NIPH
Norwegian Institute of Public Health

GUIDELINE

2021

Guidelines for the submission of documentation for single technology assessments (STAs) of medical devices and diagnostic interventions

Guidelines for the submission of documentation for single technology assessments (STAs) of medical devices and diagnostic interventions

Introduction

The priority-setting criteria for publicly financed healthcare services have been established through the Norwegian Government's White Paper no. 34 (2015-2016) (1), hereafter referred to as the Priority-Setting White Paper ("Prioriteringsmeldingen"). Furthermore, the regional health authorities' responsibility for the system through which new health interventions should be evaluated, prioritized and potentially adopted were enacted into law in December 2019[1].This system requires that specialist healthcare service interventions shall be evaluated with regard to three prioritisation criteria – the benefit criterion, the resource criterion and the severity criterion. The Decision Forum, the body within the system that determines whether an intervention is to be financed or not, performs balanced considerations of these three criteria and the potential budget impact for the specialist healthcare services of publicly financing the technology in question.

These guidelines for submissions of documentation for single technology assessments (STAs) of medical devices and diagnostic interventions reflect the principles of health sector priority-setting and define the necessary components for appropriate submission of documentation for STAs of medical technologies (devices and diagnostic interventions[2]) to be financed by the public specialist healthcare services in Norway.

The guidelines describe the content and preferred methods (and alternatives) for the preparation and submission of documentation on STAs of medical devices and diagnostic interventions. Any deviation from these requirements and recommendations must be justified. The guidelines do *not* describe how the Norwegian Institute of Public Health will evaluate the documentation beyond what is stated above.

A template for the submission of documentation for STA of medical devices, diagnostic interventions and procedures must be used when preparing a submission. The template is available at: https://www.fhi.no/kk/metodevurdering/

Submitter (manufacturer or representative agent) are required to calculate a cost-effectiveness ratio reflecting the use of resources in relation to benefit. This should be done in a health economic evaluation, typically involving decision analytic modelling. The cost-effectiveness ratio will be weighed against the severity of the relevant condition/disease. For more severe clinical conditions, it may be possible that the Decision Forum is willing to consider a relatively higher cost-effectiveness ratio. Benefit should be measured in Quality-Adjusted Life Years (QALYs). The benefit depends on the relative efficacy of the intervention on patient survival and health-related quality of life.

[1] https://nyemetoder.no/nyheter/lovfesting-av-prioriteringskriteriene-og-system

[2] They would in theory also apply to procedures, but NIPH does not anticipate submissions in this area.

The Decision Forum will base its overall evaluation of an intervention on both the factors described above and discretionary considerations. The latter may be linked to evaluation of the quality and the level of uncertainty in the documentation, as well as the budget impact. To assess whether the established prioritisation criteria are satisfactorily fulfilled, the manufacturer or representative agent is required to submit an evaluation of the cost effectiveness of the interventions in question as an integral element of the STA, i.e. a health economic analysis. In this all effects (benefits and adverse effects) and resource use/costs of the intervention should be compared to that of the most relevant alternative. Details on how the STA, including the health economic analysis is required to be carried out are described in this guideline.

These guidelines may be updated as necessary, for example, if new guidance, new evidence or experience etc. require it.

Contents

INTRODUCTION 1

CONTENTS 3

LIST OF ABBREVIATIONS 7

1. GENERAL INFORMATION ABOUT SUBMISSION OF DOCUMENTATION 9

1.1 The STA in brief 9

1.2 Template 9

1.3 About NIPH 9

1.4 References 9

1.5 Submitter contact information 9

1.6 Language 9

1.7 Confidentiality 10

1.8 Medical devices and in vitro medical devices: regulatory Issues 10

1.9 The Norwegian legislation 10
1.9.1 The Medical Devices Act 10
1.9.2 The Medical Devices Regulation 10
1.9.3 Regulation on the use of medical devices 11
1.9.4 New EU regulations on medical devices 11

2. SCOPE 12

2.1 Scope 12

2.2 Process 12

3. DESCRIPTION OF THE DISEASE/CONDITION, PATIENT POPULATION AND COMPARATOR 12

3.1 Description of the disease/condition and Norwegian clinical practice 12

3.2 Patient population 12

3.3 Intervention 13

3.4 Comparator(s) 13
3.4.1 Main rule 13
3.4.2 Several comparators 13

4. LITERATURE SEARCH AND SELECTION OF RELEVANT DOCUMENTATION 14

5. DOCUMENTATION OF CLINICAL EFFICACY AND SAFETY 16

5.1 Efficacy of intervention and comparator(s) **16**
5.1.1 Data sources 16

5.2 Effectiveness-modifying factors **16**
5.2.1 The learning curve 16
5.2.2 Product modification 17

5.3 Safety: Description of harms and adverse events **17**

5.4 Description of wider organizational implications **17**

5.5 Patient/user experience **17**

5.6 Ongoing studies **17**

6. DOCUMENTATION OF RELATIVE EFFICACY/EFFECTIVENESS 18

6.1 General information **18**
6.1.1 Direct comparisons 18
6.1.2 Indirect comparisons 18

6.2 Observational studies **18**

6.3 Statistical methods **18**

6.4 Extrapolation of relative effectiveness **19**

6.5 Use of Real World Data **19**

7. DIAGNOSTIC INTERVENTIONS 20

7.1 Introduction **20**

7.2 Diagnostic test accuracy **20**

7.3 In-house diagnostic equipment **20**

8. HEALTH-RELATED QUALITY OF LIFE 21

8.1 Instruments for measuring health-related quality of life **21**

8.2 Tariffs for setting values of health-related quality of life **21**

8.3 Mapping of quality of life data **22**

8.4 Age adjustment of health state utility values (HSUV) **22**

8.5 Treatment-specific HSUVs for the same condition **22**

8.6 Effect on the health-related quality of life of caregivers **23**

9. HEALTH ECONOMIC ANALYSES 24

9.1 Analysis methods 24
9.1.1 Cost-utility analysis (CUA) 24
9.1.2 Cost-minimisation analysis 24

9.2 Analysis perspectives 24

9.3 Resource use and costs 24
9.3.1 More about unit costs 25
9.3.2 Patients' and caregivers' time– and unit costs 26
9.3.3 Projection of unit costs 26
9.3.4 Capital costs and fixed medical equipment 27
9.3.5 Considerations for diagnostic interventions 27

9.4 Present value and discounting 27

10. MODELLING 28

10.1 Modelling of endpoints 28
10.1.1 Consistency between studies, Norwegian clinical practice and modelling 28
10.1.2 Presentation 29

10.2 Sequence modelling 29

10.3 Modelling diagnostic interventions 29

10.4 Organizational implications 29
10.4.1 Organizational changes 29
10.4.2 Minimum level of use 29

10.5 Time horizon 30

11. CALCULATION OF SEVERITY 31

11.1 Types of economic analysis 31

11.2 Treatment interventions 31

11.3 Interventions which treats several diseases/conditions 31

11.4 Calibrating two data sources 31

11.5 Preventive measures 32
11.5.1 Case 1 – Only one disease/condition is prevented 32
11.5.2 Case 2. Several diseases/conditions are prevented 32

11.6 Comorbidity and harm 32

12. UNCERTAINTY 34

12.1 Terminology used to describe uncertainty 34

12.2 Dealing with uncertainty in the analyses 34
12.2.1 Deterministic sensitivity analysis 34
12.2.2 Probabilistic sensitivity analysis (PSA) 35
12.2.3 "Value of Information" analysis 35

13. BUDGET IMPACT 36

13.1 Implantable/wearable and other non-shared medical devices **37**

13.2 Fixed and other shared medical devices **38**

14. ACKNOWLEDGEMENTS 40

15. REFERENCES 41

APPENDIX 1. DOCUMENTATION OF RELATIVE EFFICACY ININDIRECT COMPARISONS 44
1.1 General 44
1.2 Literature search 44
1.3 Assumptions 44
1.4 Statistical methods 44

APPENDIX 2. USE OF TIME TO EVENT DATA IN HEALTHECONOMIC ANALYSES 49
2.1 Introduction 49
2.2 Parametrisation of data from clinical studies 49
2.2.1 Curve fitting to observed study data 50
2.2.2 Plausibility of the extrapolated part of the curve 51
2.2.3 Algorithm and implementation in the health economic model 52
2.3 Studies where patients can switch to active (new) intervention 53

APPENDIX 3. QUALITY OF LIFE DATA 55
3.1 Example of age adjustment of future expected HSUV using the multiplicative method 55

APPENDIX 4. CALCULATING SEVERITY 57
4. Detailed procedure for calculating absolute shortfall (AS) for treatment interventions 57
4.1.2 Expected remaining QALYs for the general population 57
4.1.3 Prognosis 57
4.1.4 Absolute shortfall 58
4.2 Examples – calculation of degree of severity for treatmentinterventions 58
4.3 Example of calibrating two data sources – level adjustment 60
4.4 Examples – calculation of degree of severity for preventive measures 60
4.5 Justification of the suggested principle for weighted AS for measureswhich prevent and/or treat several types of disease 61
4.6 Expected remaining QALYs in the general population 62

List of abbreviations

AFT	Accelerated failure time model
AIC	Akaike Information Criteria
AS	Absolute shortfall
MTCCDx	Companion diagnostics
CEA	Cost-effectiveness analysis
CEAC	Cost-effectiveness acceptability curves
CrI	Credible intervals
CUA	Cost-utility analysis
DSU	Decision Support Unit
EMA	European Medicines Agency
EQ-5D	EuroQol- 5 dimensions
EUDAMED	European database of medical devices
EUnetHTA	European network for health technology assessment
EVPI	Expected value of perfect information
EVPPI	Expected value of partial perfect information
FDA	Food and Drug Administration (U.S.)
Helfo	Health Economics Administration
HR	Hazard ratio
HRQoL	Health-related quality of life
HSUV	Health state utility value (also referred to as QALY weight)
HTA	Health technology assessment
ICER	Incremental cost-effectiveness ratio
IPD	Individual patient data
ISF	Activity-based financing («Innsatsstyrt finansiering»)
ITC	Indirect treatment comparison

ITT	Intention to treat
KM	Kaplan-Meier
LIS	Norwegian Hospital Procurement Trust, Division Pharmaceuticals (Sykehusinnkjøp HF divisjon legemidler (LIS))
LYG	Life years gained
MAIC	Matching Adjusted Indirect Comparisons
MeSH	Medical Subject Headings
MTA	Multiple Technology Assessment
MTC	Mixed treatment comparison
NIPH	The Norwegian Institute of Public Health
n eff	Effective sample size, ESS
NICE	National Institute for Health and Care Excellence
NMA	Network meta-analysis
OS	Overall survival
PH	Proportional hazards
PICO	Patient population, intervention, comparator and outcome measures.
PSA	Probabilistic sensitivity analyses
QALYs	Quality-Adjusted Life Years
RCT	Randomised controlled trial
RHA	Regional health authority
ROPE	Region of practical equivalence
RWD	Real World Data
SSB KOSTRA	Statistics Norway (Municipality-State-Reporting)
STA	Single Technology Assessment
STC	Simulated Treatment Comparisons
TTE	Time-to-event
UDI	Unique Device Identification
VoI	Value of information analysis

1. General information about submission of documentation

1.1 The single technology assessment in brief

Your submission must be based on the Norwegian Institute of Public Health's (NIPH) template (see below) and include: description of the disease or condition for which the intervention is designed and how it fits in to Norwegian clinical practice. You should provide comprehensive documentation of its efficacy and safety. A health economic analysis should be ppl be included, and a model in either TreeAge© or Microsoft Excel© should be enclosed. You should calculate the severity of the disease/condition for which the intervention is designed in terms of absolute shortfall. Finally, you should account for the potential fiscal consequences for the Norwegian specialist health services in a budget impact analysis.

1.2 Template

The NIPH has developed a template that you must use when you prepare documentation for single technology assessment (STA). The template is available on the institute's webpage about health technology assessment. The template outlines the information that is required for the submission of an STA for the medical technologies in question, including overarching topics, subheadings in these, and tables to facilitate the summarization. If needed appendices or supplementary information may be attached to the main submission.

1.3 About NIPH

The NIPH (www.fhi.no), is responsible for performing health technology assessments at the national level, including specific tasks in the National System for Managed Introduction of New Technologies in the Specialist Health Service, often referred to as New Methods[3]. The institute carries out full HTAs, also known as multiple technology assessments of pharmaceuticals, medical devices, diagnostic interventions and procedures. NIPH also reviews STAs of medical devices and diagnostic interventions submitted by the manufacturer or a representative agent. An STA incorporates documentation of the intervention's effectiveness, safety and cost-effectiveness (the latter in the form of a health economic analysis), and an analysis of the budget impact for the specialist healthcare services of adopting the technology in question.

1.4 References

Enclose the most important references with your submission. References for all pivotal studies and references used as the *basis for input data* in the health economic analysis and in calculations of severity and budget impact must be provided. References used in documents, models and spreadsheets must be formatted so that they are linked directly to the individual publication/file (in PDF). We encourage you to include relevant published, and unpublished data in your submissions. The use of unpublished data ("data on file")/confidential data must be discussed with NIPH in advance.

1.5 Submitter contact information

State the name of the person(s) responsible for preparation of the submitted documentation, and the names of others who have taken part in the work.

1.6 Language

Documentation for STAs of medical devices and diagnostic interventions must be written in Norwegian, English, Swedish or Danish.

[3] https://nyemetoder.no/english

1.7 Confidentiality

NIPH operates within the Public Administration Act and the Freedom of Information Act. Please contact NIPH for further information as to how the institute handles confidentiality in health technology assessments. If any part of the information you provided is considered confidential, you should clearly state this in the submission.

1.8 Medical devices and in vitro medical devices: regulatory Issues

The Norwegian Medicines Agency is the competent authority for regulatory approval of medical devices in Norway. This implies that the Agency has administrative and advisory responsibilities related to legislation and supervisory authority over manufacturers, distributors, economic operators and notified bodies. A medical device is a device intended by its manufacturer specifically for use to diagnose, prevent, monitor, treat or alleviate disease, injury or disability in humans. Some birth control products, as well as aids for individuals with handicaps are also considered medical devices (See Directive 93/42 article 1, 2a for the complete English definition). In vitro medical devices (IVD) are intended for the analysis of samples from the human body for medical purposes.

Medical devices must comply with the essential requirements in the regulations and directives before the device can receive CE marking which confirms that regulatory requirements have been met. The process for CE marking depends on the risk classification of the device. For higher risk classes, a notified body conducts conformity assessments to assess whether a medical device complies with the regulatory requirements. The regulations on medical devices also stipulates requirements for clinical investigations in Norway. Notification of the trials should be sent to the Norwegian Medicines Agency (further information on clinical investigation).

The Norwegian Radiation and Nuclear Safety Authority is the national authority and expert body in matters concerning nuclear security, radiation use, natural radiation and radioactive contamination in the environment. A medical device or method involving ionising or non-ionising radiation must comply with requirements stated in the Radiation Protection legislation and also in the Pollution Control legislation, when relevant. The Norwegian Radiation and Nuclear Safety Authority is part of the “New Methods" to ensure that the justification regarding the evaluation of new methods and applications in medical use of radiation on a general basis is documented before such methods and applications become available for general use. A method is justified if the total diagnostic or therapeutic benefit, for the individual and society, is higher than the disadvantages related to the use of radiation. The radiation harm should therefore be included in the evaluation and the methods safety.

1.9 The Norwegian legislation

The EU Directives on medical devices are implemented in Act of 12 January 1995 no. 6 on medical devices and Regulation of 15 December 2005 no. 1690 on medical devices.

1.9.1 The Medical Devices Act

The Medical Devices Act regulates the manufacture, marketing, trade and use of medical devices. The purpose of the Act is to prevent harmful effects, incidents and accidents, and to ensure that medical devices are evaluated and used in a professional and ethical manner.

1.9.2 The Medical Devices Regulation

The Regulation applies to everyone who manufactures, markets, and trades medical devices for use in Norway, and Notified bodies. The aim of the regulation is to ensure that medical devices do not cause harmful effects to patients, users or any other person's safety in association with manufacturing, construction, trade and use of these devices.

1.9.3 Regulation on the use of medical devices

This regulation applies to the healthcare services' use of medical devices and any use of electro-medical devices by public and private business. The aim of the regulation is to ensure that, at any given time, medical devices are safe, correctly maintained, and used correctly in accordance with their intended purpose.

