confidence interval. This allows MSAC to assess whether the confidence interval contains the minimal clinically important difference.

Thus, an assessment report should support any conclusion for non-inferiority with the information contained in its assessment report Sections as referred to below.

### A3.2 Service delivery information

As part of the information provided in Section B of the assessment report (Sub-section B4 or equivalent), ensure that the service delivery relativity used in the clinical trials is appropriate. Any conclusion of non-inferiority should be accompanied by a determination of how the compared services are each delivered to achieve equi-effectiveness.

### A3.3 Non-inferiority threshold

As part of the information provided in Section B of the assessment report, explain and justify on clinical or other grounds the value of the non-inferiority threshold difference in treatment effect between the proposed medical service and its main comparator. Show how a difference greater than this nominated non-inferiority threshold difference would be clinically important. A specifically designed non-inferiority direct randomised trial would have specified a non-inferiority threshold in its power calculation, and so might have provided one or more grounds to justify this threshold as a pre-specified minimal clinically important difference (MCID). Demonstrate that a systematic approach has been taken in the search for relevant and appropriate references to support the nominated threshold and provide the supporting citations, including any references to one or more regulatory agencies that might have provided guidance on any such thresholds in medical conditions similar to the proposed main indication.

If the basis of the clinical evaluation is an indirect comparison of randomised trials and the nominated non-inferiority threshold relates to an absolute comparison (e.g. absolute risk difference or weighted mean difference) rather than a relative comparison (e.g. relative risk or odds ratio), discuss the issues raised by relying on an indirect comparison of the difference between absolute treatment effects rather than on an indirect comparison of the ratio of relative treatment effects.

### **A3.4 Method of analysis**

Also as part of the information provided in Section B of the assessment report, indicate whether the analysis of each trial was conducted on a per protocol basis (which is appropriate for an analysis in support of a conclusion of non-inferiority, because it helps examine any impact on the conclusions of losses to follow-up or poor compliance), as well as the standard ITT basis (which is the generally preferred basis for an analysis).

If one or more specifically designed non-inferiority direct randomised trials are available, also describe the primary analysis of non-inferiority in detail for each such trial, including the pre-specified non-inferiority threshold (or MCID) used in the power calculation and whether the preferred per protocol basis rather than the ITT basis was used in the context of this non-inferiority analysis. Comment on any differences in the pre-specified non-inferiority thresholds across these trials and with the nominated non-inferiority threshold.

For any direct randomised trial that was not designed as a non-inferiority trial, also describe its primary analysis in detail, including the pre-specified MCID used in the power calculation.

### **A3.5 Presenting an assessment of non-inferiority**

#### **Assessing non-inferiority based on an indirect comparison of randomised trials**

As part of the information provided in response to Part II, Section B6, present the results of each comparative analysis using, where possible, both the per protocol and the ITT basis of each trial with their 95% confidence intervals in a way that allows for direct comparison with the nominated non-inferiority threshold identified. Comment on any differences between the results for the per protocol and ITT populations. Where there is more than one trial reporting the same outcome, statistically combine these results using the random effects method and, where possible, both the per protocol and the ITT basis. Report each result with its 95% confidence interval in a way that similarly allows a comparison with the nominated non-inferiority threshold. Comment on any differences between the results for the per protocol and ITT populations. If the per protocol basis differs across trials, justify the approach to resolve this in the meta-analysis.

If one or more specifically designed non-inferiority direct randomised trials are available, also report the results and stated conclusion of the primary analysis of non-inferiority for each such trial. Report whether the entire 95% confidence interval of the treatment effect between the two medical services is more favourable to the proposed medical service than the pre-specified non-inferiority threshold corresponding to the proposed medical service being less effective. If so, there is statistical support to the conclusion of non-inferiority based on an appropriate pre-specified trial design.

If the primary analysis of a specifically designed non-inferiority direct randomised trial does not present the 95% confidence interval, and/or adopt a per protocol population basis for the analysis, and/or compare this interval with the non-inferiority threshold identified, then present the results, where possible, using both the per protocol and the ITT basis of each trial with their 95% confidence intervals in a way that allows for direct comparison with this threshold for non-inferiority. Discuss whether these results might influence the conclusion of the primary analysis of the trial.

For any direct randomised trial that was not designed as a non-inferiority trial, also report the results of the primary analysis as pre-specified. Report whether the entire 95% confidence interval of the treatment effect between the two medical services is more favourable to the proposed medical service than the pre-specified MCID corresponding to the proposed medical service being less effective. If so, there is post hoc statistical support to the conclusion of non-inferiority. Investigate whether the conclusion of non-inferiority is impacted by a comparison of an analysis conducted on a per protocol basis and/or whether the 95% confidence intervals compared with the non-inferiority threshold identified would modify this conclusion. Report these investigations.

Supplementary analyses might be helpful to support conclusions of non-inferiority that have to rely on primary outcome analyses that were not adequately powered to assess non-inferiority. Base these supplementary treatment comparisons on the results for secondary outcomes that are known to be most responsive to change.

### **Assessing non-inferiority based on an indirect comparison of randomised trials**

The general approach described above for direct randomised trials needs to be adapted for an indirect comparison of randomised trials in response. Report the point estimates for the indirect relative treatment effect with their 95% confidence intervals in a way that allows for direct comparison with the nominated non-inferiority threshold for inferiority. Report whether the entire 95% confidence interval of the treatment effect between the two medical services is more favourable to the proposed medical service than this non-inferiority threshold corresponding to the proposed medical service being less effective. If so, there is indirect statistical support to the conclusion of non-inferiority.

Where possible (and appropriate, noting that there is no basis for a pre-specified non-inferiority design for an indirect comparison of randomised trials), provide additional investigations and supplementary analyses as described above for direct randomised trials.

## **A3.6 Assessing comparative harms in the context of non-inferiority**

As part of the information provided in Section B of the assessment report, examine whether the extended assessment of comparative harms also supports a conclusion of non-inferiority.

## **A3.7 Interpretation of the clinical evidence**

As part of the information provided in Section B of the assessment report, discuss any results to support a conclusion for non-inferiority in the context of the similarity or otherwise of the mechanism of action(s) of the proposed medical service and the main comparator.

If providing a narrative conclusion in relation to whether the proposed service is non-inferior to the main comparator, provide justification why a statistical conclusion in this regard cannot be reached or is not feasible. For example, for applications that predominantly contain non-randomised evidence statistical proof on non-inferiority may not be feasible. However, where the quality of the evidence base is such that it lends itself to statistical methods to come to a conclusion of non-inferiority, these methods as opposed to narrative conclusions is preferred.



## Appendix 4      Utility valuation of health outcomes

---

### A4.1 Use of health-related QALYs gained and cost-utility analysis

The QALY is a measure of adjusted survival time where the adjustment is by means of health-related quality-of-life preference weights derived for specific health states. Expected survival time in each of these health states is adjusted using the preference weights and then summed across the duration of survival to generate expected QALYs gained. The use of preference weights distinguishes QALYs from other quality-of-life measures.

The QALY has become widespread as a measure of health outcome in the economic evaluation of medical services. The key characteristics of the QALY are as follows:

- It combines extension of life and quality of life in a single index that allows comparison across medical services.
- The utility weight index measures strength of preference on a cardinal index anchored on a 0 to 1 interval of death to full (perfect) health, with equal intervals measured in such a way as to have equal value and an allowance for the existence of health states perceived to be worse than death (i.e. <0).
- The utility weights that underpin the QALY measure are based on a sample of individual preferences. These preferences are obtained in a way that involves a trade-off between quality and quantity of life. This provides some validity to the QALY as representing societal trade-offs and therefore social values.

The implication of using this scale is that one year of life in full health is counted as one QALY. Even though one year of life in normal health is less than one QALY, this does not necessarily mean that all incremental QALY gains are numerically smaller than incremental life-year gains. This is because incremental QALY gains can also encompass the possibility of improving quality of life, and such improvements can happen for a long period before any improvement in survival happens.

Theoretically, at least, the QALY provides a measure of health outcomes that is comparable across medical services. This form of analysis should therefore be considered whenever it is appropriate to the outcomes of the proposed medical service. However, many concerns over the estimation of QALYs have been documented.

Guidance on when a cost-utility analysis should be presented is provided in Part II, Sub-section D1.

Other relevant factors (see Part II, Section F and Appendix 1) should be considered alongside, not within, a cost-utility analysis. These include prognosis, severity, age, distributional effect, context (e.g. emergency or prevention), and other equity and ethical issues that are ignored in measurements using a MAUI. Therefore, an assessment report should draw these issues to the attention of MSAC where this is thought important and relevant.

