NIPH
Norwegian Institute of Public Health

2020

CONSULTATION DRAFT:
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 guidelines for submission of documentation for single technology assessments (STAs) of medical devices and diagnostic interventions reflect the principles of health sector priority-setting, described in the Norwegian Government's White Paper no. 34 (2015-2016) (1), henceforth referred to as the Priority-setting White Paper ("Prioriteringsmeldingen").

These guidelines define the necessary components of submitted documentation for STAs of medical devices and diagnostic interventions[1] to be financed by the public specialist health services in Norway. The priority-setting criteria and the regional health authorities' responsibility for the system through which new health interventions are adopted were enacted into law in December 2019[2].

The system requires all specialist health service interventions to 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 implemented or not – will consider the intervention in the light of these criteria, together and weighed against each other. As a submitter, you are required to calculate a cost-effectiveness ratio, which reflects the use of resources in relation to benefit. You should do this in a health economic evaluation, typically involving decision analytic modelling. The cost-effectiveness ratio will be

[111] They would in theory also apply to procedures, but NIPH does not anticipate submissions in this area.
[2] https://nyemetoder.no/nyheter/lovfesting-av-prioriteringskriteriene-og-system

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. discretionary evaluation.

The Decision Forum will base its overall evaluation of an intervention on both the factors desbribed 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.

These guidelines describe the requirements and recommendations for the submission of documentation on benefit, resource use and severity of condition as well as budget impact. The guidelines describe the 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.

The template for submitting documentation for single technology assessments (STA) of medical devices, diagnostic interventions and procedures must be used when preparing a submission. The template is available at:
https://www.fhi.no/contentassets/580963b306984ec68da2676ed5fd48b2/template-medical-devices.pdf

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

Contents

Abbreviations

Abbreviation	Definition
AFT	Accelerated failure time model
AIC	Akaike Information Criteria
AS	Absolute shortfall
AUP	Pharmacy maximum sale price
BIC	Bayesian Information Criteria
CCTR	The Cochrane Controlled Trials Register
CDx	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

Abbreviation	Definition
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

Abbreviation	Definition
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
RWE	Real World Evidence
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 Template

The Norwegian Institute of Public Health (NIPH) has developed a template that you must use when you prepare documentation for single technology assessment (STA). This is available on the institute's website: : https://www.fhi.no/contentassets/580963b306984ec68da2676ed5fd48b2/template-medical-devices.pdf. The template includes a number of tables and subheadings that require you to summarize information relevant to the assessment. It is possible to send in appendices or supplementary information attached to the main submission if information does not fit easily into the space provided in the template.

1.2 About NIPH

The Norwegian Institute of Public Health, NIPH (www.niph.no) is charged with performing health technology assessment 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 the "New Methods" system.(nyemetoder.no)[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).

1.3 Comparators

Clinical and cost-effectiveness analyses require comparing the proposed intervention to a relevant alternative treatment. Chapter 3.4 discusses the principles underlying the choice of comparator. Contact NIPH for guidance if there is any doubt about the choice of comparator or any other issues related to the submission.

1.4 Health economic model

The submitted health economic model must be capable of estimating results for the most likely scenarios associated with the intervention. You must ensure that it is possible for NIPH to modify (in Microsoft Excel or TreeAge) the variables in the model.

1.5 References

It is not necessary to include all references in the documentation. However, you must include references for all of the most important efficacy studies and references used as the *basis for input data* in the health economic analysis and in calculations of severity and budget impact. References (in documents and models/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.

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

1.6 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.

1.7 Language

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

1.8 Confidentiality

NIPH operates within the Public Administration Act[4] and the Freedom of Information Act[5]. 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.

1.9 Medical devices and in vitro diagnostics: Regulatory Issues

The Norwegian Medicines Agency (NoMA) is the competent authority for medical devices in Norway. This implies that the Agency has administrative and advisory responsibilities related to legislation and supervisory authority over manufacturers, distributors 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 dir 93/42 article 1, 2a for the complete English definition)[6] . In vitro diagnostics (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 trial shall be sent to the Norwegian Medicines Agency (further information Clinical investigation).

1.10 The Norwegian legislation

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

1.10.1 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.