1.9.4 New EU regulations on medical devices

On April 5th 2017, the EU adopted two new regulations on medical devices and in vitro diagnostic medical devices. They were entered into force on May 25th 2017, and will progressively replace the existing directives. The new regulations are fully applicable in May 2021 for medical devices and May 2022 for in vitro diagnostic medical devices. The new regulations were published in the Official Journal of the European Union, L117, May 5th 2017 (eur-lex.europa.eu).

The new regulations establish a modernized and more robust EU legislative framework to ensure better protection of public health and patient safety, with the intent of boosting confidence in the medical devices industry. The new EU regulations on medical device and in vitro-diagnostic medical device will be incorporated into the EEA Agreement, and then implemented into Norwegian law. For more information: Ministry of Health and Care Services has performed a public consultation on implementing the Regulations into Norwegian law.

The new requirements involve elements meant to improve patient safety, such as Unique Device Identification (UDI), enhanced tracking, monitoring and database systems.

You must provide references to documentation showing that the device or diagnostic equipment complies with these requirements.

You must also provide a description of the status of the device or diagnostic equipment with regard to CE-marking.

2. Scope

2.1 Scope

The scope of the STA submission includes a short description of the indications for the intervention in question; the patient population to be considered; a brief description of intervention; the alternative interventions for comparison; and the most important outcome measures in the analysis (as defined by the PICO).

The documentation must be submitted in accordance with the request from the Commissioning Forum, "Nye metoder" (Bestillerforum for nye metoder). NIPH must agree in advance to any variation from the requested documentation. Communication throughout the duration of the assessment will be with NIPH. Submitters should not contact the Commissioning Forum or the Regional Health Authorities, RHA, directly.

Describe briefly which method of health economic analysis you have employed (cost-utility analysis, cost-minimisation analyses etc.).

In-house diagnostics
Note that in-house diagnostics are not covered by these guidelines, see Chapter 7 for details.

2.2 Process

Pre-submission meeting is held to clarify expectations to the submission file. The submitter notifies NIPH within 4 weeks from the pre-submission meeting about whether they will submit documentation. If yes: NIPH will expect the actual submission within 4 to 6 months. From the time point when a submission is considered to cover formal requirements, NIPH has 180 days to perform the assessment. After the pre-submission meeting, the submitter may contact NIPH for clarifications regarding the documentation in the submission. The STA report will be publicly available on NIPH's website.

3. Description of the disease/condition, patient population and comparator

3.1 Description of the disease/condition and Norwegian clinical practice

Describe the disease and Norwegian clinical practice.
Briefly describe the relevant disease or condition targeted by the proposed intervention and how patients are currently diagnosed and treated in Norway. Ideally refer to national guidelines and to current Norwegian clinical practice. Specify any clinicians or key opinion leaders you have consulted in order to confirm clinical practice.

Provide information on the prevalence and incidence of the disease/condition in Norway, and developments during the last five years. For very small patient groups, also include the worldwide prevalence. NIPH accept that Norwegian data may not be available for very small patient groups.

3.2 Patient population

Describe the patient population and any relevant subgroups.
Describe as precisely as possible, the patient population in Norway the intervention is intended for. Specify if the analysis covers only a portion of the intervention's indications/areas of use, or specific subgroups of the population. Identify the age group that is most affected by the disease or condition and state the mean age (or median age) of the relevant patient group in Norway (not the age of the study population or populations). Confirm the mean (median) age by citing clinical experts, registry data or other relevant sources. In the case of diagnostic testing, the age should reflect the current mean age of the population of patients with the disease the diagnostic test is designed to detect. If you believe there are subgroups of patients for whom the intervention

may have an efficacy and safety profile that differs from that of the overall population under consideration, detail the reasons for the anticipated differences. Refer to relevant data and specify whether the sub-groups were pre-defined in clinical studies. Describe relevant diagnostic tests and methods used to select patients.

3.3 Intervention

Describe the intervention and treatment pathway.
Describe the intervention in accordance with the template for submission of documentation (see Chapter 1), including its main characteristics and mechanisms of action. The description must outline any institutional or organisational structures that need to be established to ensure rational use of the intervention, as well as requirements for data management and interpretation of results. In addition, specify requirements for quality assurance procedures.

You should also describe the position the intervention is supposed to fill in the pathway for the handling of the defined population.

3.4 Comparator(s)

Account for the choice of comparator(s) based on the following guidelines (3.4.1 and 3.4.2). Contact NIPH for guidance if you have any questions about the choice of comparator(s).

3.4.1 Main rule

The relevant comparator is the intervention (treatment, procedure, diagnostic, etc.) currently used for the population described in section 3.1 or the intervention that will most likely be completely or partially replaced if the proposed intervention is implemented in clinical practice. If there is more than one relevant comparator, the proposed intervention must be evaluated with respect to each comparator (see Section 3.3.2).

The relevant comparator will often be the current established practice (for example, indicated by the national clinical guidelines) or the treatment that is most commonly used (number of patients). The comparator(s) may for example take the form of an alternative diagnostics, pharmacological treatment, procedure or other "standard of care", including "wait and see" strategies. Comparison with no treatment will only be relevant in exceptional cases.

3.4.2 Several comparators

When there is no clear single alternative, but there are several commonly used alternatives, then more comparators should be included.

The comparators must be presented in their individual form, i.e. not as mergers of two or more alternatives using, for example, average effects, costs etc. A comparison using a combined alternative cannot show whether the intervention is cost-effective relative to each of the individual comparators.

Some randomised, controlled trials have an "investigator's choice" control arm. In such cases, it is not always possible to individualize the alternatives, and even if it is possible, this can lead to a reduction in the strength of the results. Whether the "investigator's choice" or one of the individualized alternatives are to be used in the STA must be justified in each case.

4. Literature search and selection of relevant documentation

The aim of the literature search is to identify relevant documentation for:

- Efficacy and safety data used to document an intervention's and the chosen relevant comparators relative efficacy and safety
- Quality-of life data: Health state utility values (HSUV) for all relevant health states (if the data from the literature is used in health economic analyses or calculations of severity)
- Any other key data for which a literature search will help improve the quality of the documentation

The search for literature should be based on internationally validated methods (2-5). It should be clearly stated which databases that have been searched:

Mandatory databases:

- Cochrane Central Register of Controlled Trials
- Embase
- MEDLINE/PubMed
- US National Library of Medicine Clinical Trials
- WHO International Clinical Trials Registry Platform

Optional databases:

- INAHTA database (The International Network of Agencies for Health Technology Assessment)
- NICE Guidance (NICE Diagnostics guidance; NICE Highly specialized technologies guidance; NICE Interventional procedures guidance; NICE Medical technologies guidance, or NICE technology appraisal guidance as appropriate)
- Epistemonikos
- Subject specific databases (PsycINFO, Pedro, etc.)

The search strategy should be reported in such detail that it allows the search in at least one chosen database to be reproduced. This means that following information should be available:

- Search strings, MeSH-terms and truncations
- Number of hits
- Search date
- A list of all databases used in the search

Inclusion and exclusion criteria for study selection should be documented by stating

- A precise formulation of the research question (including specified PICOs)
- An *a priori* definition of the inclusion and exclusion criteria, and reasoning for these
- Justification for the chosen timeframe
- A description of the selection process (including information on whether one or more reviewers were involved and how disagreements were handled, e.g., by an independent professional colleague)
- A list of excluded studies detailing the reason for studies that a reader may expect to see among the included studies should be given in the appendix. This covers studies may, on the surface, appear to meet selection criteria band the basis for exclusion.

If the original literature search is more than six months old, it must be updated by repeating the search for the following period. Relevant studies identified in the updated search should be listed in the appendix. A manual search in other sources for e.g. conference posters and conference abstracts may also be relevant.

Additional information:

- EUnetHTA [web document]. Germany/Belgium: Guideline Information Retrieval. Available from: https://eunethta.eu/wp-content/uploads/2020/01/EUnetHTA_Guideline_Information_Retrieval_v2-0.pdf
- Cochrane Handbook [web document]. England: Chapter 4 [Oppdatert 09.20, lest 02.02.21]. Available from: https://training.cochrane.org/handbook/current/chapter-04
- NICE Guidelines [web document]. England: The manual [Updated 15.10.20, read 02.02.21]. Available from: https://www.nice.org.uk/process/pmg20/resources/developing-nice-guidelines-the-manual-pdf-72286708700869

5. Documentation of clinical efficacy and safety

5.1 Efficacy of intervention and comparator(s)

Present details from clinical studies.
In accordance with the template for submitting documentation to NIPH, you must present the clinical studies associated with the intervention under consideration and relevant comparator(s). The following information must be included: study design, study objective, intervention, comparator, study duration and follow-up, patient population with inclusion and exclusion criteria, sample size, and outcome measures. It should be informed whether these studies form part of the evidence supporting the health economic model included in the health technology assessment. Present the results for the primary and most important secondary endpoints, including the study duration and follow up.

5.1.1 Data sources

You may submit the following data to support the efficacy/effectiveness parameters used in the health economic analysis:

- Relevant published/unpublished data, information on terminated or ongoing studies
- Expert opinion used, for example, to describe operator experience in clinical practice and describe patient pathway, patient group/stakeholder opinion used to describe user experiences regarding experiences of the intervention effectiveness, adherence, and adverse events. Note that any data considered "commercial in confidence" must be clearly identified
- A critical appraisal of the data included in the assessment should be performed;
    - internal validity, i.e. Risk of Bias
    - external validity, i.e. to what extent the data are generalizable to Norwegian clinical practice

Present any additional data from other sources
If you have included additional endpoints in the health economic model, such as utility weights, specify the sources for the additional data and explain why the additional endpoints are relevant for the technology assessment.

5.2 Effectiveness-modifying factors

Medical devices and diagnostic interventions can be interpreted as complex interventions whose total effectiveness depends on multiple factors. If relevant, account for factors such as those mentioned below that apply to the intervention and comparator in the submission (this list is not necessarily an exhaustive guide) which may modify the decision on whether to publicly finance the intervention or not.

If relevant, you may submit information pertaining to the following:

5.2.1 The learning curve

The learning curve reflects the extent to which the skills and experience of the operator of the medical device, diagnostic intervention or performer of a procedure will influence the total effectiveness. If the device or procedure requires important skill acquisition, the impact of the learning curve should be appraised. Describe how learning is likely to affect effectiveness over time (6). Examples of underlying mechanisms that influence the curve might include "user experience, community experience/system learning and case-mix of patients in a given centre" (7, 8). Indicate the expected amount of time needed for the curve to stabilize.

Describe how, if at all, the impact of the learning curve was accounted for in the clinical evidence, for example, with respect to single- or multi-centre trials and if experts were involved.

5.2.2 Product modification

Frequent product modifications and upgrades may limit a firm's ability to identify a "steady state" period during which it is appropriate to evaluate a medical device in a randomized controlled trial (9). It is acceptable to include evidence to support earlier versions of the device or diagnostic intervention considered, together with descriptions of any upgrades or changes regarding specific properties. The significance of these changes should also be discussed. Perform a similar assessment, with respect to any anticipated further changes, for devices currently under assessment in ongoing studies.

5.3 Safety: Description of harms and adverse events

Describe the harms/adverse events associated with the intervention and comparator.
Harms or adverse events may have an impact on patients' adherence, mortality, quality of life and resource use (10). Therefore, in accordance with the template for submission of documentation, you must submit an overview of harms/adverse events as reported for both intervention and comparator in the studies (see above section).

Moreover, you must describe the harms/adverse events of significance to the technology assessment in terms of frequency, severity, and duration. Further, describe the treatment of harms/adverse events in clinical practice (monitoring, follow-up, resource use and costs). If there are critical levels of exposure or accumulated risk e.g. radiation, these should also be specified.

State why these harms/adverse events are included or excluded in the health economic model and how their impact is included in the modelling (e.g. reduced quality of life, treatment costs). If the device/ diagnostic intervention under assessment involves exposure to radiation, please seek advice from the Norwegian Radiation and Nuclear Safety Authority (www.dsa.no/en/) for further guidance before submitting the assessment.

5.4 Description of wider organizational implications

Describe, when relevant, additional changes to the health system necessitated by the proposed intervention.
Examples could be additional staff training, new patient administration or referral routines, changes in storage capacity, etc. Specify anticipated organizational implications as precisely as possible.

5.5 Patient/user experience

If you have access to qualitative data or other sources of information that can provide insight into user experiences associated with the intervention, NIPH recommends that you describe these and provide references.

5.6 Ongoing studies

Submit information on ongoing studies and their status (i.e. not yet recruiting, recruiting) incorporating the devices, methods, or procedures relevant to the technology assessment. As pointed out in Section 5.2.2 on incremental innovation, studies on modified versions of the intervention may be included if the modified properties and their significance are clearly described.

6. Documentation of relative efficacy/effectiveness

6.1 General information

Your documentation of relative efficacy and safety should be based on systematic literature searches (see Chapter 4). NIPH prefers efficacy and safety data from randomised controlled trials over data from studies of other designs. However, NIPH does acknowledge that relevant RCTs are not always available and that observational studies can provide additional information and may therefore be accepted.

6.1.1 Direct comparisons

NIPH prefers that you submit head to head comparisons of relevant randomised controlled clinical trials or relevant systematic reviews of such studies. Other designs such as observational studies, may also be accepted (see below).

6.1.2 Indirect comparisons

In cases where no direct comparisons between the intervention and a relevant comparator are available, you may perform indirect comparisons. These may include matched pairwise comparisons, network meta-analyses or other valid methods. In such cases, perform quantitative and qualitative evaluations, using tools validated for this purpose to describe the risk for systematic biases in the studies and data sources.

Non-adjusted indirect comparisons will generally not be accepted. Regarding the presentation and description of the studies, please refer to the template for submission of documentation. Matched indirect comparisons should be performed using appropriate, transparent, and validated statistical methods. You should present and discuss underlying assumptions. For further details, see Appendix 1.

6.2 Observational studies

Professional discussions of appropriate study designs for medical devices note (11) that "the very nature of devices can make performing a randomised controlled trial (RCT) difficult, biased or even unfeasible". Factors related to blinding, and recruitment/drop-out may explain why RCTs for devices are difficult to perform. Such factors may lead to weaknesses in the estimation of effectiveness and should be discussed.

Given the difficulty of conducting RCTs for medical devices, observational studies are often an important source of evidence of clinical effect and safety of these interventions. They are, however, likely to involve a higher risk of bias than RCTs (12, 13). Such biases include selection bias and confounding factors, which means that sophisticated statistical methods are required when relative effectiveness needs to be determined. Describe and justify such statistical methods when used in the submission.

6.3 Statistical methods

In cases where both direct and indirect comparisons are available, a mixed treatment comparison (MTC) may be performed. For more details, refer to Appendix 1. If, however, no coherent network of studies is available to link the intervention and comparator, relative effectiveness must be documented by means of single-arm clinical studies. Appropriate statistical methods should nevertheless be employed to estimate relative effectiveness. If individual patient data (IPD) are available, then analyses such as Matching Adjusted Indirect Comparisons (MAIC) or Simulated Treatment Comparisons (STC) may be carried out subject to other conditions being met.

6.4 Extrapolation of relative effectiveness

Justify the assumptions regarding estimated differences in effectiveness extending beyond the duration of the clinical trials. Estimation of time-to-event data may be performed. An example of time-to-event data might be time to progression in a certain disease, or time to failure of a certain device. Specific guidance for performing time-to-event estimation is taken from the "Guidelines for the submission of documentation for single technology assessment (STA) of pharmaceuticals" from the Norwegian Medicines Agency, and is outlined in Appendix 2 in this document.

6.5 Use of Real World Data

Real World Data (RWD) refer to observational data as opposed to data gathered in an experimental setting such as randomized controlled trials (RCT). Example of sources can be cohort studies, case control studies, types of pragmatic clinical trials and registry data from clinical practice.

Pivotal randomized clinical studies are the preferred source of efficacy data in a cost-utility model. However, you can use RWD to support evidence of, for example, epidemiology, treatment durationin clinical practice, resource use, survival, or adherence to treatment in the Norwegian clinical practice. If RWD are used as a source for modelling the comparator arm when relevant clinical datafor comparator is lacking, you must provide a detailed discussion of RWD source quality, study design (including endpoint definition, inclusion criteria, timing of data collection), patient characteristics, statistical considerations (e.g. how missing data were managed). Similarities and differences between the pivotal clinical trial and RWD should be examined. A discussion on how representative RWD are of the population should also be provided. Any source of bias should be highlighted.