## A4.2 Obtaining utility weights

Several approaches to obtaining utility weights are discussed in these Guidelines:

- using a MAUI in a direct randomised trial;
- creating scenarios to indirectly elicit utility weights;
- directly eliciting utility weights in a randomised trial;
- obtaining a sample of patients matched to trial participants and eligible patients, and using a MAUI;
- mapping results of other quality-of-life instruments to the utility weight anchors of a 0 to 1 interval of death to full (perfect) health; and
- reporting utility weights from published sources.

The generally preferred method of measuring QALYs is by the repeated application of a valid, reliable and responsive MAUI questionnaire to participants in a direct randomised double-blind trial, together with the application of an appropriate scoring algorithm.

However, it is recognised pragmatically that such instruments are not routinely included as an outcome measure in many trials, so it is anticipated that there will be a lag time before this preference can be met routinely. It is also recognised that in many cases it will be necessary to attach utility weights to health states that are not observed within a trial; for example, because they are the result of events that occur outside the trial time frame. Accordingly, guidance is also provided on alternative approaches (see Sub-sections A4.4 and A4.5 of this Appendix). In some circumstances, it is possible that an alternative approach would be preferred to the use of a trial-based MAUI (see Sub-section A4.4 of this Appendix).

### Post-trial transformation to estimate preference weights ('utilities')

Preference weights are preferably generated directly from a trial using MAUIs or might subsequently be elicited with the aid of scenarios. Several other approaches have been presented in major assessments, and are discussed and assessed briefly below in Sub-section A4.5 of this Appendix. MAUIs and scenario-based elicitation of preference weights are further assessed in Sub-sections A4.3 and A4.4 of this Appendix, respectively.

### MAUIs (multi-attribute utility instruments)

MAUIs have three defining elements:

- A generic health-related quality-of-life instrument. Those recommended in Part II, Sub-section B.5 have been assessed according to the criteria for such instruments identified. This element of a MAUI is a descriptive system (a questionnaire containing a set of items or statements with multiple response categories) that provides a description of the health-related quality of life of each respondent.
- A scaling technique, such as TTO or SG. This is used to derive preference-based rankings for a sample of the health states covered by the descriptive system.
- A model, which is used to extrapolate from this sample to generate cardinal weights for all health states covered by the descriptive system (i.e. to develop a preference-based scoring algorithm for the MAUI). Both mathematical and statistical models have been used to provide utility weights for any health state that can be described by

the instrument in terms of its dimensions and levels. For these utility weights to be meaningful for an economic evaluation, the scaling technique must reflect the trade-offs that individuals are willing to make between health outcomes.

Together, these elements generate the unique advantage of trial-based measurement with a MAUI, which is that the direct observation of the actual health states experienced in the trial can be used to generate utility weights in an acceptable way using utility scores of the health states that have been generated in a separate population-based study. Therefore, it is the combination of these three elements that enables acceptable post-trial transformations to estimate utility weights (see Sub-section A4.3 of this Appendix).

### **A4.3 Trial-based utility valuation of health outcomes**

#### **Measurement of QALYs using a trial-based MAUI**

For MAUIs, the measurement of the health state happens in the trial itself, which enables more accurate and unbiased measurement of the health states as experienced by the patients receiving the relevant treatments. The valuation step is then inferred using an acceptable scoring algorithm, which means that the valuation is conceptually and practically separated from the assessment of the particular disease or treatment, and therefore not subject to bias.

To maximise comparability across assessment reports, it would be ideal to request that a single ‘off-the-shelf’ MAUI be used in randomised trials across all assessment reports presenting a cost-utility analysis. Criteria to guide the selection of such an instrument include that it is valid, reliable and responsive, and that it uses an acceptable scoring algorithm and an acceptable preference elicitation technique. However, in practice, no single MAUI has demonstrated unequivocal superiority against all the others and no single MAUI has been universally accepted. There is also debate about whether generic MAUIs are sufficient to capture all important disease-specific factors that might be relevant for particular disease pathways and treatments. The advantages and disadvantages of trial-based MAUIs are discussed further below.

#### **Advantages of relying on trial-based MAUI data**

Trial-based MAUI data has the following advantages:

- (a) It promotes comparability across cost-utility analyses.
- (b) It minimises bias by eliminating the need for an analyst intermediary.
- (c) It can appropriately minimise observer bias by assessing the subjective outcome of health-related quality of life under appropriate blinded conditions.
- (d) It minimises the information asymmetry of the health state being assessed because the trial participant is directly measuring the health-related quality of life of the health state as it is being experienced.
- (e) It applies the scoring algorithm of the general population (which can minimise a source of uncertainty if this was elicited in an Australian population or possibly from socioeconomically similar countries with similar life expectancy) to take responses from the MAUI questionnaires to generate utility weights using an acceptable technique. In other words, the utility scores in the scoring algorithm have been elicited separately from the reporting of the responses in the trial context for each MAUI. The utility weights are calculated by a validated linkage between the response from the MAUI questionnaire in the trial and the utility score inferred for that response from respondents in the general population using the scoring algorithm.

- (f) As a direct translation, it minimises the number of steps between the direct trial-based measurement of health-related quality of life and its valuation.
- (g) It estimates some of the distribution and heterogeneity variation of health states in a population.
- (h) It maintains a fixed period of assessment to which the MAUI applies.
- (i) Repeatedly applying the MAUI during the trial allows for direct conversion into the net present value of the future flow of realised QALYs gained and incremental QALYs gained and might provide a basis for extrapolation beyond the horizon of the trial.
- (j) It provides a benchmark against which to compare any more specific elicitation of preferences presented as supplementary evidence (e.g. using a scenario-based approach; see Sub-section A4.4 of this Appendix).
- (k) It provides advantages for applicants and analysts in terms of time and cost to assess the appropriateness of using an acceptable ‘off-the-shelf’ MAUI in a trial.
- (l) It provides efficiency advantages for respondents and analysts, because no MAUI developed so far takes more than five to eight minutes to complete when self-administered (and less when using computer-based, interviewer-administered questionnaires) and because analysis of each of the main MAUIs is well developed.
- (m) The main MAUIs have been developed with the objective of having international applicability, so it is anticipated that this preference for trial-based MAUI utility weights will have increasing relevance over time to the multinational trial programs for new medical services.
- (n) It is possible to conduct an independent and peer-reviewed verification of any preferred MAUI — including its reliability, validity and responsiveness, the clinical importance of any differences detected by the instrument, and other desirable psychometric properties.
- (o) The use of a consistent MAUI would allow replication (and potentially meta-analysis) of results across similar direct randomised trials conducted between the proposed medical service and its main comparator.

### **Disadvantages of relying on trial-based MAUI data**

Trial-based MAUI data has the following disadvantages:

- (a) The MAUI might be relatively insensitive to the patient-relevant outcomes affected by the proposed medical service, particularly if its main treatment effects or the impacts of the medical condition do not fall within the domains examined by the MAUI. This interpretation of the results needs to be assessed against the possibility of a true negative (i.e. that the proposed medical service has no overall perceptible incremental effect on utility; see also Sub-section A4.4 of this Appendix). The MAUI should therefore be demonstrated not to fit the context of the proposed medical service and the medical condition by comparing the results from the MAUI with an accepted nonutility quality-of-life instrument, such as the SF-36.
- (b) It is unlikely that, in the near future, a randomised trial would be designed to have the MAUI as its primary outcome. The trial might therefore be underpowered to detect a difference using the MAUI. As with all secondary outcomes, the results of the MAUI would need to be assessed with reference to the conclusion from the primary analysis of the trial.

- (c) Trial participants might not be directly representative of the population for whom listing is requested, although an assessment of the distribution and heterogeneity of the results of this outcome might provide a basis for applying them to the targeted population.

### **Trial-based direct elicitation of utility weights**

Conceivably, direct methods might be used within a trial to ask patients to value their current health state at baseline (or over a recent period of time at baseline), and at one or more time points during the trial follow-up (or over a recent period of time at each time point). Advantages (a)–(d), (f) and (h) listed above would also apply to trial-based direct elicitation of utility weights.

The main disadvantage for direct elicitation in the trial setting is the time horizon assumption for TTO or SG (i.e. the trial participant is required to answer a hypothetical question assuming that they remain in the current health state for the rest of their life expectancy). In a scenario-based setting, the entire framework is hypothetical, so there is less risk of any distortion arising from the respondent first having to conceptualise what it might mean to remain in the current health state for a prolonged period.

This approach might also raise potentially important issues to do with adjusting utility weights for groups of patients in certain disease groups (e.g. quadriplegics) and with different adaptations. The defined range of a utility scale is full health (1) to death (0), but people with cancer and other diseases adapt (or adjust up) their estimate of utility closer towards 1 — such people's 'normal health' might be considerably less than 1, but they adapt up to 1. This potentially biases against the allocation of further health resources (so-called double jeopardy). Some groups, when making the adjustment, could also eliminate their capacity to benefit.