[4]Forvaltningsloven https://www.regjeringen.no/no/dokumenter/forvaltningsloven/id449156/

[5] Offetnlighetsloven https://lovdata.no/dokument/NL/lov/2006-05-19-16

[6] COUNCIL DIRECTIVE 93/42/EEC of 14 June 1993 concerning medical devices. https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CONSLEG:1993L0042:20071011:EN:PDF

1.10.2 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.

1.10.3 Regulation on the use of medical devices

This regulation[7] applies to the health care 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.

1.10.4 New EU regulations on medical devices

On 5 April 2017, the EU adopted two new regulations on medical devices and in vitro diagnostic medical devices. They entered into force on 25 May 2017 and will progressively replace the existing directives. The new regulations will be fully applicable in May 2020 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, L 117, 5. May 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[8] .

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.

[7] https://lovdata.no/dokument/SF/forskrift/2013-11-29-1373

[8] https://www.regjeringen.no/no/dokumenter/horing-om-gjennomforing-av-forordning-2017745-om-medisinsk-utstyr-og-forordning-2017746-om-in-vitro-diagnostisk-medisinsk-utstyr/id2652683/?expand=horingsbrev)

2. Scope

The scope of a submission includes a short description of the STA indications; the relevant patient population for the STA; the type of intervention; the alternative(s) for comparison; and the most important outcome measures in the analysis (as defined by the PICO).

Submit documentation in accordance with the request from the Commissioning Forum, Nye Metoder (Bestillerforum RHF). NIPH must agree in advance to any variation from the requested documentation. Communication throughout the duration of the assessment will be with NIPH. Submitters will 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.)

3. Description of the intervention and the therapeutic area

3.1 Description of the disease/condition and Norwegian clinical practice

Briefly describe the relevant disease or condition targeted by the proposed intervention and how patients are currently 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 5 years. For very small patient groups, also include the worldwide prevalence.

3.2 Description of the intervention and patient population

Describe the intervention in accordance with the template for submission of documentation (see Chapter 1), including its main characteristics and how it is intended to work. The description must outline any institutional or organisational structures that need to be in place for the intervention to work in an optimal manner.
If relevant, specify needs for IT support for the operability of the intervention, as well as requirements for data management and interpretation of results. In addition, specify requirements for quality assurance procedures.

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. 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.

In addition, describe the position the intervention is supposed to fill in the treatment pathway for the defined population.

3.3 Comparator(s)

Account for the choice of comparator(s) based on the following guidelines (3.3.1 and 3.3.2). Contact NIPH for guidance if you have any doubt about the choice of comparator(s).

3.3.1 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 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 take the form of diagnostics, prevention, curative treatment, palliative treatment or "wait and see" initiatives. Only in exceptional cases will comparison with no treatment be relevant.

Different treatment sequences may also be evaluated if robust clinical evidence for performing the comparison exists. Contact NIPH for more information.

3.3.2 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 can not 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 cumentation

The aim of the literature search is to document the methods used to determine the central data sources used in the STA. Use the literature search to identify relevant documentation for:

- Efficacy/safety data used to document an intervention's relative efficacy
- Health state utility values (HSUV) (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

Base the literature search on internationally validated methods (2-5).
Documentation of the literature search process must include, at least, the following elements

- Details of the search strategy, as performed in one of the searched databases, that allows the search to be reproduced:
  - Precise formulation of the research question (including specified PICOs)
  - Search strategy with the associated search strings, MEsH terms, truncations, etc.
  - Description of the MeSH terms used
  - An *a priori* definition of the inclusion and exclusion criteria, and reasoning for these
  - Justification for the chosen timeframe for the search (how far back in time)
  - A list of all databases used for the search
- Data extraction:
  - Describe the process for selecting studies (including whether one or more reviewers have been involved, how disagreements were handled, e.g., by an independent professional colleague).
  - Record excluded studies and the basis for exclusion.

If the original literature search is more than a year old, it must be updated by repeating the search for the following period. Include a list of new, relevant studies.

At a minimum, the following databases should be included:
Efficacy and safety:

- The Cochrane Controlled Trials Register (CCTR)
- MEDLINE/PubMed, EMBASE
- NIH Clinical trials
- International Clinical Trials Registry Platform
- More specific databases or other relevant electronic databases not covered by CCTR (e.g., PsychInfo, Pedro etc.).
- Systematic review databases (Cochrane, Epistemonikos, etc.)