7. Diagnostic interventions

7.1 Introduction

Diagnostic tests are used to inform clinical decision making, for example by predicting individuals who may benefit from a specific medical treatment. When the use of a diagnostic test is linked to a treatment decision, the evaluation of the diagnostic test should ideally be supported by studies that follow patients from testing via treatment to final clinical outcome, so-called end-to-end studies (14, 15). If end-to-end studies are available, they must be submitted to NIPH. Then clinical and relative efficacy can be reported according to Chapter 5 and Chapter 6 in this guideline. If end-to-end studies are not available, it is acceptable to provide separate evidence for test characteristics, analytical validity, clinical validity, and clinical treatment outcomes (clinical utility) to demonstrate how they are linked and estimated in the health economic model.

7.2 Diagnostic test accuracy

Submitted information if a diagnostic test is under evaluation should include:

- The target condition
- The scope of the test (diagnostic, monitoring, screening, prognostic)
- The type of test, i.e. technology and how it is done
- Prevalence of the target condition in the relevant test population
- Cut-off values (rule in/rule out)
- Associated decision rules/algorithms
- Detection limits

Specify whether the test is predictive, diagnostic or both. State whether the test will replace another test and if it is a stand-alone or complementary test. Describe the position of the test in an integrated diagnostic process and in the clinical pathway. Explain how the test is performed in clinical practice, and provide information on turn-around time, amount of biological material needed (if applicable), ease of interpretation of the test, if the test is qualitative or quantitative, training and equipment needed to perform the test. Also describe characteristics that may be important to the patient but are not captured by the test outcomes (e.g. feasibility, risk of adverse events, comfort).

Provide information on the decision rule/algorithms and whether it is in the public domain. Specify parameters such as reference standard, prevalence, and test results in terms of sensitivity and specificity. The internal validity and applicability of included diagnostic accuracy studies should be critically appraised using an appropriate instrument such as QUADAS-2 (16). Special considerations regarding resource use and health economic modelling on diagnostic intervention are mention in Chapter 9 and 10 in this guideline.

7.3 In-house diagnostic equipment

In some cases, diagnostic equipment or devices are manufactured within health institutions ('in-house') since the equipment versions available on the open market do not necessarily tend to different institutions' specific requirements. As mentioned in Chapter 1, the EU has implemented a new and more rigorous set of regulations that will have a bearing on the use of diagnostic interventions. Specific regulations apply to in-house diagnostic devices when applicable patients considered to be at high risk for a known disease. For example, clinical studies are required unless the available data is deemed sufficient[4]. In-house diagnostics are not covered by these guidelines but are likely to be the subject of specific guidelines to be issued in the future.

[4] The Norwegian Medicines Agency, presentation on in-house medical equipment.

8. Health-related quality of life

NIPH prefers that you use Quality-Adjusted Life-Years (QALYs) at the group level as the benefit measure for the health economic analysis, and patient-reported measures based on EQ-5D.

Quality of life data used in STAs must be reported in line with the template for submission of documentation. Quality of life data may be taken directly from the clinical studies used to document relative efficacy or by performing a separate literature search for relevant quality of life data. If the clinical studies used to document relative clinical efficacy and safety also report measures of quality of life based on EQ-5D, justification must be provided if these data have not been used in the health economic analysis. If quality of life data from other literature sources are used, provide documentation of the systematic literature search that was performed (Chapter 4).

Uncertainty in health state utility values (HSUV)[5] must be examined in sensitivity/scenario analyses (Chapter 12).

8.1 Instruments for measuring health-related quality of life

Health-related quality of life, as defined by Gold et al. and Sanders et al. (17, 18), must, as a rule, be based on generic preference-based instruments. To facilitate comparison between different STAs, EQ-5D (19) must, as a general guidance, be used. If appropriate measurements of quality of life, based on disease-specific instruments, exist in the included studies, report these as supplementary information.

Use of EQ-5D can be waived if there are no data based on EQ-5D methodology for the disease in question, or if EQ-5D has been deemed unsuitable for capturing relevant aspects of quality of life for the patient population in question. If EQ-5D is deemed inappropriate for the relevant patient population, provide evidence to support this claim[6]. For guidance in deciding whether EQ-5D is suitable for the evaluation, see NICE DSU Technical Support Document 8 (20).

Two versions of EQ-5D are currently available. The original version (EQ-5D-3L) describes each health dimension using three severity levels, while the new version (EQ-5D-5L) describes the same dimensions using five severity levels. Because the two EQ-5D versions result in slightly different measures of quality of life, it is important to use one of them as the 'reporting' standard to make comparisons with other studies possible. Until the new 5L version more fully replaces the original 3L version in applied studies, data from 5L should, therefore, be converted to 3L using the method described by van Hout et al (21). The use of EQ-5D-3L as the standard in STAs is based on recommendations from NICE (22).

8.2 Tariffs for setting values of health-related quality of life

In STAs, as a rule, valuation of quality of life must be based on tariffs (value sets) from the preferences of the general population. This is done to ensure consistency across STAs, and to ensure internal consistency between measures of severity and health economic analyses in every STA. In principle there should be agreement between the tariffs used to calculate benefit in the health economic analyses and those that form the basis for calculating severity.

If, in a STA, there are particular reasons for using an experience-based tariff, this should be justified. There should be an explanation for how this tariff varies from a general population-based tariff.

[5] Also referred to as QALY weights

[6] For example, NICE has evaluated EQ-5D as being less suited to measurement of quality of life in connection with loss of hearing, restricted vision or schizophrenia.

The tariff used should be relevant to the adult population living in Norway. As yet there is no representative Norwegian tariff for EQ-5D. However, a representative Norwegian tariff has been estimated for the 15D instrument (23). For consistency, we recommend that the EQ-5D with the UK population-based EQ-5D-3L tariff (24) should be used for STAs in Norway until a more relevant and applicable tariff is available. The Norwegian 15D tariff can be applied in scenario analyses (23).

8.3 Mapping of quality of life data

Where there is a lack of patient-reported EQ-5D data, other generic preference-based instruments may be used (SF-6D, 15D, HUI, AQoL, QWB). The preference-based values from such alternative instruments must then be mapped to EQ-5D values, in accordance with validated methods. Compare the mapping results to published quality of life data for the relevant patient group.

If there is no data from generic instruments, but only from disease-specific instruments, these must be mapped to predict EQ-5D values.

Describe the method used for mapping. For a more detailed description of the methods for mapping quality of life data to EQ-5D, refer to the NICE DSU Technical Support Document 10 (25).

Mapping conversions are necessary to provide comparability across economic evaluations based on alternative methods for determining health state utility values.

8.4 Age adjustment of health state utility values (HSUV)

You should adjust utility values for age

As age cohorts in the general population grow older, its members will in general, gradually experience loss of function and increased morbidity, resulting in a reduction in health-related quality of life. Given this background, the development of HSUVs should be adjusted for age in health economic models. It is the *development* of the HSUVs used over time that should be adjusted, not the level of the HSUV used as the starting age in the models. If the HSUVs are not adjusted for age, this must be justified[7].

Adjusting for age will, in addition, ensure consistency with the severity calculations in STAs, where age-adjusted HSUV should be used in the calculations of expected remaining QALYs for the general population (Appendix 3).

In order to maintain consistency in the methodology for STAs, we recommend that age related adjustments are performed based on the multiplicative method, as described in the NICE DSU Technical Support Document 12 (26). Provide an explanation if another method is chosen.

Calculating HSUV over time, based on the multiplicative method, can be briefly described as the original value for the HSUV multiplied by an adjustment index[8], and gives an age adjusted HSUV. Appendix 3 provides an example of how to perform this calculation.

8.5 Treatment-specific HSUVs for the same condition

If different treatment-specific HSUV are used for the same condition[9], this must be fully justified and documented. For different treatment-specific HSUVs to be accepted, the differences in health-related quality of life should be shown in clinical studies. Different treatment-specific HSUV should have a clinical explanation.

[7] For example, when a health economic analysis has a short time perspective

[8] Is set to 1 in the starting year in the health economic model and decreases with increasing age

[9] i.e. If different HSUV are used for the intervention and the comparator for the same condition in the health economic model. Example: HSUV X is used for the intervention and HSUV Y for the comparator for the health state progression-free survival in a HTA of a cancer medicine

8.6 Effect on the health-related quality of life of caregivers

An intervention's effect on the health-related quality of life of a caregiver may be included if proper documentation is provided. The basic same requirements that apply for documenting changes in patient's quality of life also apply for a caregiver. The effects can be quantified in QALYs, which are then included in the cost-effectiveness ratio. Include cost-effectiveness results *both* with and without the effect on the caregiver's quality of life. In cases where there is good reason to expect considerable changes in the caregiver's health-related quality of life, but where no good documentation exists, include this information in a discussion, but *not* as part of the cost-effectiveness ratio computed in the economic model.

The central effect that can be taken into account is how changes in the patient's health-related quality of life affects the health-related quality of life of the caregiver(s). If the intervention affects the *life expectancy* of the patient, the effects on the caregiver's quality of life of the increased life expectancy in itself should *not* be taken into account. There are both ethical and methodological reasons for this.

9. Health economic analyses

9.1 Analysis methods

9.1.1 Cost-utility analysis (CUA)

The recommended analysis method for health economic evaluations is usually CUA.
When the intervention affects survival, the results must be presented separately as cost per QALY gained and cost per Life Year Gained (LYG).

9.1.2 Cost-minimisation analysis

Use of cost-minimisation analysis requires documentation showing that the efficacy and safety profiles for the intervention and the comparator are approximately equal. In practice, the prerequisite for cost-minimisation analysis can be fulfilled by demonstrating that the intervention is not less effective than the comparator.

9.2 Analysis perspectives

What follows is a description of the benefits and costs that must/must not be included according to the guidance from the Priority-setting White Paper. These are costs and benefits that are expected to arise or change because of the intervention. In practice, the guidance implies a form of *extended* health-service perspective.

The following benefits must be included (if relevant):
Effects on
- The patient's lifespan
- The patient's health-related quality of life
- The health-related quality of life of caregiver(s). The analyses must be presented both with and without inclusion of this effect.

The following costs must be included (if relevant):
- Treatment or prevention costs, funded by the health service or by the patient/relatives
- Transport costs linked to travelling to and from treatment, whether it is funded by the health service, or by the patient/relative
- Patients' and relatives' use of time in connection with treatment

In accordance with the Priority-Setting White Paper the following must not be included:
- Productivity changes as a result of the intervention
- Consequences for patients' future use of public services and receipt of public benefits/pensions
- Unrelated health service costs and savings. For example, the health service costs related to future unrelated illness will not be taken into consideration.
- Taxation expenses associated with public financing
- Public benefits, pension payments, value added tax and other transfer payments

9.3 Resource use and costs

Account for resource use in the analysis

By *resource use* we primarily mean use of goods and services, use of time and use of capital. Market prices in the private sector should, as far as possible, be used as the basis for estimates of unit costs/calculation prices (27). Present and justify unit costs and resource use separately in addition to showing total (average) costs per resource for the alternative interventions. As a rule, Norwegian unit costs must be used, and any deviation from

this must be justified. Show any exchange rate used for converting calculations of costs in other currencies to Norwegian kroner (NOK).

Assumptions and justifications for costs included must be well documented. These must be reported in detail and the way the costs have been arrived at must be transparent, so the calculations can be assessed. This must be presented in accordance with the template for submission of documentation.

9.3.1 More about unit costs

With respect to devices and diagnostic equipment costs, including consumables, you must perform the analyses using the expected retail price without value added tax (VAT). A curve showing the relationship between the ICER and a percentage variation in the retail price (e.g. 10% up or down) for the device or equipment being evaluated, must be presented. It must be possible to change the device or equipment price in the model so that NIPH can carry out its own analyses, refer to Chapter 10 for the requirements for the model.

Transportation costs linked to travel to and from treatment are to be included. If it is relevant and well-documented, necessary transport costs for caregivers can also be included.

If unit costs are not taken from Norwegian sources, they can be taken from other cost studies/publications, but this must be justified. The average cost can generally be used. It is possible to rely on "standard" calculations for average cost per resource type, e.g., visits to doctors, hospital treatment, nursing home costs, laboratory services etc. Examples of some "standard" calculations of average costs based on resource type include:

- **Hospital services:** The cost per hospital admission[10] or outpatient clinic attendance[11] can be calculated by multiplying the DRG-points by the relevant unit price. This gives an estimate of the total costs per admission/attendance for the hospitals[12]. The patient's co-payment for outpatient consultations can be ignored because the contribution is already captured by DRG weighting. If there is no information in the data about the relevant DRG code, then a cost per day or per consultation can be taken from the Norwegian Directorate of Health's database (SAMDATA) for the specialist health services (covers somatic, mental health services and multi-disciplinary specialised addiction treatment). More information about activity-based financing, DRG weighting, unit prices and SAMDATA can be found on the Norwegian Directorate of Health's website.
- **General practitioner and specialist services:** As a rule, the cost per contact (consultation) is calculated by multiplying the remuneration amount from "Normaltariffen" (28) (i.e. the tariff rate in Norwegian kroner) by two (x2). This is because the cost of general practitioner consultations and specialist services (for example, one consultation) is covered by both the remuneration (the total of the reimbursement amount and the patient's contribution) and the public subsidy (basic subsidy to general practitioners, operating subsidy to specialists). The calculation gives a rough estimate. Multiplying the remuneration

[10] Applies to admissions to somatic departments.

[11] Applies to outpatient contacts/consultations in somatic departments, mental health services and multidisciplinary specialist addiction treatment.

[12] Activity based financing (ABF) (in Norwegian: "Innsatsstyrt finansiering – ISF") means that the hospital receives a refund for a share of the total cost of an activity/procedure (the ISF-share). The remaining hospitals costs are covered by its basic funding allocation. Calculation of the ISF-refund is done using the following formula (taken from the Directorate of Health's annual document "Innsatsstyrt finansiering [YEAR]" which is available on the Directorate of Health's website):

$$\text{ISF-refund} = \sum \text{ISF-Point} \times \text{Unit price} \times \text{ISF-share}$$

In the formula it is clear that ISF-points multiplied by the unit price is an estimate of 100 % of the cost of an activity/ procedure. The ISF-share defines how much of the cost of the activity/procedure is refunded. DRG-points will often be the same as ISF-points. In some cases, further adjustments are made to the DRG-points to calculate ISF-points. In these cases, it will normally still be useful to use DRG-points in the estimate of costs. This implies replacing ISF-points with DRG-points in the formula above and uses an ISF-percentage of 100 % to estimate the costs of hospital services.

amount by two is considered to give a better cost estimate than using the tariff rate directly. The Norwegian Medical Association publishes an overview of tariffs, patient contributions, and subsidies on its website.

- **Clinical laboratories and radiology services:** For these services the approach is similar to that described for general practitioner and specialist services. The unit costs are calculated as the total of the tariff per investigation (consultation) and the patient's contribution, multiplied by two (x2). The Health Economics Administration (Helfo) publishes information on tariffs, patient contributions and subsidies on their website.
- **Nursing homes:** Statistics Norway publishes information (KOSTRA) on its website about the cost per day of nursing homes.

9.3.2 Patients' and caregivers' time- and unit costs

Use of time as an input for the intervention and comparator must be included.
The intervention and the comparator can, in some cases, lead to differences between interventions in treatment duration, time to administer the treatment, and/or travel time. In these cases, documented differences in use of time (for patient, and if relevant, for the caregiver) must be estimated and the results of the analysis must be presented with these costs.

Valuing time for patients and caregivers
We recommend setting value of time at equal to the value of leisure time for *all* patients and relatives *regardless* of their employment status.

The value of an hour of increased/decreased leisure time is equivalent to the average wage in Norway after tax (29). Changes in time available for work and/or other daily activities/leisure resulting from the intervention (productivity changes) must not be included[13].

9.3.3 Projection of unit costs

Unit costs, normally, must be held constant throughout the time horizon assumed for the analysis. This can be a reasonable approach to uncertainty about future technological and/or market developments. If, however, there are good reasons for using projections of changes in unit costs over time, these must be described and justified.

For example, competition from similar interventions or improved efficiency in production can lead to substantial price reductions over time, potentially affecting the results of the economic evaluation. If, then the probable price paths must be included if a price change is considered relatively imminent. The anticipated price paths must be justified and the related uncertainty discussed.

[13]This is linked to the fact that treatment can allow the patient to experience more time in good health. If this time is used for paid work (return to work, or work more hours), this is called positive productivity changes, i.e., production gains. Such productivity changes must not be included in the analysis.