### **Presenting trial-based direct elicitation and results**

If utility weights have been directly elicited in a randomised trial, provide details of the method used and justify the selection of the approach taken (e.g. SG or TTO; interview-based and/or computer-based). The same considerations for the design of the preference elicitation task apply in this context as in a scenario-based approach (see Sub-section A4.4 of this Appendix). Report and assess the results as for MAUIs, above.

## **A4.4 Scenario-based utility valuation of health outcomes**

### **Background**

As discussed in Sub-sections A4.2 and A4.3 of this Appendix, obtaining utility weights using a MAUI within the context of a direct randomised double-blinded trial is the preferred method. This Section of this Appendix presents a less preferred alternative, because there is an expected lag time before most major assessments would be able to report utility weights on this basis. Furthermore, given that most randomised trials are designed overseas, few randomised trials would be conducted primarily to ensure that useful economic information is generated from this preferred source of evidence for MSAC and similar decision makers.

An assessment report might seek to justify the inclusion of a scenario-based approach to valuing health states in utility weights as supplementing trial-based utility weights. Alongside this justification for providing these supplementary estimates, present both sets of methods and results, and comment on the interpretation of the results compared to each other. As with the interpretation of the results of any measure of health outcomes, any

claim for an improved sensitivity in quantifying the utility weight of smaller advantages needs to be assessed against the possibility of a true negative (i.e. that the proposed medical service has no overall perceptible incremental effect on utility; see also Sub-section A4.3 of this Appendix). Document the evidence that supports any claim that any difference in results between trial-based utility weights and scenario-based utility weights is attributable to the special characteristic of the health state and not some idiosyncrasy in the utility measurement procedures that have been adopted. This would help justify any apparent diminution in comparability across assessment reports that provided trial-based utility weights. Similarly, if using a scenario-based utility valuation to capture the impacts of health outcomes only occurring beyond the horizon of the trial, document the evidence that supports any claim that the scenario-based utility weights reflect the trial-based utility weights (e.g. by including one or more health states captured and valued within the trial as part of the scenario-based utility valuation study).

Other situations where a scenario-based approach might supplement trial-based utility weights include those in which:

- the health states are associated with quantitatively important ‘ex ante’ anticipated factors (in which one or more elements of the health state are anticipated rather than experienced, so that concepts such as anxiety, risk aversion, fear, hope or dread might be captured) or non-health outcome factors, such as convenience; and
- the health outcomes are significantly affected by prognosis.

If the introduction of the proposed medical service is expected to induce a succession of changing health states that have a significant interactive effect on utility and the composite utility is not equal to the sum (in which a profile of health states would need to be valued), this then suggests that the QALYs approach is unlikely to be suitable, and an alternative and technically more complex approach might be more appropriate, such as a healthy-year equivalents approach.

An assessment report might need to present a scenario-based approach to valuing health states as utility weights in the absence of any trial-based utility weights. In this situation, the main objective of achieving a comparable approach across assessment reports is diminished. Furthermore, many of the issues in interpreting scenario-based utility weights in the absence of trial-based utility weights are similar in nature to the issues in interpreting any results of non-randomised studies in the absence of a direct randomised trial. In particular, it is difficult to minimise the many sources of analyst bias that are intrinsic to this approach (including in the unblinded nature of the construction and presentation of the scenarios, the design of the methods to elicit values and the analysis and interpretation of the results, which are all conducted after the trial results are known).

A particular source of potential biases can be identified with post-trial scenario-based approaches to valuing health outcomes. This is because there is a justifiable preference for eliciting these values from individual respondents drawn from the general population (because they might better reflect the perspective of society overall as representing the balance of taxpayers and patients) rather than of patients alone (who are likely to recognise that they would be the beneficiaries of any new subsidised intervention). However, this inevitably leads to an information asymmetry for the respondent in relation to each specific post-trial scenario in a scenario-based utility study. Seeking to address this information asymmetry by loading more information into the scenarios raises the problem that respondents might manage this burden by unknown filter mechanisms used subconsciously when assimilating the information provided about the scenarios.

On the other hand, giving insufficient descriptions of the scenarios raises the problem that respondents might manage the gaps by unknown extrapolations, also used subconsciously, when assimilating the information provided about the scenarios. It is likely that both assimilation processes are operating simultaneously whenever a respondent is interpreting the presentation of scenarios. It would therefore be expected that their responses would be sensitive to the construction and presentation of the background and scenarios by the analyst. However, any examination of the sensitivity of the results to these sources of bias would be limited by the number of scenario variations that can be examined for any one respondent or in any one study. In contrast, these sources of bias can be more successfully minimised by the trial-based MAUI approach outlined in this Appendix, which separates the scoring of each health state by the fully informed but appropriately blinded patient who is actually experiencing it from the previous generation of the valuation of that health state by members of the general population (thereby avoiding the need for a further analyst to act as an intermediary after the trial).

The post-trial scenario construction process has a number of implications. The scenario-based approach runs the risk of presenting ‘extremes’ of health states for valuation rather than reflecting the distribution. Given the limited number of health states presented for valuation, there is rarely a basis to examine this source of uncertainty in sensitivity analyses. Using a MAUI in the context of a randomised trial (see Part II, Sub-section B.5) avoids this problem. Furthermore, a key implication of analyst bias is the potential for the scenario-based approach to focus on particular symptoms and attributes, which would not necessarily be the way that a person experiencing the health state would perceive it. This leads to a distortion along the lines that ‘nothing seems as important as when you are asked to think about it’.

### **Presenting the methods of generating scenarios and of presenting them to respondents**

If preference weights in utility units have been derived with the use of hypothetical health state scenarios, provide details of the methods used in the utility study as part of the information provided in Part II, Sub-section C.1. Provide data and references that support the validity and reliability of these methods.

Describe the approach taken to construct the scenarios. The scenarios should be developed rigorously, including by demonstrating that consideration has been given to the following:

- Describe the basis of the derivation of the health state scenarios for the survey. Discuss the relationship between these scenarios and the quantified estimates supporting the conclusions presented in Section B of the assessment report or modified in Section C of the assessment report. Given the inherently subjective nature of this process, report any attempt to minimise selection bias in the process and its impact. A more convincing case would be based on a randomised trial that measured health-related quality of life frequently with one or more valid and reliable generic instruments, and the construction of the scenarios is justified and compared with the detailed quality-of-life information from the trial results using these instruments.
- Explain the derivation of the descriptions in each scenario. Discuss the approaches taken to reflect the experience of patients experiencing these health states in the text of the scenarios. For example, describe the derivation of the health state scenarios and weighting and whether they were derived directly using one or more facilitated focus groups (such a group should include Australians — users of the proposed medical

service and people with some experience of the medical condition, as well as medical experts). In particular, explain how the five to nine attributes (see guidance in relation to text below) were selected for inclusion in each scenario from the range of patient experiences. Discuss the need for, and implications of, choosing a proxy (e.g. a carer, a family member or a health care professional) in place of patients for this step.

- Examine whether the description of each scenario was understandable to Australian respondents. For example, report whether initial scenarios developed were piloted using in-depth interviews on all aspects of the respondents' thoughts and comments before undertaking the full survey. If a pilot study was conducted, advise whether it identified any issues and how these were addressed before the scenarios were used in the utility study.
- Report any assessment of the scenarios developed in terms of validity, reliability, responsiveness to change, and clinical importance. Report any assessment of the duration of the period covered in each scenario compared with the duration assumed in the choice-based preference elicitation task (see below).
- Clearly distinguish between elements in the scenarios relating to health and elements not relating to health (such as convenience of use, increased availability of options and any other externality). If non-health elements are included, ensure that elicited preferences can be presented separately as health elements alone or as health elements combined with other elements. The base case should be based on health elements alone. Use sensitivity analyses to examine the impact of including any other elements.