Quality of life:

- MEDLINE/PubMed, EMBASE, and other more specific databases. For a detailed description of how a literature search for quality of life can be carried out, as well as which databases are relevant, see NICE DSU Technical Support Document 9 (5).

A manual search in other sources (e.g., ClinicalTrials.gov, WHOs), conference posters, conference abstracts, reference databases and other types of documentation not covered by electronic databases) as well as grey literature searches can also be relevant.

5. Documentation of clinical efficacy and safety

5.1 Efficacy of intervention and comparator(s)

In accordance with the template for submitting documentation to NIPH, you must present the clinical studies associated with the intervention under consideration, and include the following information: study design, name of intervention, comparator, sample size, patient population, outcome measures and 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.

If you have included additional endpoints in the health economic model, 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):

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 with regard to 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

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 tjat you describie these and provide references.

5.6 Data sources

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

- Relevant published/unpublished non-confidential data, information on terminated or ongoing trials
- Critical appraisal of the data included in the assessment; internal validity, i.e. Risk of Bias
- 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.

5.7 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 as long as the modified properties and their significance are clearly described.

6. Documentation of relative effectiveness

6.1 Relative effectiveness

6.1.1 Direct comparisons

The underlying clinical evidence supporting relative effectiveness should be based on the literature review (see Chapter 4). Efficacy and safety data from RCTs are preferred. If relevant systematic reviews of the intervention exist, these may be submitted as part of the documentation.

6.1.2 Indirect comparisons

In cases where no direct comparisons between the intervention and a relevant comparator are available, indirect comparisons may be performed. 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. Underlying assumptions should be presented and discussed. For further details, see Appendix 1.

6.2 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 in order 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.

6.3 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 increase 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 such statistical methods and justify their use in the submission.

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 was published by NoMA in 2015 (more details in Appendix 2).

7. Diagnostic interventions

7.1 Use of diagnostic interventions

The purpose of a diagnostic intervention is to determine the presence of a disease or condition in an individual who may benefit from medical treatment. The four main uses of diagnostic tests are: diagnosis, monitoring, screening and prognosis (14). Once individuals with the condition have been identified, they may be offered treatment. However, diagnostic interventions may also involve adverse effects such as misdiagnosis, anxiety, unnecessary treatment and undetected cases.

When submitting information on a diagnostic test, you should include: The test's scope, i.e. what is its purpose (screening, treatment triage, disease state assessment or risk stratification) (14). The type of test, its associated decision rules/algorithms; its cut-off values (rule in/rule out), detection limits and the prevalence of the condition it is designed for.

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 relevant characteristics (e.g. feasibility, risk of adverse events, comfort) that may be important to the patient, but are not captured by the test outcomes.

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[9]. 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.

7.2 Diagnostic intervention studies

Research in the field of diagnostic interventions may be grouped into four categories: (15) i) Technical performance, ii) Diagnostic accuracy iii) Patient outcomes and iv) Cost-effectiveness. Ideally, diagnostic interventions should be supported by studies that follow patients from testing via treatment to final clinical outcome, so-called end-to-end studies (14). Documentation that may provide evidence for the significance of the diagnostic test for clinical outcomes are of particular interest.

[99] The Norwegian Medicines Agency, In-house medisinsk utstyrs: https://legemiddelverket.no/Documents/Medisinsk%20utstyr/Tredje%20presentasjon%20h%C3%B8ringsm%C3%B8te%20om%20nye%20regler%20for%20medisinsk%20utstyr.pdf

If end-to-end studies are available, you must submit them to NIPH.
If such evidence is not available, it is acceptable to provide separate evidence for test characteristics, analytical validity, clinical validity, and clinical treatment outcomes (clinical utility) in order to demonstrate how they are linked and estimated in the health economic model.

For studies on test accuracy, provide information on the decision rule/algorithms and whether or not it is in the public domain. Specify parameters such as reference standard, prevalence and test results in terms of sensitivity and specificity.

Studies should be critically appraised using an appropriate instrument such as GRADE for diagnostic interventions. See for example, a guide on applying the GRADE instrument to diagnostic interventions (16).