9.3.4 Capital costs and fixed medical equipment

Regarding larger, often hospital-based, medical equipment, it is necessary to calculate the cost of each session of patient use. In such cases, you need to estimate several variables, such as the equipment's expected lifetime, the patient volume over a given period, and the overhead cost rate. Maintenance and removal costs should also be included. The capital costs should be amortised over the equipment's lifetime; amortisation provides for "a constant annual cash flow that has the same net present value as the project over the project's lifetime" (30).

NIPH currently has no specific recommendations concerning allocation of overhead costs, as long as the result is deemed reasonably representative of Norwegian clinical practice and the methods are clearly described (for guidance see e.g. the standard text, Drummond et al 2016 (32), Chapter 7). For example, allocation of overhead costs of hospital-based medical equipment may be based on capital asset value, and allocation to procedures could be based on relative human resource input in hours (10). Moreover, the impact of adopting the equipment on the overall hospital infrastructure and capacity (for example, are there sufficient staff numbers available to operate equipment) should be addressed in the submission (10). If the equipment has the potential for use in multiple indications, the average costs of the indication under evaluation should be used. If this can be justified, some of the fixed costs may be assigned to other indications in the sensitivity analysis (14).

9.3.5 Considerations for diagnostic interventions

Cost estimates for the diagnostic intervention should reflect average costs in its expected setting. If the equipment has a potential for use in multiple indications, you should estimate the average costs of the indication. Some fixed costs may be assigned to other indications given proper justification (14). Costs associated with detection and follow-up of true positive and false positive cases should be included. This also applies to health-related quality of life consequences. If introduction of the technology requires additional infrastructure, these costs should be incorporated in the analysis, by inclusion in the average cost. In the case of devices involving radiation for example, this may include additional staffing, need for special professions or competency.

9.4 Present value and discounting

To compare benefits and costs in a cost-utility analysis which occur in different years, the annual benefits and costs must be converted to present values. In calculation of present value both benefits and costs are discounted by 4% per year.

Long time horizons

In the Priority-setting White Paper (St. meld. 34 2015-16) it is stated that the discount rate should be equal to the applicable rate at any time given by the Ministry of Finance. As described in "Rundskriv R-109-2014" the rate should be 4% per year for the first 40 years after the planned start of the intervention, i.e. in the years 0-39. In the years 40-74 a discount rate of 3% per year is to be applied, while thereafter, from (and including) year 75 and onwards, 2% per year is to be used.

10. Modelling

You must justify your choice of health economic model. The model should be as simple, straightforward, and transparent as possible, but still capture all relevant factors that could affect a decision. The model's construction, assumptions, and methods for modelling different input data must be documented and described fully.

There must be consistency among the clinical documentation of relative efficacy, Norwegian clinical practice, and the cost-effectiveness model. Therefore, models should be validated to the extent possible. Internal and external validity should be described. Check carefully whether the calculations are precise and consistent (internal validity). The results from the model should be checked against independent sources (external validity). This can include comparing clinical events predicted by the model against data not used in the model, for example, epidemiological studies. Costs and effects must be estimated within the same model.

You may use international models, but they must be adapted to Norwegian conditions in terms of clinical practice, costs and any relevant health effects. It should be clear how such models have been adapted for Norway. If they have not been adapted, this must be justified. Indicate the consequences any lack of adaptation may have for the results.

NIPH must be able to change all relevant variables and parameters in the model. This includes any parameterising functions. The model must be able to update the sensitivity analysis automatically.

The model must not be locked, time limited, password-protected unless the password is made available, or have any hidden elements that are not described or cannot easily be changed. The model should be submitted in either Microsoft Excel© or TreeAge©.

Data sources used in the model will be considered in terms of three criteria: Context, credibility and consistency (10).

- Context refers to the data's scope and fitness for purpose (i.e. the differences between the data and the decision problem), for example in terms of inclusion and exclusion criteria
- Credibility involves a "perceived lack of bias" (10), i.e. systematic deviation of a parameter estimate from its true value. In other words, the methodological rigour of the data should be sufficient to generate a credible estimate
- Consistency also entails that the data is suited to its intended purpose, as well as being collected and measured in a consistent manner

10.1 Modelling of endpoints

If efficacy data are only available for intermediate endpoints (for example cholesterol levels or blood pressure), the analysis must report how changes in these affect the endpoints in the model (for example, heart attack or stroke). A documentation of causal relationships between the intermediate endpoints and the hard endpoints should be made available.

10.1.1 Consistency between studies, Norwegian clinical practice and modelling

The health economic model must reflect an accurate representation of the clinical course of the disease, with enough detail to capture meaningful costs and benefits, and must reflect Norwegian clinical practice. The data used in the model must originate in the included clinical studies or in the indirect comparison/meta-analysis. As a rule, the effect estimates for the primary endpoint, or the hard endpoints, should be used in the model.

Explain connections and/or any deviations among the data used in the model, clinical data and Norwegian clinical practice.

If the clinical studies used in the health economic analysis also include quality of life data, i.e. health state utilities values (HSUVs) (See, Chapter 8), or data that can be translated into HSUVs, provide a justification if these data are not used in the analysis.

10.1.2 Presentation

Present in a clear table, as described in the template for submission of documentation, all outcome parameters (clinical efficacy, harms/adverse reactions and quality of life) used in the health economic model. The table should also include information on how these have been determined. The definition of the outcomes from different sources must also be presented.

If the results from the studies and the estimates used in the health economic model are not the same, this must be described and justified.

Effectiveness-modifying factors

Describe how any uncertainty associated with the effectiveness-modifying factors, i.e. the learning curve (see chapter 5), is captured in the model.

10.2 Sequence modelling

In some cases, it can be relevant to model treatments as part of a sequence. A prerequisite for this type of approach is the existence of sufficiently good quality documentation of treatment efficacy both for (1) relative differences between different treatment or test pathways and (2) the order of different treatments or tests within the pathways.

10.3 Modelling diagnostic interventions

Describe the relevant treatment alternatives associated with the disease state or target condition and potential clinical outcomes. Provide a detailed explanation of patient pathway and how this is captured in the health economic model. The model should attempt to follow the patient from diagnostic test via treatment to final outcomes. It should be an integrated model in the sense that parameters related to both testing and treatment can be varied and analysed within the same model. The link between intermediate and final outcomes should be supported by documentation (14).

10.4 Organizational implications

10.4.1 Organizational changes

Potential organizational changes that can lead to variation in parameter estimates must be described, justified and incorporated into the health economic model.

State if the intervention is likely to be more cost-effective in one setting than another (31). If there are several organizational options, choose the one most likely to be implemented in Norwegian clinical practice. Explain potential economies of scale and scope resulting from the changes. Clarify any organizational arrangements that are necessary for maintenance, (emergency) extraction and handling (at the end of the device's lifespan).

10.4.2 Minimum level of use

If there is a minimum level of use that is required in order to maintain satisfactory standards and achieve the expected clinical outcomes, specify this, as well as the average annual patient volumes expected in a Norwegian clinical setting. For example, in the case of a surgical implant, specify the expected annual number of

implantations at an average hospital. If possible, note any critical level required for operators to maintain an acceptable skill level.

10.5 Time horizon

The time horizon of the analysis must be long enough to capture all important, future differences in costs and health effects between alternatives. That is, extending the model's time horizon should not affect model results in any meaningful way. Use a lifetime time horizon if the intervention has an effect on mortality.
In some cases, it may be relevant to consider a shorter time horizon. There can be several reasons for this, for example if:

- There is no documentation/it is not likely that the relative efficacy will be maintained over a longer time horizon
- There are other reasons (explain) that justify a shorter time horizon

11. Calculation of severity

You must quantify severity using absolute shortfall (AS) in health technology assessments.
Absolute shortfall is the number of future healthy life years an average patient in the patient group will lose because of the disease, compared to the average in the population of the same age. Absolute shortfall is the same as the reduction in expected future healthy life years without the treatment under consideration (i.e., with the current standard treatment). The term 'healthy life years' contains two dimensions – lifetime and life quality – which are expressed as QALYs, see Chapter 8. Absolute shortfall is thus expressed in QALYs lost.

In the following, we specify the principles for calculating absolute shortfall. There is a differentiation between treatment interventions and preventive measures. Comorbidity is discussed separately.

11.1 Types of economic analysis

Absolute shortfall must, in general, be calculated when cost-utility analyses are used.

If the analysis submitted is in the form of a cost-minimisation analysis, it is not necessary to calculate the absolute shortfall.

11.2 Treatment interventions

The calculation of absolute shortfall is performed in stages

1. Define the mean age, at start of treatment, of the Norwegian patient group for whom the intervention is intended. We refer to the age as A.
2. Estimate the number of remaining healthy life years for an average person from the general population with the age A. We refer to this as $QALYs_A$.
3. Calculate the prognosis for the relevant Norwegian patient group. The prognosis is the average number of remaining healthy life years for the patient group with the current standard treatment. We refer to this as P_A.
4. The absolute shortfall is the difference between the estimate in point 2 and the projection in point 3:
   $AS = QALYs_A - P_A$

In calculations, the undiscounted values for $QALYs_A$ and P_A must be used.

A detailed description of this approach, with examples, can be found in Appendix 4.

11.3 Interventions which treats several diseases/conditions

The principle for quantifying severity, when an intervention has a treatment effect on several diseases in the patient group, corresponds to the principle for calculations when a pharmaceutical why pharmaceutical has a preventive effect on several diseases. This is described in Chapter 11.5.2 below.

11.4 Calibrating two data sources

In calculations, data for the prognosis for the patient group and data for the expected number of remaining QALYs for the average population will usually come from different sources.

HSUV (QALY weights) in the prognosis calculation will come from clinical studies of the pharmaceutical being evaluated, or from other studies where the quality of life for the disease/condition has been measured. The HSUV for the average population will, as a rule, have come from other sources (Appendix 4). This means that HSUV can come from different populations and may have been measured using different instruments and tariffs.

In some cases, the HSUV for symptom-free conditions in the health economic analyses, (which form the basis for the prognosis calculations), are higher than the HSUV for the average population (used in calculating the expected number of remaining QALYs). If so, this should usually be corrected for by calibration.
An example of calibration is shown in Appendix 4.

11.5 Preventive measures

Calculating the severity must be linked to the disease that is being prevented, for the subgroup who would have developed the disease in the absence of the new intervention, measured from the time the disease would be expected to occur in the average patient. Examples of calculating the degree of severity for preventive measures are shown in Appendix 4.

11.5.1 Case 1 – Only one disease/condition is prevented

To calculate the absolute shortfall for conditions, the following must be taken into account:
- Not all the individuals in the group will actually be affected by the disease/event
- There is a time difference between when the prevention starts and when the disease/event may occur

Procedure:
1. First consider which of the individuals/patients in the group must be included in the calculations. Severity is only calculated for that part of the group which is expected to be affected by the disease the preventive measure is aimed at in the current situation. The current situation includes any preventive measures already being carried out (the current standard prevention) but does not include the new preventive measure which is to be evaluated.
2. Then calculate the average prognosis and absolute shortfall for the subgroup expected to get the disease with the current standard prevention and expected standard treatment of the disease from the time the disease occurs.

11.5.2 Case 2. Several diseases/conditions are prevented

Calculation and weighting of severity can be performed in several stages:

1. Calculate the absolute shortfall for each of the diseases/conditions for the relevant population with current preventive practice (the comparator in the health economic analysis). This is explained in "Case 1 – Only one disease/condition is prevented".

2. After this, calculate a weighted absolute shortfall for the diseases/conditions. Example: for prevention of two diseases/conditions, the disease that is most important for the estimated benefit (gained QALYs) of the new preventive measure must be weighted heaviest in the calculation of the weighted absolute shortfall. Absolute shortfall for disease A must be weighted at 90 % in the weighted absolute shortfall if 90 % of the benefit, measured in QALYs, can be attributed to prevention of disease A.

The justification for this type of weighting is given in Appendix 4. This weighting principle is also relevant if the intervention can both treat and prevent disease. An example would be an MRI machine used to screening and monitoring the development of cancer.

11.6 Comorbidity and harm

For interventions directed towards one main condition, it is the overall degree of severity of the main condition and the issues resulting from the main condition that are to be assessed and calculated.

For interventions aimed at symptoms *resulting* from the main condition (and do not affect the main condition), it is the degree of severity for the resultant symptoms alone, and not of the main condition, that must be evaluated

and calculated. For example, if a disease causes pain, the pain medication should be assigned a degree of severity that corresponds to the absolute shortfall for the pain alone, independent of the main condition.

For interventions aimed at treating adverse reactions *resulting from* the treatment of the main condition, i.e., reflecting adverse reactions, not comorbidity. It is the degree of severity of the adverse reaction, and not of the main condition – that must be evaluated and calculated. For example: If the treatment for a disease leads to nausea, then the medicine for nausea should be assigned a degree of severity which corresponds to the absolute shortfall for nausea alone, so the degree of severity is the same regardless of who is affected.

For interventions aimed at symptoms that *are not related to* the main condition, it is the degree of severity of the symptoms – and not the main condition – that must be evaluated and calculated.

12. Uncertainty

You must explore and discuss uncertainty in the health economic analyses.
In this section, we describe different sources of uncertainty in health economic analyses and ways of dealing with different types of uncertainty.

12.1 Terminology used to describe uncertainty

It is useful to differentiate between the following in relation to uncertainty in health economic models (32, 33).

- **Stochastic uncertainty**: This means that patients with the same risk may experience different outcomes of the disease or intervention due to random variability
- **Parameter uncertainty**: This relates to uncertainty about the "true value" of a parameter. This applies to variables estimated from sample data or are based on other data/sources. This will typically be costs, HSUVs, treatment effects, and the probability of events. Parameter uncertainty can reflect sampling data, contradictory studies, lack of internal or external validity, limited generalisability, or lack of data
- **Model uncertainty or structural uncertainty**: This relates to uncertainty about assumptions and choices made in the construction of the model. Examples are the relationships between variables in the model, the chosen functional form for modelling the time-to-event data, extrapolation of treatment effect, and the choice of which health states are included in the model
- **Heterogeneity**: The effect of patient heterogeneity (variation in patient characteristics) on the model's results is not related to uncertainty, and is best analysed by sub-group analysis
- **Methodological uncertainty**[14]: This will typically be about areas within health economics where there is methodological disagreement. An example is the choice of instrument to measure health-related quality of life

12.2 Dealing with uncertainty in the analyses

Uncertainty in the health economic analysis must be explored and presented through sensitivity analyses. This should be done using both deterministic and probabilistic sensitivity analyses, described in more detail below. Not all uncertainty can be reflected this way. It can, for example, be very difficult to analyse structural uncertainty and generalisability fully in sensitivity analyses.

The impact of uncertainty on the outcomes of the analysis must be discussed in order to highlight what factors drive the uncertainty, whether the uncertainty can be reduced, whether additional data can be expected, whether any bias is present, and how the results of the analysis are affected by changes in the parameters or assumptions.

12.2.1 Deterministic sensitivity analysis

In deterministic sensitivity analyses selected variables are changed to explore how sensitive the model outcomes are to these changes. This type of analysis is performed as one-way, two-way or multiway sensitivity analyses and in scenario analyses.

We recommend analysing methodological and structural uncertainty, as well as uncertainty linked to generalizability, by using deterministic sensitivity analyses, as far as possible.

Deterministic sensitivity analyses alone will not be able to show all the uncertainty, and should be supplemented by probabilistic analyses and discussion. One-way sensitivity analyses cannot capture correlation between variables and the impact of joint parameter uncertainty on model outcomes. For two-way and multiway

[14] Methodological uncertainty is reduced by the recommendation of a preferred method/approach, for example our recommendation to the use of one quality of life instrument (EQ-5D) and a set discount rate (4 %).

sensitivity analyses, the number of possible parameter combinations can easily become insurmountable, making it difficult for decision makers to easily judge how likely different outcomes are. Deterministic sensitivity analyses alone are therefore not sufficient to reflect the impact of parameter uncertainty.

One-way sensitivity analyses
In a one-way sensitivity analysis, the values are varied individually. For example, parameter values can be varied within their corresponding 95% confidence interval or relevant credibility interval.

Perform a one-way sensitivity analysis for each parameter in the model and present all results in a table. Present one-way sensitivity results for the most important parameters in both tables and in a tornado diagram. Time horizon, the device price or procedure cost and the comparator(s), HSUVs, parametric functions for time-to-event data as well as effect parameters must always be included.

Two-way and multiway sensitivity analyses
In two-way and multiway sensitivity analyses the values of two or multiple parameters, respectively, are varied simultaneously.