The text used to describe each health state scenario is crucial as the means to convey the basis of the utility weight elicited. Demonstrate that consideration has been given to the following:

- Respondents to scenarios are likely to be subject to cognitive overload when the number of attributes or aspects of the health state increases beyond five to nine.
- Each scenario should adopt the patient's perspective, such that respondents are to imagine that they are in the health state described. The scenarios might be presented in the first or third person.
- Each scenario should be a single static health state rather than a profile of two or more different health states.
- The 'ex post' perspective (in which the health state is as experienced with a full diagnosis without considering the risk of a future event) is preferred in the description of scenarios to ensure that all relevant and important aspects are included explicitly and that all irrelevant aspects are excluded (e.g. the process of diagnosis and a range of possible prognoses). Provide a justification to support the use of an 'ex ante' perspective in any health state scenario. A possible example is the use of a medical service that is intended to prevent a future harmful event.
- As the scenarios are to be presented to individuals with limited technical knowledge, use simple language and a logical sequence of presentation of material to allow all respondents to understand the background and the scenarios. Avoid technical terms and unnecessary words.
- Minimise the possibility of framing and labelling effects in which apparently small changes in wording of the scenario can produce substantial shifts in response. A possible way of doing this is to provide more background context, but because each scenario is essentially a subjective matter, it is difficult to anticipate where problems could arise in any particular context. Report the results of any pilot testing for obvious



framing and labelling effects (e.g. the use of emotive disease labels such as ‘cancer’ or ‘neurological disorder’ in the health state description) in the design and implementation of the scenario. An exception to the above example might be where an ‘ex ante’ perspective is justified.

- To minimise sponsor bias, the supplier should not be named during the survey. To focus on the health state, it would be preferable not to identify the medical service or the nature of the service. A justification should be provided if the service assesses some non-health outcome aspect of therapy.
- Consider including questions to confirm the respondents’ comprehension of the background information and scenarios provided, and report the results of such a validation exercise.
- Justify the number of scenarios to be presented for valuation. The burden on respondents represents an upper limit, which is influenced by the complexity of the information presented and the number of attributes, as well as the number of scenarios. If the number of scenarios to be valued is less than this upper limit, consider including one or more extra scenarios that capture any important variation in the description of one or more health states to be valued. These extra scenarios would enable the presentation of sensitivity analyses of the impact of the description of the scenarios valued for the base case. An important limitation of the scenario-based approach to valuation is that sensitivity analysis of this important source of uncertainty is rarely presented.

Provide a copy of the information provided to the respondents as an attachment to the assessment report. Include in these materials any background information, the text of all health state scenarios, any questions used to confirm comprehension and the questions used to elicit preference weights (‘utilities’). Also provide a copy of any computer program used to facilitate the presentation of information and the elicitation of utility weights.

Outline the methodology adopted in implementing the survey instrument. Demonstrate that consideration has been given to the following:

- Face-to-face interviews are preferred to facilitate comprehension of the background information provided, the description of the scenarios and the questions asked. Provide a justification to support the use of telephone interviews or posted self-administered questionnaires.
- The respondent should be asked questions throughout the background narrative to keep them involved and to ensure understanding.
- Interviewers should be carefully trained to read material at an appropriate pace, and to use conversational inflection, pauses and eye contact in the appropriate manner.
- Material should be provided in a logical sequence and illustrated where appropriate with pictures, graphs or diagrams. Include display items to improve understanding and to increase interest.

Comment on how the study addressed the controversy of whose utility weights are elicited (e.g. a patient, a proxy for the patient, such as a care-giver or a member of the general population) discussed in the background above. The possibly unattainable ideal is that these utility weights are elicited from a representative cross-sectional sample of the Australian general population that is fully informed of all health implications of each health state scenario presented.

If respondents are not from the general population, this approach might also raise potentially important issues to do with adjusting utility weights for groups of patients in certain disease groups (so-called double jeopardy, see Sub-section A4.3 of this Appendix for further explanation). Therefore, for health states reflecting a chronic medical condition, also comment on whether the approach taken reflects adaptation of patients to the experience of the health state, and the implications this has for relating the valuation to the duration of the health state.

### **Elicitation, statistical analysis, reporting of results and interpretation of scenario-based utility valuation of health outcomes**

Anchor the utility weights elicited on a 0–1 ratio scale of death to full (perfect) health. Elicit these weights using a choice-based preference elicitation task, which makes explicit that a choice or trade-off has to be made, and therefore allows for the strength of preference to be revealed. Justify the method chosen and provide details of the method used. The method chosen might be one of the following:

- SG: this method has the more direct theoretical foundation.
- TTO: this is a direct measurement tool designed specifically for use in health care evaluation. It is more appropriate for use by respondents who have difficulty in understanding probabilities. It is particularly useful in studies that compare alternatives in which TTO is the major clinical factor. The utility weight is based on how much quantity of life people are prepared to give up for additional quality of life.
- Each of these scaling techniques is confounded: TTO by time preference and SG by risk attitude. As both SG and TTO relative values are consistent in the direction of expected bias compared to each other, and comparison of the two techniques indicates that they provide similar results; therefore, either can be used as a scaling technique in an assessment report.
- The use of a MAUI to generate utility weights from a scenario is discouraged. This would not be a preference elicitation task, but rather a ‘mapping’ from one scenario to another MAUI-based scenario. If the scenario captures only a few domains covered by the MAUI, the respondent is forced to guess from the information provided what response should be given for the other domains covered by the MAUI. On the other hand, if the scenario is constructed to capture all domains, the analyst’s control of the scenario descriptions is so influential that the descriptive words chosen can tend to lead the respondent towards particular responses in each domain. In an extreme case, the analyst could effectively nominate the utility weight yielded by this approach based on their own expert opinion, and then align the text of the scenario descriptions to the text of the MAUI questions.
- Other methods for eliciting preferences, such as discrete choice experiments or other conjoint analysis methods, are still in development and thus any guidance here is preliminary. There are five main stages that characterise these types of study:
  - *Determine the attributes*: if based on one or more submitted randomised trials, the attributes should reflect the different components of the trial arms. If they are not defined on this basis, then literature reviews, patient group discussions and individual patient interviews will need to be used to solicit the attributes. These attributes should be important to the patients. If cost is used as an attribute, the technique can generate willingness-to-pay (WTP) under certain circumstances (see Sub-section A5.2 of Appendix5). To ensure that the analysis is being used to value health states rather than to value the treatments, it is important to exclude any other description or process aspect of the treatment.

- Define the characteristic levels: justify the use of cardinal, ordinal or categorical scales. The levels should be realistic, be capable of being traded off and capture all relevant outcomes.
- Choose the scenarios to be presented in the stated preference experiment: justify the presentation of the scenarios to ensure that they are realistic (e.g. ensure that the defined period of time for each scenario is consistent for both the proposed medical service and the main comparator) and that they make sense to the respondent (see guidance on constructing the scenarios in this Section of this Appendix). The number of scenarios will increase with the number of attributes and attribute levels, and it is generally not feasible to present all combinations of scenarios in a questionnaire. Use an appropriate experimental design, typically a fractional factorial design based on orthogonality, to choose the subset of scenarios to be presented in the experiment. Describe and justify the basis for generating the experimental design, including details of any software used. Provide the full experimental design in an attachment to the assessment report, including a list of all scenarios developed.
- Establish preferences using discrete choices: present each respondent with a series of pairs or groups of options (choice sets) among the scenarios and request that a selection be made defining which is the most preferred. Ranking and rating exercises have been used in conjoint analysis; however, the use of discrete choice experiments is preferred, because they are more consistent with the choice-based nature of SG and TTO, and have a more established basis in economic theory and statistical analysis.
- Analyse data: analyse the responses from the scenarios using regression techniques. Typically, a multinomial logit analysis is used because the dependent variable is a discrete random variable. Justify the modelling approach, including consideration of treatment of repeated observations and heterogeneity (e.g. use of mixed logit). Report on the extent to which the model explains the variation in preference selection. Explore the impact of possible confounding factors.

Claimed advantages of conjoint analysis include the ability to describe health state changes in terms of comparisons across the attributes, the duration of these changes and the probability of these changes occurring. Although the techniques of conjoint analysis are developing, they are still not yet sufficiently acceptable to have direct influence on MSAC decision making on their own. They are claimed to also explicitly consider non-health elements (in which case, results should be presented with and without including those elements). However, it is not clear that there is an acceptable framework outside the QALY framework in which to consider these claimed advantages in a comparable way across assessment reports.

Ensure that the sample size is large enough to measure population variance. The power of the study should be tested and between-group correlations should be demonstrated.

Present the results of the utility study as part of the information provided in response to Sub-section C2. Report the results as the point estimate of the mean utility of each health state scenario with its 95% confidence interval. In discussing these results, provide an overall assessment of the approach adopted to elicit preference weights from the hypothetical scenarios. Particularly, consider whether the methods by which the health state scenarios were:

- constructed to allow all the critical changes in quality of life associated with the intervention to be captured and presented in such a way that they are accurately perceived by the respondents; and
- derived and constructed to likely lead to bias in the valuation of health-related quality of life associated with the medical service; for example, by focusing on some aspects of health-related quality of life (for example physical functioning) while excluding or minimising the impact of others (such as mental or social health).

From these results presented in Sub-section C2, identify and justify the estimates to be used as variables in the economic evaluation presented in Sub-section D5 for the base case and Sub-section D6 for the sensitivity analyses.