7.3 Resource use

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, the average costs of the indication should be estimated. Some of the fixed costs may be assigned to other indications in the sensitivity analysis given proper justification (14). Costs associated with detection and follow-up of true positive and false positive cases should be included. If introduction of the technology requires additional infrastructure, these costs should be incorporated in the analysis, by inclusion in the average cost. ,

7.4 Modelling

Describe the relevant treatment alternatives associated with the disease state or condition tested for and the potential clinical outcomes. Provide a detailed explanation of the patient pathway and how it 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)

8. Health-related quality of life

Use Quality-adjusted life-years (QALYs) as the benefit measure for STAs at the group level. Patient-reported measures based on EQ-5D are preferred. 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 (see chapter 4).

Uncertainty in health state utility values (HSUV)[10] must be examined in sensitivity/scenario analyses (see 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 rule, 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[11]. For guidance in deciding whether or not 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 Hout et al (21). The use of EQ-5D-3L as the standard in STAs is based on recommendations from NICE (22). For studies that have used the 5L version, we recommend the newly published English tariff by Devlin et al. (23) until further notice.

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

Valuation of quality of life in STAs must, as a rule, be based on tariffs (value sets) from the preferences of the general population. This will ensure consistency across STAs, and internal

[10] Also referred to as QALY weights

[11] 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.

consistency between measures of severity and health economic analyses in individual STAs. 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, for a given STA, there are specific reasons for using an experience-based tariff, i.e. one that reflects patient preferences rather than those of the general population, this should be justified. Explain how this tariff varies from a general population-based tariff.

In principle, the tariff used in the analyses should be relevant to the Norwegian population. However, as no Norwegian tariff currently exists, we recommend using the UK population-based tariff (24) for STAs in Norway until a more relevant and applicable tariff is available. As a standard for STAs the use of EQ-5D with UK tariffs is strongly recommended.

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)

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[12].

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 (see appendix 4.1.2).

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[13], and gives an age-adjusted HSUV. Appendix 3 provides an example of how to perform this calculation.

[12] For example, when a health economic analysis has a short time perspective.
[13] Is set to 1 in the starting year in the health economic model, and decreases with increasing age.

8.5 Treatment-specific HSUVs for the same condition

If different treatment-specific HSUV are used for the same condition[14], 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.

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, included 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.

[14] 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.

9. Health economic analyses

For an overview of a reference case for health economic analyses, see Appendix 5 Reference case – health economic analyses.

9.1 Analysis methods

9.1.1 Cost-utility analysis (CUA)

The recommended analysis method for health economic evaluations is 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

See the Norwegian Directorate of Health's guidance "Economic evaluation in the health sector" (updated, 2018) for more in-depth information about the perspective of analyses and analysis methods for different types of interventions which affect health.

9.3 Resource use and costs

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 calculated directly as part of the STA, they can be taken from other cost studies/publications. 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[15] or outpatient clinic attendance[16] 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[17]. The patient's co-payment for outpatient

[15] Applies to admissions to somatic departments.

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

[17] 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 rest of the 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 for our purposes to use DRG-points in

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, but multiplying the remuneration 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[18]**.**

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.

[18]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.

9.3.3 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.

9.3.4 Capital costs and fixed medical equipment

With regard to 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 time 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, 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).

9.4 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.

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, validate models should 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 not be implemented (fully or partly) in proprietary or non-transparent programs and/or programming language.[19]

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 documented causal relationship 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 estimate for the primary endpoint, or the hard endpoints, should be used in the model.

[19] Normally it will be useful if the model is designed using Excel, but other alternatives can also be acceptable such as, TreeAge.

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 Organizational implications

10.3.1 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).

10.3.2 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.

10.4 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

Severity must be quantified 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 quality-adjusted life years (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'''. 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 – Calculation of severity.

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, see appendix 4, section 4.1.2. 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 populaton (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 section 4.3.

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, section 4.4.

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, section 4.5.

This weighting principle is also relevant if the intervention can both treat and prevent disease. An example would be an MRI machine used to screen for, and monitor 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

Uncertainty in health economic analyses must be explored and discussed. 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**[20]: 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.

[20] 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 %).