Scenario analyses
A scenario analysis is used to evaluate the impact of alternative values for selected sets of parameters on the model outcomes. Selection is often made so that it represents, for example, a “base case” and “worst case” analysis, or alternative plausible scenarios.

12.2.2 Probabilistic sensitivity analysis (PSA)

In a PSA a range of chosen variables are defined as stochastic variables, with an associated probability distribution. Justify the choice of variables included in the PSA and their probability distributions. The probability distribution of the variable and its most important moments (usually the expected value and standard error), should be based on empirical data. If there is a lack of empirical data, a plausible probability distribution must be chosen for the variable. Each type of variable, costs for example, will usually only have a few types of probability distributions that are relevant for use in PSAs (see for example Drummond 2015 (32)).

PSAs should be used to capture the impact of joint parameter uncertainty. In principle, model uncertainty can also be explored in the PSA, for example by assigning probability weights and distributions to alternative assumptions. This is recommended if it is possible and appropriate.

The results of the PSA must be presented as a scatter plot of the simulated Incremental cost-effectiveness ratio (ICERs) and as cost-effectiveness acceptability curves (CEACs).

12.2.3 “Value of Information” analysis

Value of Information analysis (VoI) can be conducted using results from the probabilistic sensitivity analysis. Such an analysis can include estimation of the Expected Value of Perfect Information (EVPI), which combines the probability of making a wrong decision with the consequential losses of that decision. The EVPI should be calculated when a PSA has been performed and there is decision uncertainty (when the probability that the new treatment is cost effective is less than 100 %, but higher than 0 %, for a range of common willingness to pay thresholds). The EVPI should be presented in a graph for a range of willingness to pay thresholds.

Further analyses can be requested to investigate whether the decision to introduce the intervention should be postponed, either to obtain, or in anticipation of, further evidence. This can include estimation of the Expected Value of Partial Perfect Information (EVPPI) to identify key parameters. See the relevant literature for more information about the method and presentation of VoI analyses (33-35).

13. Budget impact

In general, the budget impact analysis covers the financial, rather than the economic, costs of the intervention versus its comparator over a five-year period at the national level.

You should calculate costs using two scenarios – one in which the new intervention is adopted by the specialist health services, and another in which it is not. The budget impact is the difference between the two scenarios in each of the first five years after adoption.

The budget impact has three components:

- The costs of the device or diagnostic equipment for the specialist healthcare services
- Other related costs borne by the specialist healthcare services
- Other related costs in the health and care services (outside of the specialist health services)

You may, if you provide reasonable justification (e.g. organizational complexity), present the budget impact using other methods than those recommended below (contact NIPH for guidance).

The time horizon for the budget analyses of pharmaceuticals is five years. This is because it is assumed that the broad usage of new pharmaceuticals is well established after five years. For other technologies, the time horizon may vary depending on the economic lifespan and/or depreciation of the technology. Provide a justification if the time horizon deviaties from five years.

Calculation of the additional costs shall be based on the following factors:

1. Costs incurred by the specialist health service during the calculation/analysis period.
2. The estimated market share of the new technology, in relation to the patient group the technology targets, in each of the relevant years after the decision to use the technology is made.
3. Deductions of: costs of competing technologies that will be completely or partially replaced by the new technology, any increases in patient payments and increments in user fees during outpatient treatment.
4. Other costs related to the technology assessment (change in bed-days, commodity costs, personnel costs, nursing costs, depreciation, travel expenses covered by the specialist health care service, administrative expenses, etc.) should only be included if there are significant differences between the competing technologies and/or if the differences constitute a large proportion of the additional costs.

Budget calculations/analysis should cover both the new technology and competing technology(ies) if the extent of use is affected by the possible introduction of the new technology. This will in turn make it possible to calculate a total budget impact. The budget impact is the difference between the two scenarios in each of the relevant years of the analysis (see tables below). Year 1 is the first full calendar year after a decision is made about introducing the new technology into the specialist health care service.

The budget impact calculations must show the following:

1 What proportion of the total additional costs is the result of an increase in patient numbers and what proportion is due to the transition to a more expensive technology.
2 The basis for key assumptions in the calculations.

Additionally, the following calculations may apply in special cases:

1 Subgroup analyses such as in cases where it is prudent to prioritise giving the new technology to only a subset of the total population.
2 Analyses with added costs/impact on other patient groups not targeted by the new technology but whom none the less use the technology.

3 Sensitivity analyses where key assumptions and data are tested in order to check to what extent results and estimates used are sensitive to changes. This is particularly relevant if critical assumptions in the analyses are very uncertain.

The budget impact calculations will also depend on whether the proposed intervention involves smaller medical devices that are assigned to individual patients, or larger medical equipment intended for use by many patients. The respective calculations are described in the sections below.

13.1 Implantable/wearable and other non-shared medical devices

First, estimate the market share of the intervention and comparator(s), in terms of number of patients using the respective technologies over the next five years. The costs of the intervention and comparator are then applied to the respective numbers to obtain an estimate of annual budget impact. Expenditure must not be calculated cumulatively for the first five years.

The cost of a medical device should be based on the expected retail price including VAT. The tables below show how the calculation of budget impact for the healthcare services should be presented. Patients' co-payment must not be included.

Number of patients

Table 1 shows the number of patients expected to be treated with the intervention, and the comparator in the first five years in scenarios with and without the intervention being implemented in the specialist healthcare sector.

Table 1 Number of patients expected to be treated over the next five-year period – if the intervention is implemented

	Year 1	Year 2	Year 3	Year 4	Year 5
Intervention					
Comparator 1					
Comparator 2					

Table 2 Number of patients expected to be treated during the next five-year period – if the intervention is NOT implemented

	Year 1	Year 2	Year 3	Year 4	Year 5
Intervention	0	0	0	0	0
Comparator 1					
Comparator 2					

Expenditure per patient
Calculate the expenditure per patient per year for the different treatments. The estimates should be consistent with the corresponding calculations in the CUA. If you choose to use the health economic model for these calculations, note that they should be inclusive of VAT and the relevant costs, as set out in the various sections in this chapter, but without discounting.

Budget impact
Multiply the expenditure per patient per year by the number of patients per year for the intervention and comparator(s). Total these costs for each year and enter this figure into the table below. In the bottom row of the table, present the estimated budget impact of adopting the intervention.

Table 3 Expected budget impact of adopting the intervention for the relevant indication.

	Year 1	Year 2	Year 3	Year 4	Year 5
The intervention is adopted	X1	X2	X3	X4	X5
Minus: The intervention is not adopted	Y1	Y2	Y3	Y4	Y5
= budget impact	X1 - Y1	X2 - Y2	X3 - Y3	X4 - Y4	X5 - Y5

The budget impact is the difference between the two scenarios in each of the first five years.

13.2 Fixed and other shared medical devices

Medical devices or equipment which are used by several patients over a period of time (located in a hospital or in a mobile unit) should be included in the budget impact analysis in a manner similar to that described above for small devices and implants.

As in the above case, first estimate the market share of the intervention and comparator(s). However, rather than quantify the number of patients, estimate the number of shared devices for both intervention and comparator(s) that is likely be in use in the specialist care sector each year in the next five years. If a comparator is not a shared device, refer to the section below for details. Do this for scenarios with and without the intervention. Again, as in the above case, the costs of the intervention and comparator are applied to the respective numbers to obtain an estimate of annual budget impact. If the comparator is a drug or a non-shared device, the method in 13.1 should be used to estimate the annual costs for this comparator.

The cost of a shared medical device should be based on the expected retail price including VAT (if expenditure must be calculated without discounting). The following tables show how the calculation of budget impact for the specialist healthcare services should be presented

Number of shared devices

Tables 4 and 5 shows the number of devices expected to be in use with the intervention, and the comparator in the first five years in scenarios with and without the intervention being implemented in the specialist healthcare sector.

Table 4 Number of shared devices expected to be in use during the next five-year period – if the intervention is implemented

	Year 1	Year 2	Year 3	Year 4	Year 5
Intervention					
Comparator 1					
Comparator 2					

Table 5 Number of shared devices expected to be in use during the next five-year period – if intervention is NOT implemented

	Year 1	Year 2	Year 3	Year 4	Year 5
Intervention	0	0	0	0	0
Comparator 1					
Comparator 2					

Expenditure per device

Calculate the annual expenditure per device for the different devices. The estimates should be consistent with the corresponding calculations in the CUA. If you choose to use the health economic model for these calculations, note that they should be inclusive of VAT and the relevant costs, as set out in the various sections in this chapter, but without discounting.

Budget impact

Multiply the expenditure per device per year by the number of devices per year for the intervention and comparator(s). Total these costs for each year and enter this figure into the table below. In the bottom row of the table, present the estimated budget impact of adopting the intervention.

Table 6 Expected budget impact of adopting the intervention for the relevant indication.

	Year 1	Year 2	Year 3	Year 4	Year 5
The intervention is adopted	X1	X2	X3	X4	X5
Minus: The intervention is not adopted	Y1	Y2	Y3	Y4	Y5
= budget impact	X1 - Y1	X2 - Y2	X3 -Y3	X4 - Y4	X5 -Y5

14. Acknowledgements

The Norwegian Medicines Agency has published guidelines for pharmacoeconomic analyses since 2002. Following a mandate from the Ministry of Health and Care Services (HOD), the Norwegian Medicines Agency updated its guidelines in cooperation with a multi-agency working group consisting of representatives from the Norwegian Institute of Public Health, the Norwegian Directorate of Health, the four regional health authorities and the Norwegian Hospital Procurement Trust, Division Pharmaceuticals (LIS, now a part of Sykehusinnkjøp HF). The group was chaired by Morten Aaserud.

The updated version of these Agency's guidelines came into force on January 1st 2018 and an English version was published in April the same year. The guidelines were further updated in May 2020. This version of the guidelines has been adapted by NIPH to specifically address the economic evaluation of medical devices and diagnostic interventions.

Staff at the Norwegian Medicines Agency provided input regarding the regulatory context in Chapter 1. At NIPH, the following people have been involved in developing these guidelines: Espen Movik, Jan Abel Olsen, Lene Kristine Juvet and Øyvind Melien (draft sections and edits). Arna Desser (language and other edits) Julia Bidonde Torre, Trygve Ottersen, Anders Huitfeldt, Elisabet Hafstad, Gerd Flodgren, Tor Rosness and Ulrikke Højslev Lund (edits and comments). Chapter 4 was revised in February 2021 by Elisabet Hafstad, Ingrid Harboe og Gunn Eva Næss.

The guidelines were subject to external review between September 15th and December 15th 2020. The following institutions submitted comments:

Helse Midt RHA
Helse Sør-Øst RHA
Melanor
The Directorate of Health
The Norwegian Medical Association
The Norwegian Medicines Agency
The Norwegian Radiation Protection Authority
The Secretariat for the "New Methods" system

NIPH wishes to thank all external parties for their comments and it is worth stressing that the responsibility for these guidelines remains with NIPH.

The final version of the guideline was reviewed and edited by Kjetil Gundro Brurberg, Jan Marcus Sverre, Liv Merete Reinar and Anna Lien Espeland.

15. References

1. Helse- og omsorgsdepartementet. Meld. St. 34 Verdier i pasientens helsetjeneste. 2015-2016.
2. Kirkehei I, Ormstad SS. Litteratursøk. Norsk epidemiologi. 2013;23(2).
3. Dias S, Welton NJ, Sutton AJ, Ades A. NICE DSU Technical Support Document 1: Introduction to evidence synthesis for decision making 2011 [1-24]. Available from: http://www.nicedsu.org.uk.
4. Egger M, Juni P, Bartlett C, Holenstein F, Sterne J. How important are comprehensive literature searches and the assessment of trial quality in systematic reviews?: Empirical study. NCCHTA; 2003.
5. Papaioannou D, Brazier, J.E., Paisley, S. NICE DSU Technical Support Document 9: The identification, review and synthesis of health state utility values from the literature. 2011 [Available from: http://www.nicedsu.org.uk.
6. Tarricone R, Torbica A, Drummond M. Key Recommendations from the MedtecHTA Project. Health economics. 2017;26 Suppl 1:145-52.
7. SSB. Dødelighetstabeller, 2016 [Available from: https://www.ssb.no/befolkning/statistikker/dode/aar/2017-03-09?fane=tabell&sort=nummer&tabell=297661.
8. Rothery C, Claxton K, Palmer S, Epstein D, Tarricone R, Sculpher M. Characterising Uncertainty in the Assessment of Medical Devices and Determining Future Research Needs. Health economics. 2017;26 Suppl 1:109-23.
9. Sorenson C, Tarricone R, Siebert M, Drummond M. Applying health economics for policy decision making: do devices differ from drugs? Europace : European pacing, arrhythmias, and cardiac electrophysiology : journal of the working groups on cardiac pacing, arrhythmias, and cardiac cellular electrophysiology of the European Society of Cardiology. 2011;13 Suppl 2:ii54-8.
10. CADTH. Guidelines for the Economic Evaluation of Health Technologies: Canada guidelines for the Economic Evaluation of Health Technologies: Canada. Canadian Agency for Drugs and Technology in Health; 2017.
11. Chapman AM, Taylor CA, Girling AJ. Are the UK systems of innovation and evaluation of medical devices compatible? The role of NICE's Medical Technologies Evaluation Programme (MTEP). Applied health economics and health policy. 2014;12(4):347-57.
12. Willeme P, Dumont M. Machines that Go 'Ping': Medical Technology and Health Expenditures in OECD Countries. Health economics. 2015;24(8):1027-41.
13. Craig JA, Carr L, Hutton J, Glanville J, Iglesias CP, Sims AJ. A review of the economic tools for assessing new medical devices. Applied health economics and health policy. 2015;13(1):1527.
14. Nation Institute of Health and Care Excellence. Diagnostic Assessment Programme Manual. 2011.
15. Fryback DG, Thornbury JR. The efficacy of diagnostic imaging. Medical decision making. 1991;11(2):88-94.
16. Whiting PF, Rutjes AW, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. Ann Intern Med 2011;155(8):529-536.
17. Gold MR, Siegel JE, Russell LB, Weinstein MC. Cost effectiveness in health and medicine: Oxford University Press, Oxford; 1996.
18. Sanders GD, Neumann PJ, Basu A, Brock DW, Feeny D, Krahn M, et al. Recommendations for conduct, methodological practices, and reporting of cost-effectiveness analyses: second panel on cost-effectiveness in health and medicine. Jama. 2016;316(10):1093-103.
19. EuroQol. EQ-5D [updated 16.05.2017]. Available from: http://www.euroqol.org/.
20. Brazier JE, Longworth, L. NICE DSU Technical Support Document 8: An introduction to the measurement and valuation of health for NICE submissions. 2011 [Available from: http://www.nicedsu.org.uk.
21. van Hout B, Janssen MF, Feng YS, Kohlmann T, Busschbach J, Golicki D, et al. Interim scoring for the EQ-5D-5L: mapping the EQ-5D-5L to EQ-5D-3L value sets. Value in health : the journal of the International Society for Pharmacoeconomics and Outcomes Research. 2012;15(5):708-15.