## **A4.5 Other methods for obtaining utilities**

The following methods have all been presented in assessment reports for MSAC. Each raises a series of concerns, as detailed below.

### **Mapping of generic and disease-specific scales**

In contrast with MAUIs, although other generic and disease-specific scales might be based on sophisticated psychometric techniques for instrument construction, none of those scales is capable of representing individual preferences on a scale of 0 = death and 1 = full (perfect) health, and so none can be used to calculate QALYs without some transformation. Despite this, a number of attempts have been made to ‘map’ from scores reported in randomised trials using generic or disease-specific quality-of-life measures into utility weights, which are then used to construct QALYs. Approaches vary from a simple intuitive mapping to the use of statistical techniques. For example, responses on a visual analogue scale of 0 to 100 to the question asking respondents to rate their health today have been divided by 100 and (wrongly) claimed to therefore measure utility weights on a 0 to 1 scale. Another example is the use of regression to ‘map’ an association between two sets of responses from a survey of respondents, each completing both the quality-of-life instrument and a MAUI, or other acceptable technique of eliciting preference weights. This regression ‘map’ is then used to transform into ‘utilities’ the responses to the quality-of-life instrument reported by respondents in another trial.

These are not well-established procedures. Where statistical techniques have been used, tests of reliability might include the predictive value of the technique across a range of quality-of-life values and changes in quality of life within, and differences between, respondents with the relevant medical condition. Where this approach is adopted, extensive sensitive analysis around the estimates generated should be undertaken to examine the sensitivity of results of the economic evaluation to this variable. Where such ‘mapping’ is presented, special attention needs to be given to establishing that the results generated are plausible and unbiased, particularly where the preference weight estimates generated have a substantial impact on the results of the economic evaluation.

It is difficult to illustrate the assessment of plausibility and bias in these circumstances. An approach that does not ‘map’ to an adequate utility instrument (i.e. that satisfies characteristics (b) and (c) of the QALYs shown in Section A4.2 in this Appendix) would not meet an essential prerequisite in estimating a preference weight index. An approach that is not based on a study that concomitantly measured the quality-of-life measure and such an index would also not meet an essential prerequisite to generate an association. Other issues to assess include the difficulties of ‘mapping’ ordinal (ranking) scales to the cardinal utility scale, the presence of floor and ceiling effects in most quality-of-life measures, and whether an acceptable range of important dimensions are adequately captured (the latter two have been assessed as acceptable for the MAUIs recommended in Sub-section B5). A more structured approach might be taken to map specific dimensions of a generic quality-of-life instrument to corresponding dimensions of a MAUI (possibly best exemplified by the mapping of the SF-36 to the SF-6D), but this involves a much greater amount of developmental research work.

### **Population-matching studies**

Another alternative occasionally used involves recruiting a separate sample of patients with characteristics similar to those in the randomised trials and for whom listing is requested. These matched patients then complete a MAUI reflecting their current health state (as a surrogate for a trial participant directly completing the MAUI), which is then used to estimate utility weights for the economic evaluation.

This population-matching approach is also subject to multiple sources of bias and thus uncertainty, particularly related to how similar the sampled patients are to those in the economic evaluation and the inability to blind the sampled patients from the objectives of the study. This can be context specific; for example, if there are important side effects, it might be particularly important to ensure that the sampled patients are exposed to the medical service and its side effects at the time the MAUI is completed.

This approach might be strengthened by getting the sampled patients to complete another quality-of-life instrument that was completed in the trials, and using the results of this concurrent instrument to more closely match a subset of sampled patients with trial participants and with the population for whom listing is requested. It can also be used to develop sample-based statistics of variance around the utility weights, which can be used in the sensitivity analysis of the economic evaluation.

### **Preference weights (‘utilities’) sourced from the literature**

‘Off-the-shelf’ utility estimates might be available from the literature, and have been most often used when seeking to examine the impact of quality-adjusting a survival claim estimated in terms of life-years gained. As for any presentation of secondary (or even tertiary) data or analysis, the validity of the utility estimate depends on the methods used to elicit the estimate. Accordingly, present and assess the results against the preferred characteristics of a primary utility study, including:

- how the studies were identified (e.g. systematic search preferred to selective reporting);
- how representative the health state in each identified study is of the health state in the presented economic evaluation (including in dimensions of the type and severity of symptoms, and the duration of the health state);
- how the health state was captured (e.g. MAUI versus scenario based);

- how the preference was elicited (e.g. SG or TTO);
- what sample was chosen to respond to the MAUI questionnaire or scenario (e.g. members of the general public, patients, care givers, health care professionals);
- what assessment was made of the nature and direction of bias that might arise given the sample and methods; and
- how the sensitivity analyses examined variation in the identified utility options.

A particular difficulty in interpretation has occurred when a cost-utility analysis relies on combining utility weights across different sources for different health states within an economic evaluation, particularly across different sources that used different methods.

## Appendix 5     Monetary valuation of health outcomes

---

### A5.1 Preference for cost-utility analyses over cost-benefit analyses

Cost-benefit analyses are not preferred by MSAC because they are not likely to be helpful to most MSAC deliberations. The reasons for this are as follows:

- Cost-benefit analyses are typically applied in the context of a fixed-decision rule, which does not incorporate the breadth of equity and ethical considerations that are relevant to MSAC decision making (see also Appendix 1).
- The use of willingness to pay (WTP) to elicit monetary valuation for a cost-benefit analysis, which will be influenced by an individual's income and assets, is inconsistent with the principles of MSAC as a subsidy program to ensure equity of access.
- There remain considerable problems with interpreting WTP responses in the context of the Australian health care system where individuals do not typically face market prices. It could be argued further that the MBS, which uses fixed levels of co-payment and safety nets to achieve its objective in minimising low income as a barrier to accessing medical services that are MBS-funded in Australia, removes price signals even more than other elements in the Australian health care system.
- The methods for deriving monetary valuations of health gains presented to date have not satisfactorily minimised the hypothetical nature of the responses elicited or the incentives for the respondents to provide values that reflect a desire to have the MBS subsidy proceed in the full knowledge that the respondent will not directly incur this cost. Although it is theoretically possible to improve the realism of the scenarios and of the questions asked to elicit plausible monetary values (see Sub-section A5.2 of this Appendix), there remains a residual uncertainty in aligning the provision of resources valued in monetary units with welfare outcomes, which are apparently valued in the same monetary units.
- Cost-benefit analyses typically assign preference weights including to other welfare changes beyond the primary focus of MSAC on health outcomes (these include production changes and process changes), which have tended to reflect the construction of the scenario or attribute used to elicit the monetary valuation rather than to reflect the weights assigned by MSAC when considering a fuller range of other relevant factors, particularly equity.
- For the above reasons, there is unlikely to be a consistent exchange rate between monetary valuation and the utility weight that is the preferred basis for assessing strength of preference (see Sub-section A4.1 of Appendix 4). Therefore, considering these two approaches to valuing outcomes in parallel would predictably result in inconsistent decisions across assessment reports. This is undesirable.

- Although it is possible to use utility-based instruments in randomised trials to estimate the strength of preference for different health outcomes (see Sub-section B5), this is not yet practical for monetary-based instruments. Therefore, the advantages outlined in Sub-section A4.3 of Appendix 4 for trial-based utility weights cannot be generated for monetary valuation. There are therefore disadvantages in common between scenario-based utility valuation (see Sub-section A4.4 of Appendix 4) and scenario-based monetary valuation (see Sub-section A5.2 of this Appendix).

Given the above reasoning, monetary valuation of health outcomes is allowed but is considered to be supplementary to utility valuation. Therefore, if both a cost-utility analysis and a cost-benefit analysis are presented in an assessment report, discuss the differences in the results and any differences in conclusions. In the absence of a cost-utility analysis, discuss why only a cost-benefit analysis is thought to be informative and why a cost-utility analysis is not possible. For example, consideration of such analyses might be justified in some situations to provide informative insights to the perception of the respondents to the clinical performance of a proposed medical service; however, such analyses should be interpreted cautiously in the absence of a worthwhile gain in health outcomes. Further guidance is provided in Sub-section A5.2 of this Appendix.

## **A5.2 Scenario-based monetary valuation of health outcomes**

### **Background**

Monetary valuation of health outcomes is typically scenario based. The issues raised in Sub-section A4.4 of Appendix 4 regarding the use of scenarios as a basis for eliciting the strength of preference in a utility metric largely overlap with their use as a basis for eliciting the strength of preference in a monetary metric. It is conceivable that monetary valuation could be elicited in the context of a randomised double-blind trial, but the practicalities of addressing the issues raised below suggest that this will not occur in the near future.