22. National Institute for Health and Care Excellence. Position statement on the use of the EQ-5D-5L valuation set 2017 [Available from: https://www.nice.org.uk/Media/Default/About/what-we-do/NICE-guidance/NICE-technology-appraisal-guidance/eq5d5l_nice_position_statement.pdf.
23. Michel YA, Augestad LA, Barra M, Rand K. A Norwegian 15D value algorithm: proposing a new procedure to estimate 15D value algorithms. Quality of Life Research. 2019;28(5):1129-43.
24. Dolan P. Modeling valuations for EuroQol health states. Medical care. 1997;35(11):1095-108
25. Longworth L, Rowen D. NICE DSU Techinical Support Document 10: The Use of mapping methods to estimate health stse utility values 2011 [Available from: http://www.nicedsu.org.uk.
26. Ara R, Wailoo, A.J. NICE DSU Technical Support Document 12: The use of health state utility values in decision models 2011 [Available from: http://www.nicedsu.org.uk.
27. Direktoratet for økonomistyring. Veileder i samfunnsøkonomiske analyser. 2014.
28. Legeforeningen. Normaltariffen 2017 [Available from: https://www.legeforeningen.no/jus-og-arbeidsliv/avtaler/Normaltariffen/.
29. Finansdepartementet. Rundskriv R-109/14 - Prinsipper og krav ved utarbeidelse av samfunnsøkonomiske analyser mv. 2014.
30. Finansdepartementet. NOU 2012:16 Samfunnsøkonomiske analyser. 2012.
31. Drummond M, Griffin A, Tarricone R. Economic evaluation for devices and drugs--same or different? Value in health : the journal of the International Society for Pharmacoeconomics and Outcomes Research. 2009;12(4):402-4.
32. Drummond MF, Sculpher MJ, Claxton K, Stoddart GL, Torrance GW. Methods for the economic evaluation of healthcare programmes: Oxford university press; 2015.
33. Briggs AH, Weinstein MC, Fenwick EA, Karnon J, Sculpher MJ, Paltiel AD, et al. Model parameter estimation and uncertainty: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force-6. Value in Health. 2012;15(6):835-42.
34. Claxton K, Sculpher M, Drummond M. A rational framework for decision making by the National Institute For Clinical Excellence (NICE). Lancet. 2002;360(9334):711-5.
35. Briggs AH, Claxton K, Sculpher MJ. Decision modelling for health economic evaluation: Handbooks in Health Economic E; 2006.
36. Kruschke J. Doing Bayesian data analysis: A tutorial with R, JAGS, and Stan: Academic Press; 2014.
37. Jansen JP, Trikalinos T, Cappelleri JC, Daw J, Andes S, Eldessouki R, et al. Indirect treatment comparison/network meta-analysis study questionnaire to assess relevance and credibility to inform healthcare decision making: an ISPOR-AMCP-NPC Good Practice Task Force report. Value in Health. 2014;17(2):157-73.
38. Signorovitch JE, Sikirica V, Erder MH, Xie J, Lu M, Hodgkins PS, et al. Matching-adjusted indirect comparisons: a new tool for timely comparative effectiveness research. Value in Health. 2012;15(6):940-7.
39. Dias S, Welton NJ, Sutton AJ, Ades A. NICE DSU technical support document 2: a generalised linear modelling framework for pairwise and network meta-analysis of randomised controlled trials 2011 [Available from: http://www.nicedsu.org.uk].
40. Dias S, Sutton AJ, Welton NJ, Ades A. NICE DSU Technical Support Document 3: Heterogeneity: subgroups, meta-regression, bias and bias-adjustment. Sheffield: Decision Support Unit ScHARR. 2011:1-24.
41. Dias S, Welton NJ, Sutton AJ, Caldwell DM, Lu G, Ades A. NICE DSU Technical Support Document 4: Inconsistency in networks of evidence based on randomised controlled trials 2011 [Available from: http://www.nicedsu.org.uk.
42. Dias S, Welton NJ, Sutton AJ, Ades A. NICE DSU Technical Support Document 5: Evidence synthesis in the baseline natural history model: Citeseer; 2011 [Available from: http://www.nicedsu.org.uk.
43. Ades A, Caldwell DM, Reken S, Welton NJ, Sutton AJ, Dias S. NICE DSU Technical Support Document 7: Evidence synthesis of treatment efficacy in decision making: a reviewer's checklist 2012 [Available from: http://www.nicedsu.org.uk.

44. Phillippo DM, Ades A, Dias S, Palmer S, Abrams KR, Welton NJ. NICE DSU Technical Support Document 18: Methods for population-adjusted indirect comparisons in submissions to NICE 2016 [Available from: http://www.nicedsu.org.uk.
45. Bradburn M, Clark T, Love S, Altman D. Survival analysis Part III: multivariate data analysis-choosing a model and assessing its adequacy and fit. The British Journal of Cancer. 2003;89(4):605.
46. Bradburn MJ, Clark TG, Love S, Altman D. Survival analysis part II: multivariate data analysis-an introduction to concepts and methods. The British Journal of Cancer. 2003;89(3):431.
47. Clark T, Bradburn M, Love S, Altman D. Survival analysis part IV: further concepts and methods in survival analysis. The British Journal of Cancer. 2003;89(5):781.
48. Clark T, Bradburn M, Love S, Altman D. Survival analysis part I: basic concepts and first analyses. The British Journal of Cancer. 2003;89(2):232.
49. Dohoo I, Martin W, Stryhn H. Methods in Epidemiologic Research. Charlottetown, Prince Edward Island: AVC. Inc; 2012.
50. Latimer NR. Survival analysis for economic evaluations alongside clinical trials—extrapolation with patient-level data: inconsistencies, limitations, and a practical guide. Medical Decision Making. 2013;33(6):743-54.
51. Latimer N. NICE DSU technical support document 14: survival analysis for economic evaluations alongside clinical trials-extrapolation with patient-level data 2011 [Available from: http://www.nicedsu.org.uk.
52. Bagust A, Beale S. Survival Analysis and Extrapolation Modeling of Time-to-Event Clinical Trial Data for Economic Evaluation An Alternative Approach. Medical Decision Making. 2014;34(3):343-51.
53. Latimer NR, Abrams KR. NICE DSU Technical Support Document 16: Adjusting survival time estimates in the presence of treatment switching 2014 [b12]. Available from: http://www.nicedsu.org.uk.
54. Sun S, Irestig R, Burstrom B, Beijer U, Burstrom K. Health-related quality of life (EQ-5D) among homeless persons compared to a general population sample in Stockholm County, 2006. Scand J Public Health. 2012;40(2):115-25.
55. Burström K, Johannesson M, Diderichsen F. Swedish population health-related quality of life results using the EQ-5D. Quality of life research. 2001;10(7):621-35.
56. SSB. Dødelighetstabeller, 2017 [Available from: https://www.ssb.no/befolkning/statistikker/dode.

Appendix 1. Documentation of relative efficacy in indirect comparisons

1.1 General

The research question /scope must be clearly formulated.
Justify why it is necessary to use an indirect comparison.
Before an indirect comparison is carried out the known effect modifiers and prognostic factors must be described as fully as possible from previous knowledge.

1.2 Literature search

Carry out a full systematic literature search. Describe the literature search in detail, both for the relevant intervention and for the chosen comparator(s). For literature searches intended to support the documentation of relative efficacy, PICO[17] must be taken into account.

All relevant data from the literature search must be described according to the template for submission of documentation.

1.3 Assumptions

Describe which assumptions form the basis for the indirect comparison and evaluate whether the assumptions have been satisfied. Describe how differences, heterogeneity and (lack of) consistency have been dealt with.

1.4 Statistical methods

Justify the choice of the statistical method. Use appropriate statistical methods and describe these in detail. Present all relevant aspects of the statistical analyses in a transparent way. This applies, among other things, to how the adjusted indirect comparisons are carried out, how multi-armed studies are dealt with, use of random effects or fixed effect models, technical details, programming codes, how outliers and particularly influential studies/datasets are dealt with, and sensitivity analyses.

The choice of a fixed effect or random effects model must be based on the extent to which the studies have been carried out with sufficient similarity. Meta-analyses include studies which are clinically and methodologically diverse, and heterogeneity in the study effects is to be expected. For this reason, the random effects model is usually preferable.

If Bayesian statistics is used, then the following must also be described as a minimum (36):

[17] PICO: Patient, Intervention, Comparator, Outcome

Choice of priors: If informative priors are used, a sensitivity analysis with non-informative priors should be presented as well. If informative priors are used, there must be documentation showing which assumptions and which data these informative priors are based on.

Calculation of credible intervals (CrI): Describe the methods for calculating and defining credible intervals (CrI).

Definition and discussion of region of practical equivalence (ROPE): Describe the criteria and the information sources the ROPE is based on.

Enclose a graphical presentation of the relevant posterior distributions with the chosen prior for the most relevant outcome measures.

If MAIC or STC is used, the following must be done as a minimum:

Describe in detail the population the STA is relevant to and describe the extent to which the adjusted population (MAIC or STC) deviates from this.

Describe and discuss on the basis of clincial evidence, whether the studies being compared overlap sufficiently in terms of study design, inclusion criteria, patient characteristics, definition of outcome measures and reporting of data.

Account for those effect modifiers (for MAIC and STC) and prognostic (for MAIC) factors which are not balanced in the studies being compared, and assess the extent to which there is enough information in the studies to correct completely for all these factors. Account for covariates which cannot be taken into account in the analysis. Discuss the risk of unmeasured confounding factors which could affect the analysis.

In a MAIC, patients from the study with individual patient data (IPD) are assigned weights[18], so that the weighted average patient characteristics equal to what is reported from the studies without IPD (published, aggregated data). Effective sample size (*n eff*) should be reported for the "balanced" population, ie, how much of the information from the index population contributes to the adjusted outcome measures in the indirect comparison.

Based on clinical evidence, describe and justify the possible consequences of a variable being excluded from the weighting.

For a more detailed description of how to carry out an MAIC or STC we recommend Jansen et al and Signorovitch et al (37, 38) as well as DSU from NICE (6, 39-44).

[18] Patients in a treatment arm (study with IPD) are weighted with inverse odds in order to be in the relevant treatment group versus the other treatment group (study with only published aggregated data).

Appendix 2. Use of time to event data in health economic analyses

2.1 Introduction

Examples of time to event data (also known as survival or event history) are time to progression in cancer, ie, progression-free survival (PFS), time to death, ie, overall survival (OS), or time to a cardiovascular event or treatment discontinuation. The randomisation time point is usually the starting point in time to event analyses.

In health economic analyses it is normal to use a form of parametrisation with extrapolation of the clinical time to event data beyond the actual study period. Below we specify how parametrisation and extrapolation of survival data must be carried out for health economic analyses sent to NIPH for evaluation. This applies regardless of whether the relative efficacy has been obtained by direct or indirect comparisons. For choice of time horizon, see Chapter 10.3 Time horizon.

2.2 Parametrisation of data from clinical studies

Data extrapolation beyond the study follow-up period is common in health economic analyses. In such analyses a type of a parametric function is often used. Parametric functions are based on an assumption that the underlying risk of the event (baseline risk) follows a given distribution, in contrast to non-parametric (eg, Kaplan-Meier) or semi-parametric (eg, Cox model) functions.

Different parametric functions can give very different estimates.

The choice of a parametric function is based on statistical analyses of best mathematical fit, in combination with biological criteria related to knowledge of how the risk of event is expected to develop for the current condition/disease and endpoint. For example, some conditions will have a high risk of an event initially, but will then decrease (biphasically), while for others the risk of event will increase or decrease in monotonously.

Parametrisation must be based on the actual data from the clinical studies, thus highlighting the direct effect of the treatment under consideration.

Statistical tests and graphic evaluations must be carried out systematically to allow the choice of the most accurate parametric function (45-52).

For a given function to fit satisfactorily, the following two criteria must be fulfilled:

1. The function must fit well with the observed efficacy data from the study or studies
2. The extrapolated part is clinically and biologically plausible

Justify in detail the choice of a function in light of the two criteria above. Functions which do not fulfill both these criteria are probably not suitable.

2.2.1 Curve fitting to observed study data

By curve fitting, we mean how well suited a parametric function is to the clinical data from the study or studies (usually Kaplan-Meier data). For optimal evaluation of the curve fit an extensive description and analysis of any assumptions and properties regarding the parametric functions and relevant clinical data should be submitted. In order to document the adjustment(s) to the observed study data all of the points in the list below must be included as a minimum:

- The following parametric functions should, as a minimum, be included in the health economic model: exponential, Weibull, Gompertz, gamma, log-logistic, log-normal and Generalised gamma distributions
- Statistical tests and graphical presentation for testing of proportional hazard (PH), accelerated failure time modell (AFT) and for assessing the fit of standard parametric functions (53):
    - log-kumulativ hazardplot for PH: $\log(-\log(S(t)))$ vs. $\log(t)$ with linear trendlines for the intervention and comparator
    - plot based on Schoenfeld residuals
    - Quantile-Quantile-plot for AFT $t_0(p)$ vs $t_1(p)$ with a linear trendline, using the percentiles of the inverse survival functions for the intervention and comparator:

      $t_0(p) = S_0^{-1}\left(\frac{100-p}{100}\right)$, $t_1(p) = S_1^{-1}\left(\frac{100-p}{100}\right)$
    - $\log(S(t)/(1-S(t)))$ vs. $\log(t)$ with linear trendlines for the intervention and comparator
    - inverse.normal$(1-S(t))$ vs. $\log(t)$ with linear trendlines for the intervention and comparator

- If neither PH nor AFT appears suitable, standard parametric models fitted to each treatment arm independently should be considered before other, more flexible functions are considered, such as a piecewise function, Royston-Palmer models, spline models
- smoothed and unsmoothed hazard plots for the observed data from the clinical study per treatment arm (54, 55)
- smoothed hazard plots for the observed data from the clinical study with the hazard function of all the standard parametric functions plotted in the same figure, per treatment arm (54, 55)
- Akaike's Information Criteria (AIC), Bayesian Information Criteria (BIC) and/or other suitable tests for those functions which are relevant on the basis of the criteria described above, per treatment arm.
- Graphical presentation of time to event data curves, where both Kaplan-Meier (KM) data and the parametric distribution is shown in the same figure. Similar graphical presentation should also be included in the health economic model (in the spreadsheet)
- In some cases, curves with KM data for the first part of the study period can be appropriate, and then a parametric tail which shows the extrapolation beyond this point (transition point). point must be evaluated in the individual case. As a minimum requirement an analysismust be presented where the tail is set at the time point where 50 % of the included population in each treatment arm is still "at risk".
- Parametrization of survival data should be conducted in a transparent way that allows the analysis to be reproduced.

2.2.2 Plausibility of the extrapolated part of the curve

The plausibility of the extrapolated part of the survival curve must be documented and justified biologically and clinically for the patient group in question. External data can be used to evaluate the assumptions made in the extrapolation. External data can include data from another study of a similar patient group or data from a national/international registry with long-term follow-up of a relevant patient group. The patient population must be relevant in terms of patient characteristics, pre-treatment and treatment.

External data can only be seen as indicative. Use of external data requires a balanced discussion of how far any differences in long-term survival between the projected survival curve and the external data source is due to:

- Weaknesses in the chosen parametric function and/or
- Limitations in the external data source

External data will most likely, only be available for the comparator arm, and will therefore be most useful for evaluating the plausibility of projecting the comparator arm. Therefore, the clinically valid assumptions on the duration of treatment effect will be necessary for extrapolating the effect of the intervention. The assumptions can be sourced from clinical expert statements, evaluation of the mechanism of action and biological plausibility. Different assumptions must be tested in the scenario analyses. The significance of each of these factors in assessing plausibility will depend on the current issue and will vary from case to case.

2.2.3 Algorithm and implementation in the health economic model

The figure below shows the algorithm for selection of a parametric model in time to event data analysis.