This Appendix seeks to identify those areas where monetary valuation might be informative in situations where utility valuation is problematic. Situations identified to date have tended to arise due to concerns over the lack of sensitivity of utility valuation to perceived increments in health outcomes. These have included short-term changes in health outcomes, differences in health outcomes that are too small to be detected with utility-based instruments, and differences in adverse outcomes for two medical services that are otherwise similar in terms of comparative effectiveness. An alternative metric might be justified in these circumstances, because underlying the quality-adjusted life-year (QALY) approach is the fact that survival duration is the metric, and there might be health gains that are valued, but are not sufficient for individuals to trade-off survival. However, this reduces comparability across assessment reports, because it introduces a new valuation system that is not necessarily interpreted the same way in the valuation step by the respondent as utility valuation. It also brings in other aspects, whether implicit or not, beyond valuing health outcomes.

An assessment report seeking to supplement a utility valuation of health outcomes with a monetary valuation of health outcomes should provide a justification for doing so. Alongside this justification for providing these supplementary estimates, present both sets of methods and results, and comment on the interpretation of the results compared with each other. As with the interpretation of the results of any measure of health outcomes, any claim for an improved sensitivity in quantifying the utility weight of smaller advantages needs to be assessed against the possibility of a true negative (i.e. that the proposed medical service has no overall perceptible incremental effect on strength of



preference; see also Sub-sections A4.3 and A4.4 of Appendix 4). Document the evidence that supports any claim that any difference in results between utility-based valuation and monetary-based valuation is attributable to the special characteristic of the health state, and not some idiosyncrasy in the utility measurement procedures that have been adopted. This would help justify any apparent diminution in comparability across assessment reports that provide utility weights.

An assessment report that provides monetary valuations of health outcomes without corresponding utility valuations would be more difficult to assess in terms of comparability across assessment reports.

Consistent with the request in Section D and Sub-section A4.4 of Appendix 4, an assessment report that seeks to provide a monetary valuation of any attribute other than health outcomes (e.g. a production change; see Appendix 6) should do so separately from the valuation of health outcomes. This can be done by providing a supplementary economic evaluation that adds the additional information to the base-case economic evaluation. A request in an assessment report for MSAC to consider a non-health outcome or process attribute (such as convenience of use, increased availability of options and any other externality) would need to be judged on its merits, which would be informed by the direction and extent of the impact of its inclusion on the base-case economic evaluation. This distinction is therefore important both to promote consistency of decision making based primarily on health outcomes and to allow flexibility to consider other factors that MSAC might accept as relevant.

### **Presenting the methods of generating scenarios and presenting them to respondents**

If preference weights in monetary units have been derived with the use of hypothetical health state scenarios, provide details of the methods used in the study as part of the information provided in Sub-section C1. Provide data and references that support the validity and reliability of these methods. Refer to the text under the corresponding subheading of Sub-section A4.4 of Appendix 4 to identify the information to be provided, including a clear description of the attributes that are compared between the proposed medical service and its main comparator. Additional information specific to monetary valuations includes the following:

- Describe the attributes in each scenario in a way that matches the policy question and the underlying theoretical construct to be addressed in the contingent market.
- Whenever a probability of any type is included for an attribute in a scenario, examine more than one level of probability when eliciting monetary values in order to assess the degree of understanding (e.g. that a greater probability of benefit yields a greater monetary value of WTP).
- Where scenarios are developed as changes in health states rather than as the health states themselves, describe the likelihood, extent and duration of each change.

### **Elicitation, statistical analysis, reporting of results and interpretation of scenario-based monetary valuation of health outcomes**

The most commonly used method is contingent valuation (CV) to elicit WTP. If a CV study is included in an assessment report, provide a justification for its inclusion, including why it would be informative for MSAC decision making.

The assessment report should outline the methodology adopted in designing and implementing the CV survey instrument. Demonstrate that consideration has been given to the following:

- The contingent (hypothetical) market should be a simple out-of-pocket payment to elicit the individual's strength of preference by considering the question of spending their private income to estimate the value of the change in health states being presented. Ensure that respondents understand the nature of the payment vehicle and that their responses are interpreted appropriately. The average WTP across respondents from this valuation might not necessarily be the WTP that society overall has for subsidising medical services to improve health outcomes for the population as a whole, but it is not clear that changing the hypothetical market to reflect a societal question of funding a public subsidy program would be meaningful to respondents. This market should also be described in simple language, eliminating unnecessary words and avoiding technical jargon.
- The initial WTP elicitation instrument describing the contingent market should be piloted alongside the piloting of the background information and the scenarios. Report any issues arising and how they were addressed before the full study began.
- Discuss the choice between a discrete choice format or an open-ended questionnaire format (with prompts or a payment card) to elicit responses. The closed-bid discrete choice format with randomly selected bids presented to each respondent — and only one bid per respondent — is more theoretically valid and less subject to bias than the other methods. Other issues to consider include the sample size required for the statistical analysis to infer the mean WTP from discrete choices, and the increased likelihood of nonresponse or protest response from open-ended questions. Justify the range of values used in the discrete choices or the prompts or payment cards. When conducting the survey, randomly allocate the selection of the order of discrete choices across respondents or the selection from the range of values in prompts and cards.
- To ensure some consistency within the time frames across different WTP studies, frame the questions in one of two ways:
  - as a one-off payment but constrained to within any one year, by invoking each respondent's annual (rather than lifetime) income; or
  - as a regular annual payment, with the value derived for 'this year' only, not for a 'hypothetical' year.
- Remind respondents of their budget constraints for their WTP throughout the survey.
- When conducting the survey, adopt a random ordering of questions across respondents.
- WTP studies should be conducted in a comparative sense and respondents should be made aware of any close substitutes. This would help to make clear the extent of incremental improvement in health across the alternatives.
- WTP is expected to be correlated to ability to pay. Indicate whether ability to pay has been assessed according to personal or household income (and, if the latter, whether this is adjusted for household size) and whether it has been assessed according to current income or also reflects assets that could be realised to make payments. Socio-demographic characteristics of respondents should be collected and included in the analysis.

From the above information, indicate the steps that have been taken to minimise the following sources of bias in the WTP survey:

- *hypothetical bias*: the respondent responds to a perception that the survey is hypothetical with hypothetical and therefore meaningless answers;
- *strategic bias*: the respondent varies the WTP from the ‘true’ WTP to increase the chances of getting a preferred decision by influencing the decision maker;
- interviewer bias: face-to-face or telephone interviews run the risk that valuation will be influenced (purposefully or accidentally) by the interviewer;
- *starting-point bias*: the initial prompt or bid in the bidding approach will anchor the respondent towards the starting bid, narrowing the distribution around the mean (portraying greater consensus than truly exists) and causing a loss in efficiency;
- ‘*yea-saying*’ bias: the respondent will agree with amounts as offered by interviewer;
- *range bias*: the elicitation procedure presents a range of potential WTP amounts that influences the WTP amount given by respondents; and
- *sponsor bias*: knowledge of the identity of the sponsor affects responses; minimised by not naming the sponsor of the survey or the manufacturer of the medical product.

The validity of the WTP depends on minimising sources of bias to reveal the true strength of preference in monetary terms.

Some preliminary guidance in relation to other stated preference methods, such as discrete choice experiments and conjoint analysis, is presented under the corresponding subheading in Sub-section A4.4 of Appendix 4. The methodological guidance on those methods should be considered in addition to the general guidance given above in this Section for valuing discrete health states. In addition, discrete choice experiments might also be used to calculate monetary measures of the composite of incremental health outcomes from the proposed medical service as a comparison of the alternative profiles of health outcomes over defined periods of time resulting from the proposed medical service and the main comparator. If so, justify the presentation of these profiles of health states to ensure that they realistically and accurately reflect the choice context (e.g. allowing for a ‘status quo’ or an ‘opt out’ option where appropriate for the presentation of the alternative profiles in each choice set) and that they make sense to the respondent (see general guidance on constructing the scenarios).

### **The statistical analysis, interpretation and reporting of data**

Present the results of the scenario-based monetary valuation study as part of the information provided in response to Sub-section C2. Report mean WTP values on a net present value basis for each health state and then the overall aggregate with their 95% confidence intervals, interquartile range and full range.

Assess the results of the WTP survey as follows:

- Present WTP values without adjustment for income. Also report WTP disaggregated across income group. Where the mean ability to pay in the survey differs from the national average, comment on the interpretation of the results.
- Present the results both in an unadjusted fashion and with outliers removed. Discuss any difference in these results.

- Report the response rate. Comment on the implications of the response rate and other potential sources of selection bias for the interpretability of the results of the survey.
- Report the proportions of zero and very high bids. If either or both of these are greater than 10%, discuss the possible reasons for these proportions and their implications. Ask respondents to explain their reasons for responding with a zero bid.
- Conduct regression analyses to assess the factors that might explain the WTP values given. Variables to examine include an 'interviewer' variable, a 'question order' variable, a 'prompt' variable (of the range of starting values in the prompt) and an 'income' variable.
- Assess whether the results make economic sense (i.e. that WTP increases with the size of both health gains increases and ability to pay increases).