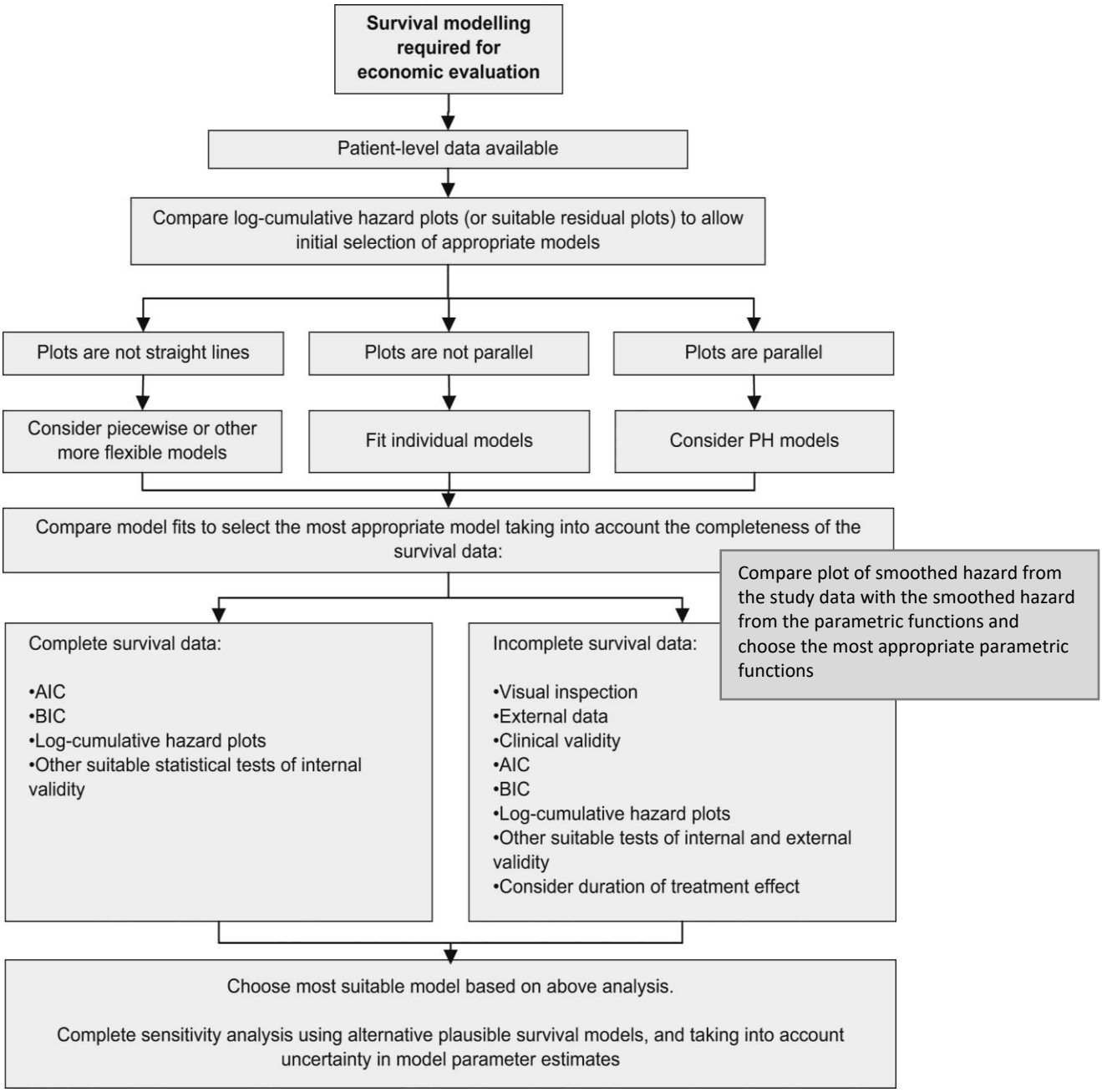

Figure 1: Algorithm for selection of a parametric model. Modified from Latimer 2013 (49)

2.3 Studies where patients can switch to active (new) intervention

For ethical reasons, many controlled clinical studies allow patients in the control arm to switch over to the intervention arm or another active treatment at a given time point, often at progression of disease (treatment switching, crossover). In the submitted documentation it must be explained why this has been done and when the patient changed treatment.

For ethical reasons, treatment switching is relatively common in cancer studies. In such cases the effect estimate for overall survival will be affected by the treatment switching. There are several correction methods which can be used to give an estimate of survival, as if the switch had not taken place. Which method is most suitable, depends on the data in question, and must be evaluated in the individual case. Often a certain method is specified in the study's statistical analysis plan. An intention to treat (ITT) analysis (or the relevant primary analysis if there is no ITT) with an estimate without adjustment for the treatment switch must always be submitted.

Analyses which have been corrected for treatment switch can be submitted. In such a case, justify why the particular correction method has been used and other correction methods have not been used, with a related discussion of the strengths, weaknesses and assumptions of the different methods (56).

Figure 2 shows the procedure both for the choice of correction method and for which considerations form the basis for parametrisation and projection depending on the adjustment method. The intention to treat analysis (ITT) or another primary analysis parameterised and extrapolated as described in appendix 2.2 above.

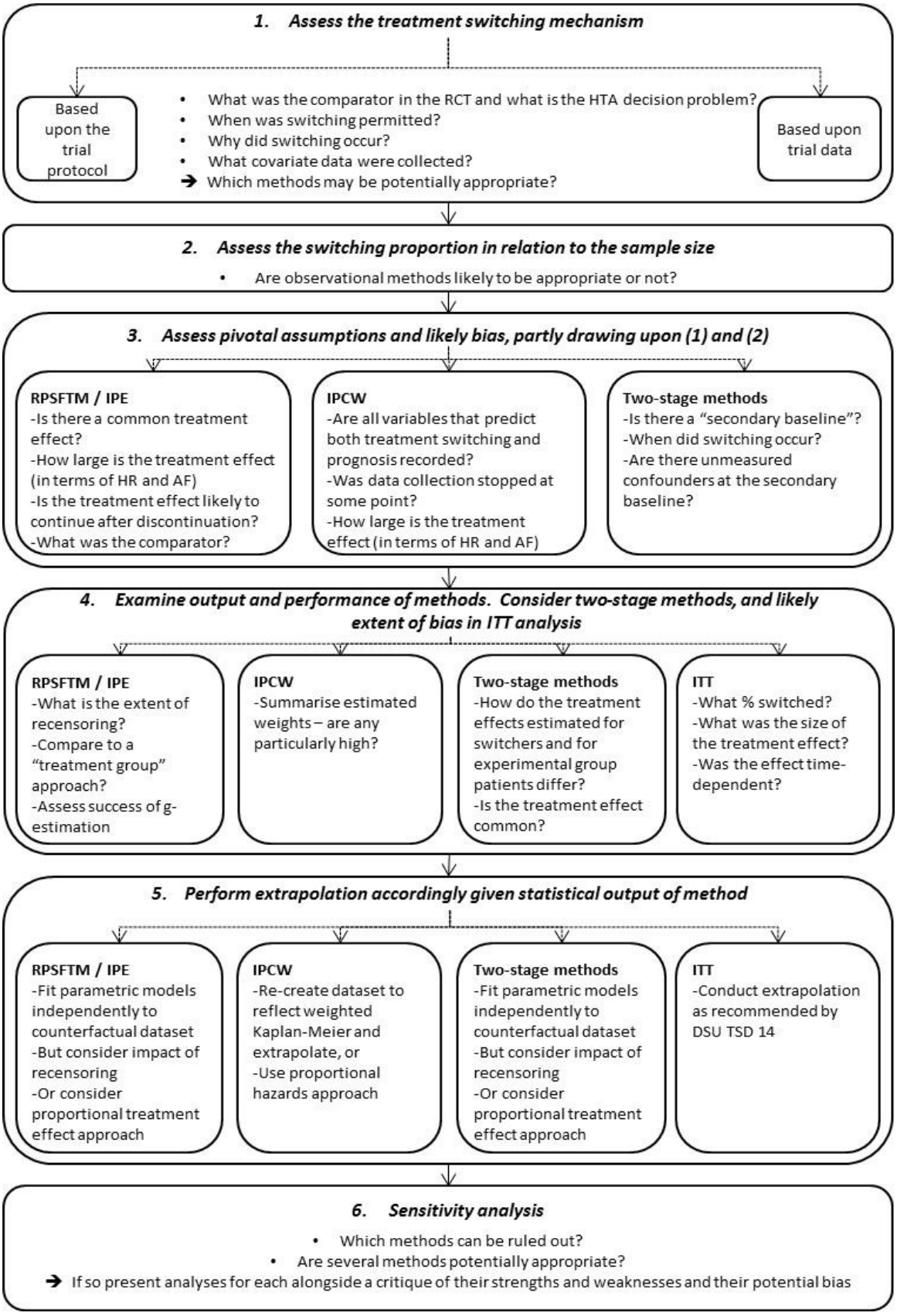

Figure 2: Flow diagram for analyses with treatment switching. From Latimer 2014 (56)

Appendix 3. Quality of life data

3.1 Example of age adjustment of future expected HSUV using the multiplicative method

Suppose that when modelling a chronic disease, we have a HSUV of 0.780 for the "best" health state that can be expected for the patients. The mean patient age is 50 years, and the health economic model is using a lifetime perspective. Without age adjustment, this HSUV will be constant for the proportion of patients who reach the "best" health status for the rest of their lifetime. Age-specific HSUV based on the study by Stavem et al (57), indicate a HSUV of 0.846 at age 50 years in the general population. Similarly, the HSUV at age 81 is 0.730 in the general population. Without age adjustment one would then be using a higher health-related quality of life for a patient population over 81 years than that assumed for the general population, as shown in Figure 3.This can be unrealistic/ unreasonable and is the justification for recommendations about age adjustment in expected future health states.

Age-adjusted HSUV for patients in this example will be a result of a HSUV of 0.780 multiplied by an adjustment index which is set at 1 at the start of the model. In this example, the index is reduced over time on the basis of age-specific HSUV based on Stavem et al (57). This is illustrated in Figure 3, where HSUV for the general population based on data from Stavem et al (57) are represented by the blue line. The yellow and the grey line show HSUVfor patients with and without age adjustment respectively.

Figure 3: Development of HSUV over time

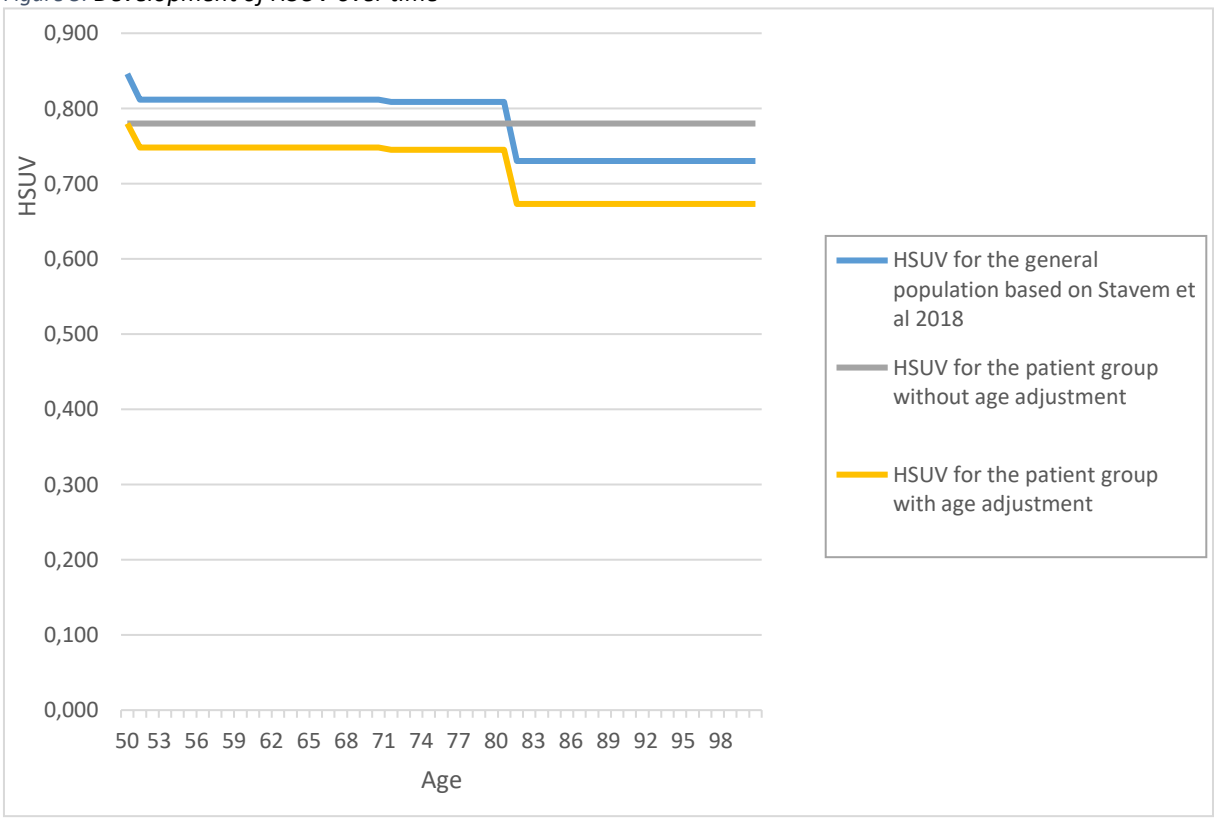

Table 6 Calculating age-adjusted HSUV

Baseline HSUV for the patient group	Age	HSUV for the general population based on Stavem et al 2018	Adjustment index	HSUV for the patient group with age adjustment	HSUV for the patient group without age adjustment
0.78	50	0.846	1.000	0.780	0.780
	51	0.811	0.959	0.748	0.780
	52	0.811	0.959	0.748	0.780
	53	0.811	0.959	0.748	0.780
	54	0.811	0.959	0.748	0.780
	55	0.811	0.959	0.748	0.780
	56	0.811	0.959	0.748	0.780
	57	0.811	0.959	0.748	0.780
	58	0.811	0.959	0.748	0.780
	59	0.811	0.959	0.748	0.780
	60	0.811	0.959	0.748	0.780
	61	0.811	0.959	0.748	0.780
	62	0.811	0.959	0.748	0.780
	63	0.811	0.959	0.748	0.780
	64	0.811	0.959	0.748	0.780
	65	0.811	0.959	0.748	0.780
	66	0.811	0.959	0.748	0.780
	67	0.811	0.959	0.748	0.780
	68	0.811	0.959	0.748	0.780
	69	0.811	0.959	0.748	0.780
	70	0.811	0.959	0.748	0.780
	71	0.808	0.955	0.745	0.780
	72	0.808	0.955	0.745	0.780
	73	0.808	0.955	0.745	0.780
	74	0.808	0.955	0.745	0.780
	75	0.808	0.955	0.745	0.780
	76	0.808	0.955	0.745	0.780
	77	0.808	0.955	0.745	0.780
	78	0.808	0.955	0.745	0.780
	79	0.808	0.955	0.745	0.780
	80	0.808	0.955	0.745	0.780
	81	0.730	0.863	0.673	0.780
	82	0.730	0.863	0.673	0.780
	83	0.730	0.863	0.673	0.780
	84	0.730	0.863	0.673	0.780
	85	0.730	0.863	0.673	0.780
	86	0.730	0.863	0.673	0.780
	87	0.730	0.863	0.673	0.780
	88	0.730	0.863	0.673	0.780
	89	0.730	0.863	0.673	0.780
	90	0.730	0.863	0.673	0.780
	91	0.730	0.863	0.673	0.780
	92	0.730	0.863	0.673	0.780
	93	0.730	0.863	0.673	0.780
	94	0.730	0.863	0.673	0.780
	95	0.730	0.863	0.673	0.780
	96	0.730	0.863	0.673	0.780
	97	0.730	0.863	0.673	0.780
	98	0.730	0.863	0.673	0.780
	99	0.730	0.863	0.673	0.780
	100	0.730	0.863	0.673	0.780

Appendix 4. Calculating severity

4.1 Detailed procedure for calculating absolute shortfall (AS) for treatment interventions

4.1.1 Age

Define the mean age at start of treatment for the relevant Norwegian patient group under consideration for the new treatment. If the age spread in the patient group is very uneven, the median age can be considered. There must be consistency between the age used in the severity calculations, the age in clinical practice and the age in the health economic model. Where there is considerable uncertainty or divergent estimates of age from different sources, it can be useful to usean age interval. Account for where in the interval the mean or median is most likely to lie.Sources for mean age estimation can be registry data, study data and/or information from clinicalexperts. Use the source which best reflects the relevant population in Norway.

4.1.2 Expected remaining QALYs for the general population

Estimate the number of remaining QALY for an average person from the general population with the age found in point 4.1.1. This can be called the quality adjusted expected remaining lifetime from the relevant age. We use the term $QALYs_A$ – short for remaining QALYs at age A. It is the remaining QALYsof both men and women, seen as one, which is used in the calculations, not the gender-specific expected QALYs.

In order for the calculations to be as comparable as possible, the following main sources are recommended for use in calculating $QALYs_A$: Use mortality data for the Norwegian population from Statistics Norway in calculating expected remaining lifetime at different ages (58). This is combined with age-specific quality of life data to calculate quality adjusted remaining lifetime for different ages. We recommend using Table 7 in appendix 4.6. The table shows the expected remaining quality adjusted life years according to age in the average population..

4.1.3 Prognosis

Calculate the prognosis for the relevant patient population at the start of treatment. The prognosis is the average number of remaining healthy life years for the patient group with the current standard treatment P_A. The prognosis is therefore calculated for the treatment the patient group would have received if the new intervention was not used, ie, the current standard treatment (comparator). If there is currently no active treatment, the choice of patient population for calculating the prognosis must be in accordance with the guidelines for choice of comparator in health economic analyses, for example, best supportive care or no treatment, see choice of comparator, Chapter 3.4. The prognosis is calculated for the rest of the patient group's lifetime and is based on the mean value. The prognosis is measured in QALYs. Calculate the prognosis from the number of QALYs the patients can expect with the comparator treatment (usually the current standard treatment) in the health economic analysis. When the health economic calculations are based on a lifecycle model (eg, Markov), it is normally

useful to have a model-based estimate to ensure consistency between thedifferent priority-setting criteria. In the following, we use P_A to denote prognosis at age A.