WTP values are context specific, so values should only be used and applied to the specific circumstances for which they were obtained. WTP values are interpreted as an upper limit to true valuation. From these results presented in Sub-section C2, identify and justify the estimates to be used as variables in the economic evaluation presented in Sub-section D5 for the base case and Sub-section D6 for the sensitivity analyses.

## **Appendix 6 Including non-health care resources and non-health outcomes in a supplementary analysis**

---

This Appendix provides additional guidance on the preparation of supplementary analyses of an economic evaluation to incorporate changes in non-health care resources and/or non-health outcomes that would be attributable to the listing of the proposed medical service (see Sub-section D1).

### **A6.1 Identifying, measuring and valuing non-health care resources**

Occasionally, because of the medical condition under treatment or the age of the patients, consideration of direct non-health care costs such as social services (home help, day care, meals on wheels, private travel to access health care, etc) might be relevant.

If incorporation of non-health care resources is relevant for a supplementary analysis, adapt the general principles as detailed in Sub-section D4 for health care resources to generate and present these variables. In brief, the resources should be identified and defined. An appropriate unit of measurement should be identified and the extent of change in the provision of the resources should be estimated. Present and justify an appropriate unit cost to estimate the value of the resources.

### **A6.2 Identifying, measuring and valuing non-health outcomes**

Occasionally, listing a proposed medical service might generate worthwhile impacts that are not captured as health outcomes, such as the value of information to the patient generated by an additional diagnostic test that does not change management of a medical condition.

If incorporation of changes in non-health outcomes (including economic outcomes) is relevant for a supplementary analysis, adapt the general principles outlined in Sub-section D4 for health outcomes, including reference to Sub-section A5.2 of Appendix 5, as appropriate. In brief, the outcome should be identified and defined. An appropriate unit of measurement should be identified and the extent of change in the outcome should be estimated. Present and justify an appropriate valuation of the outcome.

#### **Production changes**

A production change is the value estimated in monetary units of the potential working time gained or lost measured in time units (days, weeks, years, etc.), which is realised as productive activity. It might also include realising the productive change of the potential impaired working time gained or lost by a sick patient who continues to work (measured in similar time units together with a measure of any associated change in the extent of impairment). Production changes have been called indirect economic outcomes in recognition of the fact that subsequent decisions had to be made to realise the time gained as productive activity to the advantage of the rest of society rather than as any other activity.

Provide a strong justification if production changes are combined with surrogate outcome indicators in an economic evaluation because this combination is generally inappropriate.

If production changes are to be included in a cost-utility analysis, adopt a method that avoids double-counting the estimates of health-related quality-of-life changes. The utility weights in this analysis already capture these health-related changes because they incorporate the utility impacts of productive capacity to the individual receiving the proposed medical service. These health-related changes are therefore already appropriately included in the denominator of the cost-utility ratio.

Unlike direct health benefits, the economic benefit to society through patients' return to, or maintenance of, productive capacity is both difficult and controversial to estimate accurately. This is because the available methods and their application remain unresolved. Therefore, although changes in production as an outcome of therapy might be included in supplementary analyses in assessment reports for MSAC, they should not be included in the base-case analysis.

There are several difficulties in estimating the net present value of production changes. These estimates are underpinned by three assumptions:

- for short-term absence, production will be made up on the return to work;
- employers usually have excess capacity in the labour force to cover absenteeism; and
- for long-term absence, production will be made up by a replacement worker otherwise unemployed.

Where estimation of production changes can be justified in the assessment report, address each of the three underlying assumptions listed above when estimating production changes from the potential working time gained or lost (reported in time units). For example, the claim that there has been a recovery of production lost due to returning to health from an episode of illness depends on demonstrating that:

- the worker returns to work;
- the worker is productive;
- the production lost is not made up elsewhere by others in the company or the same worker following return to work (note: if the worker is highly productive, the incentives to replace that worker are stronger); and
- no temporary replacement from outside has been employed (namely, that there is full employment).

As in this example, the **marginal** increase in society's production due to the return of healthy workers to the workplace is overestimated if the human capital method is used; that is, the workers' time regained is simply multiplied by the labour market value of the average worker (usually estimated by the average wage). It is not always likely to be zero either, but some proportion in between. Provide and justify the best estimate of the true proportion based on firm evidence.

Addressing the four questions in the example above would therefore help to convert the potential working time gained or lost reported in time units into production gains or losses reported in monetary units. The friction method has been advocated as a method that provides a basis to help make this type of conversion. Although there is no evidence that it has yet been applied in Australia, it is theoretically preferable to the human capital

method for this reason. However, in the example provided above, it only offers a basis for addressing the last two of the four questions, and only does so by proposing an indirect estimate at the national level rather than a direct estimate at the patient level. The friction method therefore still generates an upper estimate compared with an approach that could address all four of the questions above. Other evidence needs to be provided to address the first two questions, because not all healthy workers would choose to deploy the time gain to return to contributing to societal production. In the example above, recognising that this choice exists is important because deploying the time gain for some other purpose, such as a leisure activity, is an intrinsic part of valuing the improved health as a gain in utility weights rather than valuing it as a production gain to society in monetary terms.

Any evidence to support an estimate of the proportion of people who choose to return to contributing to societal production would also need to account for the influence of incentives provided through various types of sickness benefit payments provided by social security systems and employers, which vary across countries. This might hinder the translation of overseas evidence to Australia.

Answering all four questions satisfactorily in the example above would therefore help minimise double-counting across the denominator and the numerator of an incremental cost-utility ratio, because it would more accurately estimate the extent of production gains to society beyond the gains valued by the population benefiting with improved health. Valued in monetary terms, these production gains would represent a more suitable estimate for inclusion in the numerator of this ratio.

The above example is intended to illustrate the application of the three more general reasons. A similar approach would be needed in other contexts, such as a medical service that prevents future episodes of illness, or a medical service that might improve production capacity in individuals who, without the proposed medical service, would otherwise stay at work, although unwell, and therefore perform at less than full production capacity.

Present the results of the economic evaluation excluding the production changes in the base case. Assess the impact of including these changes in a supplementary analysis. This separation allows MSAC to consider the impact of their inclusion on the direction and extent of change on the base case.

At the same time, MSAC can weigh up, as another relevant factor, the inevitable equity implications of varying the base case to include an element that explicitly favours those who make a greater contribution to production. Inclusion of production gains favours those medical services that improve the health of people who are able and choose to return to contributing to societal production.

The present value of production changes should be calculated. This means that where production gains are anticipated over a number of time periods (beyond one year) these should also be discounted. Discounting future costs and benefits is a standard feature of economic evaluation. Costs or benefits are discounted at an annual rate of 5%.

### **A6.3 Resources and outcomes to be excluded**

Costs should be limited to those associated with the medical condition under treatment. In other words, do not include as consequences in the economic evaluation other unrelated medical conditions that, in the fullness of time, are likely to afflict patients who live longer as a result of effective treatment that they receive now.



## Appendix 7 Measures taken by the investigators to minimise bias in non-randomised studies

---

This Appendix is relevant to Part III, Section B(ii). It is designed as a useful guide to help MSAC and the applicant review the scientific rigour of the evidence by assessing the measures taken by the investigators to minimise bias. It is not intended to discourage the presentation of data.

Categorise studies into the study types defined below. Then, for each methodological topic listed for the relevant study type, choose the description that best fits each study. If the submission includes a number of studies of the same type, tabulate the responses. In each case, the methodological descriptions are arranged in a descending order of quality (i.e. 1 being the worst).

As for the assessment of randomised trials in Part II, Section B and Part III, Section B(i), the purpose of these assessments is to provide the sponsor and MSAC with a clear idea of which studies are of greater scientific rigour. There is no minimum standard, but MSAC is most likely to be persuaded by the data of the highest scientific rigour. Submissions should therefore be particularly careful to justify using the results of studies with less scientific rigour in an economic evaluation in place of trials with greater scientific rigour.

There may be other aspects of particular non-randomised studies that might affect the results of such studies and their comparability with different studies of the same type. If these aspects are likely to be important, they should also be identified.

**Note:** In each case, if there is insufficient information available to classify the study, assign it to category 1.

### A7.1 Classical observational designs

#### Controlled cohort studies

In this study type, assignment of the groups of individuals to treatment is not random. However, individuals receiving the proposed medical service are followed forward in time from their first exposure and control individuals are followed forward in time from their enrolment in the study. Cohort studies can be concurrent or historical. In the former, the study is planned and conducted prospectively. In the latter, existing records are used to define treatment status and determine the outcomes.

#### Possibility of confounding

It is important that there are no substantial differences at the baseline between treated and control participants in respect of factors that could influence the outcome(s) being studied. Identify which of the following best describes the differences in baseline factors:

- 1) There were significant differences in baseline factors between treated and control participants that have been shown to influence the study outcome(s), and these were not adjusted for in the main analysis.

- 2) There were significant differences in baseline factors between treated and control participants that might have influenced the study outcome(s), and these were not adjusted for in the main analysis.
- 3) There were no differences in baseline factors between treated and control participants that might have influenced the study outcome(s), or any differences were adjusted for in the main analysis.

### **Adequacy of follow-up**

It is important that an attempt is made to summarise the study outcomes for all participants who were included in the study. Identify which of the following best describes the adequacy of follow-up in the study:

- 1) There were significant numbers of drop-outs with no assessment of study outcome(s) in the participants who dropped out, and drop-out rates differed between treated and control groups.
- 2) There were some drop-outs with no assessment of study outcome(s) in the participants who dropped out, and drop-out rates were (approximately) equivalent in treated and control groups.
- 3) Study outcome(s) were assessed in all or nearly all treated and control participants.

### **Blinding of outcomes assessment**

It is important that the observer responsible for measuring the study outcome is unaware of whether the participant belongs to the treated or control group. Identify which of the following best describes the blinding of outcomes assessment:

- 1) There was no attempt to blind the observer(s) to the treatment or control status of the study participants, or any attempt made was inadequate to keep the observer(s) fully blind to the treatment or control status of the study participants.
- 2) The observer(s) were kept fully blinded to the treatment or control status of the study participants.

### **Case-control studies**

In this study type, participants are defined by the presence (cases) or absence (controls) of the study outcome, and their previous use of the proposed medical service is compared.

### **Selection of cases**

It is most important that cases are selected independently of their treatment status. Identify which of the following best describes the selection of cases:

- 1) The process of referral and selection of cases was likely to have been influenced by the participants' previous use of the proposed medical service, and knowledge of the association between use of the proposed medical service and study outcome.
- 2) The process of referral or selection of cases was not influenced by the participants' previous use of the proposed medical service or knowledge of the association between use of the proposed medical service and study outcome.

### **Selection of controls**

The purpose of the control group is to provide an estimate of the odds of exposure in participants who are free from the disease in question in the source population. Identify which of the following best describes the selection of controls:

- 1) The controls were not drawn from the same source population as the cases.
- 2) The controls were drawn from the same source population as the cases (community controls).

### **Possibility of confounding**

It is important that there are no substantial differences between cases and controls in respect of factors that could influence the outcome being studied, other than the risk of exposure to the proposed medical service. Identify which of the following best describes the comparability of cases and controls:

- 1) There were significant differences in factors between cases and controls that have been shown to influence the study outcome, and these were not adjusted for in the main analysis.
- 2) There were differences in factors between cases and controls that might have influenced the study outcome, and these were not adjusted for in the main analysis.
- 3) There were no differences in factors between cases and controls that might have influenced the study outcome, or any differences were adjusted for in the main analysis.

### **Possibility of measurement bias**

It is important that assessment of treatment status (or exposure) is made in an unbiased way. Identify which of the following best describes the assessment of treatment status:

- 1) The measurement of previous use of the proposed medical service (or exposure) was made using an unstructured interview or questionnaire by an observer who was aware of the case or control status of the participants.
- 2) The measurement of previous use of the proposed medical service or exposure was made using a structured interview or questionnaire by an observer who was aware of the case or control status of the participants.
- 3) The measurement of previous use of the proposed medical service (or exposure) was made using a structured interview or questionnaire by an observer who was unaware of the case or control status of the participants or the definition of exposure preceded the outcome.

## **A7.2 Quasi-experimental designs**

### **‘Before and after’ studies**

In this type of study, participants are observed before and after using a medical service. It is really only possible to use this design if the manifestations of the illness being treated are both chronic and reversible. Typically, this will be an opportunistic study, rather than planned. In addition to the sources of bias that affect the previously mentioned observational designs, this study type has particular problems related to time (or order) effects, resulting from the participants being observed over a period, and the lack of a contemporaneous control group. There may be changes in disease severity, symptomatology or resource use that occur independently of any treatment, and it is

impossible to assess these properly without a contemporaneous control group. It is highly likely that participants would be switched to the new therapy because they have not been doing well on the old therapy, and thus their symptoms would tend to be most severe at the time of switching. Regression to the mean will make the new therapy seem better than the old one, in terms of both apparent treatment responses and resource provision.

### **Selection of participants**

- 1) The participants were selected retrospectively from case notes, and the investigators were probably aware of the responses to the old treatment at the time of selection.
- 2) The study was planned, prospective data collection was undertaken in both study periods, and selection of the participants was made without knowledge of the treatment responses.

### **Possibility of confounding**

- 1) There were within-participant differences in factors between the two study periods that were likely to influence the study outcome(s), and these were not adjusted for in the main analysis.
- 2) There were no within-participant differences in factors between the two study periods that were likely to influence the study outcome(s), or any differences were adjusted for in the main analysis.

### **Adequacy of follow-up**

- 1) Drop-out rates differed between the 'before' and 'after' study periods, with no assessment of study outcome(s) in the participants who dropped out.
- 2) There were no drop-outs in either study period (this implies prospective data collection in both periods), or study outcome(s) were assessed in all participants who were commenced on treatment.

### **Blinding of outcomes assessment**

- 1) The observer(s) responsible for outcome assessment was aware of which treatment the study participants had been receiving.
- 2) The observer(s) responsible for outcome assessment was kept fully blinded to the treatment being received by the study participants.

### **Case-series with historical controls**

Typically, this type of study is carried out by a clinical department that has introduced a new management procedure and wishes to compare the results with those of patients treated previously in the department using the old management procedure. Therefore, this type of study shares the same problems of order effects as 'before and after' studies but does not involve the same individuals in both arms.

### **Selection of participants**

- 1) The participants were selected retrospectively from case notes, and the investigators were probably aware of the responses to the old treatment at the time of selection.
- 2) The study was planned, prospective data collection was undertaken in both study periods, and selection of the participants was made without knowledge of the treatment responses.

### **Possibility of confounding**

- 1) There were differences in factors between participants in the two study periods that were likely to influence the study outcome(s), and these were not adjusted for in the main analysis.
- 2) There were no differences in factors between participants in the two study periods that were likely to influence the study outcome(s), or any differences were adjusted for in the main analysis.

### **Adequacy of follow-up**

- 1) Drop-out rates differed between the two study periods, with no assessment of study outcome(s) in the participants who dropped out.
- 2) There were no drop-outs in either study period, or study outcome(s) were assessed in all participants who began the treatment.

### **Blinding of outcomes assessment**

- 1) The observer(s) responsible for outcome assessment were aware of which treatment the study participants had been receiving.
- 2) The observer(s) responsible for outcome assessment were kept fully blinded to the treatment being received by the study participants.

### **Comparison of the results of two or more single-arm studies**

In addition to all the problems noted earlier with ‘before and after’ studies or case-series with historical controls, this approach has the added disadvantage that the outcome assessments were made by different investigators in different settings. It is not possible to compare the results of such studies with any confidence. Assess comparisons involving single arms extracted from randomised trials (when compared without a common reference) as comparisons of the results of two or more single-arm studies.

### **Selection of participants**

- 1) In the studies for either or both alternatives, the participants were selected retrospectively from case notes, and the investigators were probably aware of the responses to the old treatment at the time of selection.
- 2) The studies for both alternatives were planned, prospective data collection was undertaken for all consecutive patients in the study period, and selection of the participants was made without knowledge of the treatment responses.

### **Possibility of confounding**

- 1) There were differences in factors between participants in the study populations for the two alternatives that were likely to influence the study outcome(s), and these were not adjusted for in the main analysis.
- 2) There were no differences in factors between participants in the study populations for the two alternatives that were likely to influence the study outcome(s), or any differences were adjusted for in the main analysis.

### **Adequacy of follow-up**

- 1) Drop-out rates differed between the studies for the two alternatives, with no assessment of study outcome(s) in the participants who dropped out.
- 2) There were no drop-outs in the studies for either alternative, or study outcome(s) were assessed in all participants who were commenced on treatment.

### **Blinding of outcomes assessment**

- 1) In the studies for one or both of the alternatives, the observer(s) responsible for outcome assessment were aware of which treatment the study participants had been receiving.
- 2) In the studies for both alternatives, the observer(s) responsible for outcome assessment were kept fully blinded to the treatment being received by the study participants.