Sources for prognosis calculation: Prognosis, measured in undiscounted QALYs, for the patient group being treated with the comparator in the health economic model, will usually be useful as a source for the severity calculation. Alternative sources are relevant clinical studies, registry data or data from systematic literature searches.

4.1.4 Absolute shortfall

$AS = QALYs_A - P_A$

In the calculations, undiscounted numbers for $QALYs_A$ and P_A are used.

Uncertainty in calculating AS must be discussed. This applies for example to uncertainty in the estimates of age or prognosis.

4.2 Examples – calculation of degree of severity for treatment interventions

4.2.1 Example of calculation of absolute shortfall for disease A.

Based on a health economic model with a lifetime perspective.

1. Age. The mean age at treatment start in the patient group relevant for treatment is estimated by clinical experts to be 57 years. This is supplemented by data from national registries.
2. For a 57 year old the expected remaining healthy life years ($QALYs_{57}$) is calculated as 22.0 QALYs. See appendix. 4.6.
3. Prognosis. Patients have an expected remaining life time of 2.5 years, corresponding to 1.5 QALYs (undiscounted) with the current standard treatment (the comparator). This is based on simulations with the health economic model included in the company's documentation, after the Norwegian Medicines Agency has evaluated the documentation.
4. The absolute shortfall (AS) will then be 22.0 QALYs – 1.5 QALYs = 20.5 QALYs.

Figure 4 below illustrates the AS calculation for the treatment intervention. The figure applies on a patient group level. The Y-axis shows health-related quality of life, (HRQoL) on a scale from 0 (dead) to 1 (full health). The X-axis shows time. The new treatment is considered at age A. Without the disease, the future health would be given by the area under the solid blue line from timepoint A. This is given as $QALYs_A$, cf. the example above. The disease leads to a shortening of lifetime and a reduction in the quality of life (with the current standard treatment). The prognosis with the disease and current treatment is shown in the shaded area P_A. The absolute shortfall (AS) is shown as the difference between $QALYs_A$ and P_A.

The figure does not include any potential health loss linked to the disease before the start of treatment. This is because the Priority-setting White Paper only recommends future health loss for quantifying severity.

Note that the efficacy of the new treatment/intervention which is being evaluated, is *not* included in the calculation of severity. The efficacy is included in evaluation of the other priority-setting criteria: namely, benefit. In the calculation of severity (absolute shortfall) the efficacy (prognosis) with the current standard treatment is included.

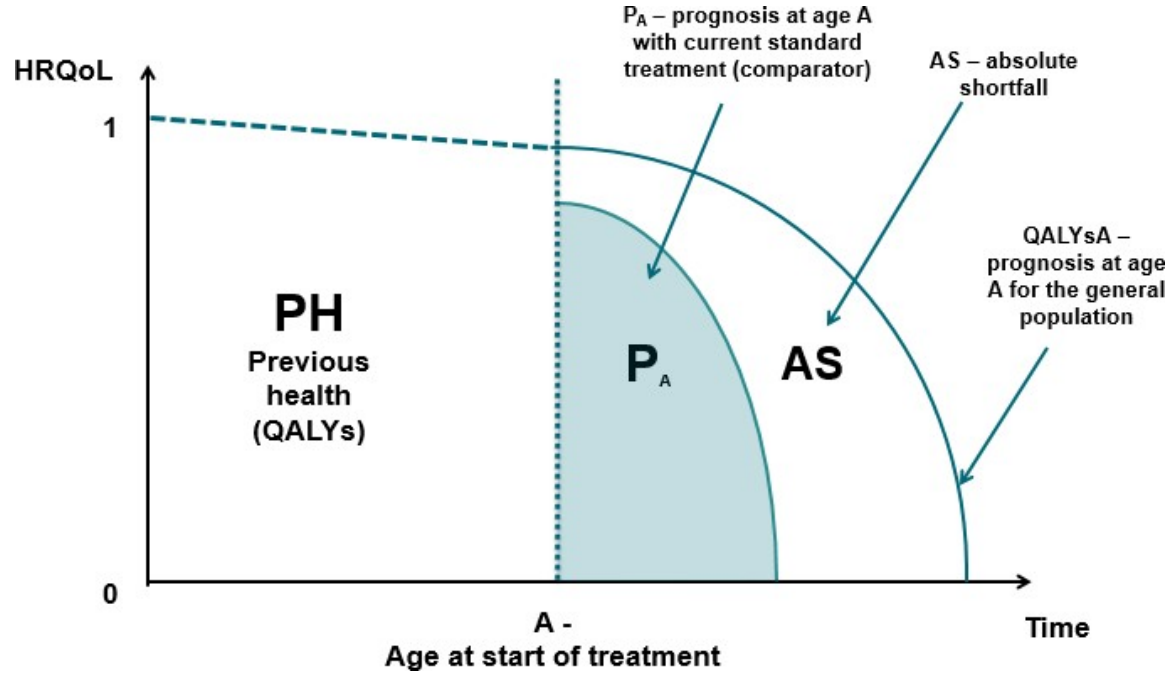

Figure 4: How to quantify severity

4.2.2 Example of calculating absolute shortfall for disease B.

Based on a health economic model with shorter time perspective than lifetime, eg, with a one-year perspective. This could be a chronic, non-fatal condition.

1. Age. The mean age for treatment start in the relevant patient group is estimated at 50 years.
2. For a 50 year old the expected remaining healthy life years ($QALYs_{50}$) is calculated as 27.3 QALYs. See appendix 4.6
3. Prognosis. The prognosis (undiscounted) in the health economic model analysis is 0.750 QALYs. But this is the prognosis on a 1-year timescale, not for the rest of life. The prognosis for the rest of life must be calculated. The calculation will depend on the disease and the disease progression with the current standard treatment. A stylized way to calculate lifetime prognosis can be as follows: assume that from another source, eg, Stavem et al (57), the HSUV for a 50 year old is 0.846. Assume also that the relative QALY loss caused by the disease is constant for the rest of life: Then the prognosis can be simply estimated in this way: Prognosis= (0.750/0.846) * 27.3 QALYs = 24.2 QALYs. In the calculation 27.3 QALYs is the expected remaining QALYs for a 50 year old.
4. AS will then be 27.3 QALYs – 24.2 QALYs = 3.1 QALYs.

4.3 Example of calibrating two data sources – level adjustment

In some cases, the HSUV for symptom-free states in the health economic analyses which form the basis for prognosis calculation are higher than the HSUV for the average population used in the calculation of remaining QALYs. This should, as a rule, be corrected for by calibration.

Example:

A single technology assessment of a new intervention where the mean age at treatment start is 70 years. The prognosis estimate for established treatment is taken from the health economic model. The prognosis is 3 QALYs. From the quality-adjusted life years tables (see appendix 4.6) the remaining QALYs for a 70 year old will be 12.9 QALYs. The AS is 12.9 QALYs – 3 QALYs = 9.9 QALYs.

The HSUV in the prognosis calculation will come from clinical studies of the intervention being evaluated or from other studies where the quality of life for the disease/condition have been measured, while the HSUV included in the quality-adjusted lifeyears table come from another source. This must be taken account of in the way shown in our example:

In the health economic analysis the condition has a "symptom-free" HSUV of 0.850. This weight is used in the prognosis calculation. In the calculation of remaining $QALYs_{70}$ however, the HSUV for an average 70 year old is lower at 0.811 and is based on Stavem et al (57).

This should be adjusted for by multiplying the prognosis estimate by the factor 0.811/0.850.

Thus the adjusted prognosis will be 3 QALYs* 0.811/0.850 = 2.9 QALYs. The adjusted absolute shortfall will then be 12.9 QALYs – 2.9 QALYs = 10.0 QALYs.

In this example, the adjustment did not lead to major changes in the calculated absolute shortfall. In other cases, it can make more difference. In general terms, when adjustment has been used, companies should consider whether the adjustment is reasonable.

4.4 Examples – calculation of degree of severity for preventive measures

4.4.1 New measure which prevents one type of disease.

1. Age. The new preventive measure is given to the relevant population from a mean age of 40 years. For the population, the disease occurs on average from age 60 with the current preventive practice (the comparator in the health economic analysis). The age that must be used in the calculation of absolute shortfall is 60 years.
2. For a 60 year old the expected number of remaining healthy lifeyears ($QALYs_{60}$) is calculated as 19.8 QALYs. (See appendix 4.6).
3. Prognosis. For this disease and the relevant population, the prognosis is 7.3 QALYs with the current standard treatment. The prognosis reflects the fact that some individuals who get the disease will die of it, while the majority will survive, albeit with somewhat reduced quality of life. Heart attack is an example of a disease/event of this type. The average prognosis will thus be a weighted average of the prognosis for those who die of the disease/event and those who survive.
4. AS is estimated as 19.8 QALYs – 7.3 QALYs = 12.5 QALYs.

Calculation of absolute shortfall for prevention can also be shown by the figure above, but then the absolute shortfall is calculated

- From the timepoint that the disease manifests (timepoint A), not from the timepoint that the preventive measure is carried out or started.
- for a patient who gets the disease the prevention is aimed at, not for a person who gets the preventive measure. The figure will thus refer to the sub-group who get the disease at a later timepoint (timepoint A).
- with prognosis based on the current standard *treatment* of the disease.

4.4.2 New measure that prevents two types of disease, A and B – Calculationof weighted absolute shortfall

Procedure for calculating weighted absolute shortfall (weighted AS):

1. Calculated AS for disease A: 10 QALYs
   Calculated AS for disease B: 6 QALYs

2. In the health economic analysis the benefit is estimated as 2.0 QALYs. This is the average incremental effect per person who receives the measure. 1.8 QALYs, ie, 90 %, of the benefit is linked to prevention of disease A. 0.2 QALYs, ie, 10 % of the benefit is linked to prevention of disease B.

The weighted AS for disease A and B in this case will then be: 90% * 10 QALYs + 10 % * 6 QALYs = 9.6 QALYs.

4.5 Justification of the suggested principle for weighted AS for measureswhich prevent and/or treat several types of disease

1. Severity must be taken into account along with the other two priority-setting criteria, benefit and use of resources, in prioritising between measures/interventions. Benefit and use of resources are included directly in a cost-effectiveness analysis for calculating the cost per QALY- ratio of the measure. Severity is included in the form of severity weights where, after the cost- effectiveness analysis, it is decided what is the highest acceptable cost per QALY- ratio. A higher AS gives a higher severity weight, and therefore a higher acceptable cost per QALY ratio.

2. All relevant benefits and costs must be included in the analysis to give the best possible basis for decision-making with regard to the *effectiveness* of the measure. Severity is a consideration of *distribution or fairness* which is considered in addition to effectiveness. If a measure is to be given a high overall severity weighting in prioritisation, it should appear as a good measure for the treatment or prevention of severe diseases. Then the benefit from the measure should be linked to the treatment or prevention of severe diseases. If the measure is aimed a several diseases, then the disease which is most important when estimating the benefit of the new measure, should be given the greatest weight when the measure is accorded a severity weighting.

3. All the benefit components which are included in the benefit evaluation of the measure, will

individually contribute towards making the measure more cost effective. If a company chooses to include benefit for prevention or treatment of *several* diseases in its analysis, in order to achieve a better cost-benefit ratio, then the company must expect that the overall severity weight (weighted AS) across the diseases will be calculated using a weight based on the different diseases' share of the benefit.

4. This means that weights according to the diseases' share of benefit in the cost-effectiveness analysis are logical and consistent for use in the severity evaluation based on the weighted AS (given the use of weighted AS in the prioritising between interventions/interventions, cf. point 1 above). Such a weighting can be used whether it is the same patient group which has/willget several diseases or different groups which each have/will get one of the diseases.

4.6 Expected remaining QALYs in the general population

NIPH has updated the population norms for EQ-5D[19] (HSUV) with the recently published population norms by Stavem et al (57). The population samples included are representative of the Norwegian general population, and the collected data are more recent than the Swedish population norms used in our previous versions (59, 60), though the number of respondents is lower. We have not changed tariff for scoring the EQ-5D index and use the population based UK tariff (61). Table 7 shows the expected remaining QALYs and (health-related) HSUV respectively, by age for the general population. Expected remaining QALYs are based on mortality data for the Norwegian population from Statistics Norway (58) and the age-specific HSUV in the right hand column.

Stavem *et al* (57) covers the age group from 19 to 97. HSUV (values in parentheses) for the age groups 19-50 years in 10-year brackets are directly incorporated from Stavem *et al* (57): 19-30 (0.906), 31-40 (0.870), 41-50 (0.846). Using the raw data[20] from Stavem *et al* (57), we have calculated a simplified weighted average[21] for the age groups 51-70[22] (0.811) and 71-80 (0.808). The raw data is also used for the HSUV for the age group above 80 (0.730). This sharper decrease in HSUV after age 80 compared with the decrease between ages 50 and 80 is supported by findings in the Tromsø Study (T7, unpublished) and in European health status surveys (62-64). Furthermore, NIPH assumes that HSUV are somewhat higher in the younger age group (0-19) and uses the same increment as before (0.02) yielding a HSUV of 0.926

[19] NIPH uses the same strategy in calculating and extrapolating the Norwegian HSUV as we did with the previous Swedish based figures.
[20] Stavem – Personal communication
[21] The raw data were available for the groups 71-75 and 76-80; the average is weighted by the fraction of responders in each of the two age groups.
[22] Stavem *et al* reported lower utility values in the age bracket 51-60 compared with 61-70 years. Such fluctuations are not reported in other comparable studies, and NIPH chose to smooth the HSUV by weighting an average for the pooled 51-70 group.

Table 7: Expected remaining QALYs and HSUV in the general population

Age	Expected remaining QALYs	HSUV	Age	Expected remaining QALYs	HSUV	Age	Expected remaining QALYs	HSUV
0	70.9	0.926	36	38.8	0.870	72	11.6	0.808
1	70.2	0.926	37	37.9	0.870	73	11.0	0.808
2	69.2	0.926	38	37.1	0.870	74	10.4	0.808
3	68.3	0.926	39	36.2	0.870	75	9.8	0.808
4	67.4	0.926	40	35.4	0.870	76	9.2	0.808
5	66.5	0.926	41	34.6	0.846	77	8.7	0.808
6	65.6	0.926	42	33.7	0.846	78	8.1	0.808
7	64.6	0.926	43	32.9	0.846	79	7.5	0.808
8	63.7	0.926	44	32.1	0.846	80	7.0	0.808
9	62.8	0.926	45	31.3	0.846	81	6.5	0.730
10	61.9	0.926	46	30.5	0.846	82	6.0	0.730
11	61.0	0.926	47	29.7	0.846	83	5.6	0.730
12	60.0	0.926	48	28.9	0.846	84	5.2	0.730
13	59.1	0.926	49	28.1	0.846	85	4.9	0.730
14	58.2	0.926	50	27.3	0.846	86	4.5	0.730
15	57.3	0.926	51	26.5	0.811	87	4.1	0.730
16	56.4	0.926	52	25.7	0.811	88	3.8	0.730
17	55.4	0.926	53	25.0	0.811	89	3.5	0.730
18	54.5	0.926	54	24.2	0.811	90	3.2	0.730
19	53.6	0.906	55	23.5	0.811	91	3.0	0.730
20	52.7	0.906	56	22.7	0.811	92	2.8	0.730
21	51.9	0.906	57	22.0	0.811	93	2.6	0.730
22	51.0	0.906	58	21.2	0.811	94	2.4	0.730
23	50.1	0.906	59	20.5	0.811	95	2.2	0.730
24	49.2	0.906	60	19.8	0.811	96	2.0	0.730
25	48.3	0.906	61	19.1	0.811	97	1.8	0.730
26	47.4	0.906	62	18.3	0.811	98	1.8	0.730
27	46.6	0.906	63	17.7	0.811	99	1.6	0.730
28	45.7	0.906	64	17.0	0.811	100	1.5	0.730
29	44.8	0.906	65	16.3	0.811	101	1.5	0.730
30	43.9	0.906	66	15.6	0.811	102	1.4	0.730
31	43.0	0.870	67	14.9	0.811	103	1.3	0.730
32	42.2	0.870	68	14.2	0.811	104	1.0	0.730
33	41.3	0.870	69	13.6	0.811	105	0.8	0.730
34	40.5	0.870	70	12.9	0.811			
35	39.6	0.870	71	12.3	0.808			

Published by the Norwegian Institute of Public Health
September 2021
P.O.B 4404 Nydalen
NO-0403 Oslo
Phone: + 47-21 07 70 00
The report can be downloaded as pdf
at www.fhi.no/en/publ/