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 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 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.
The budget impact has three components:

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

The costs are calculated 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.

You may choose not to include these costs in your budget calculations, but must then explain why it is plausible that the budget impact is negligible or negative (i.e., budget savings). You may also, if you provide reasonable justuification (e.g. organizational complexity) present the budget impact using other methods than those recommended below (contact NIPH for guidance).

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 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 specialist health care services should be presented

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 health care 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 who are 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 tables below show how the calculation of budget impact for the specialist health care 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 health care 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 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 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 (NoMA) 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 thes Agency's guidelines came into force on January 1st 2018 and an English version was published in April the same year. 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 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).

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. GuiGuidelines for the Economic Evaluation of Health Technologies: Canadadelines 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. Excellence NIfHaC. Diagnostic Assessment Programme Manual. 2011.
15. Fryback DG, Thornbury JR. The efficacy of diagnostic imaging. Medical decision making : an I nternational journal of the Society for Medical Decision Making. 1991;11(2):88-94.
16. Singh S, Chang SM, Matchar DB, Bass EB. AHRQ Methods for Effective Health Care Grading a Body of Evidence on Diagnostic Tests. In: Chang SM, Matchar DB, Smetana GW, Umscheid CA, editors. Methods Guide for Medical Test Reviews. Rockville (MD): Agency for Healthcare Research and Quality (US); 2012.
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. Devlin NJ, Shah KK, Feng Y, Mulhern B, van Hout B. Valuing health-related quality of life: An EQ-5D-5L value set for England. Health economics. 2017.
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: http://legeforeningen.no/Arbeidsliv-og-jus/naringsdrivende/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 health care 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 health care 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 r esults 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

Justify why it is necessary to use an indirect comparison.

The research question /scope must be clearly formulated.

Before an indirect comparison is conducted the known effect-modifiers and prognostic factors must be described as fully as possible based on existing knowledge.

1.2 Literature search

Conduct a full systematic literature search. Describe the literature search in detail, both for the relevant intervention and the chosen comparator(s). For literature searches intended to support the documentation of relative efficacy, the PICO[21] 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 conducted, 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 conducted 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):

[21] 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, based on clinical evidence, whether the compared studies overlap sufficiently in terms of study design, inclusion criteria, patient characteristics, definitions 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 compared studies, and assess the extent to which there is enough information in the studies to correct completely for all of these factors. Account for covariates that cannot be taken into account in the analysis. Discuss the risk of unmeasured confounding factors that could affect the analysis.

In a MAIC, patients from the study with individual patient data (IPD) are assigned weights[22], 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, i.e., 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 excluding a variable 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 (3, 39-44).

[22] 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, i.e., progression-free survival (PFS), time to death, i.e., overall survival (OS), or time to a cardiovascular event. 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 parametrization with extrapolation of the clinical time-to-event data beyond the actual study period. Below we specify how parametrization and extrapolation of survival data must be performed 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 10.4 Time horizon.

2.2 Parametrization 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 (e.g., Kaplan-Meier) or semi-parametric (e.g., 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.

Parametrization 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 performed 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 that do not fulfill both 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). At a minimum, all of the items in the list below must be included to document adjustments to the observed study data:

- Log-cumulative hazard plot for the different parametric functions as a guide for the choice of parametric function and the choice between proportional hazard model (PH) and accelerated failure time model (AFT).

- Statistical tests and graphical presentation of proportional hazard (PH), if such a model is chosen. Examples are log-cumulative hazard plot and plot based on Schoenfeld residuals, but if other graphical methods are more suitable, these must be used.

- If neither PH nor AFT appears suitable, another, more flexible function must be considered, e.g., a piecewise function, Royston-Palmer models, or spline models.

- 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

- Graphical presentation of time-to-event data curves, where both Kaplan-Meier (KM) data and the parametric distribution is shown in the same figure.

- In some cases, curves with KM data for the first part of the study period can be appropriate, followed by a parametric tail that shows the extrapolation beyond this point (transition point). The transition point must be evaluated in the individual case. As a minimum requirement, an analysis must be presented where the tail is set at the time point at which 50 % of the included population in each treatment arm is still "at risk".

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 to what extent any differences in long-term survival between the projected survival curve and the external data source reflect:

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

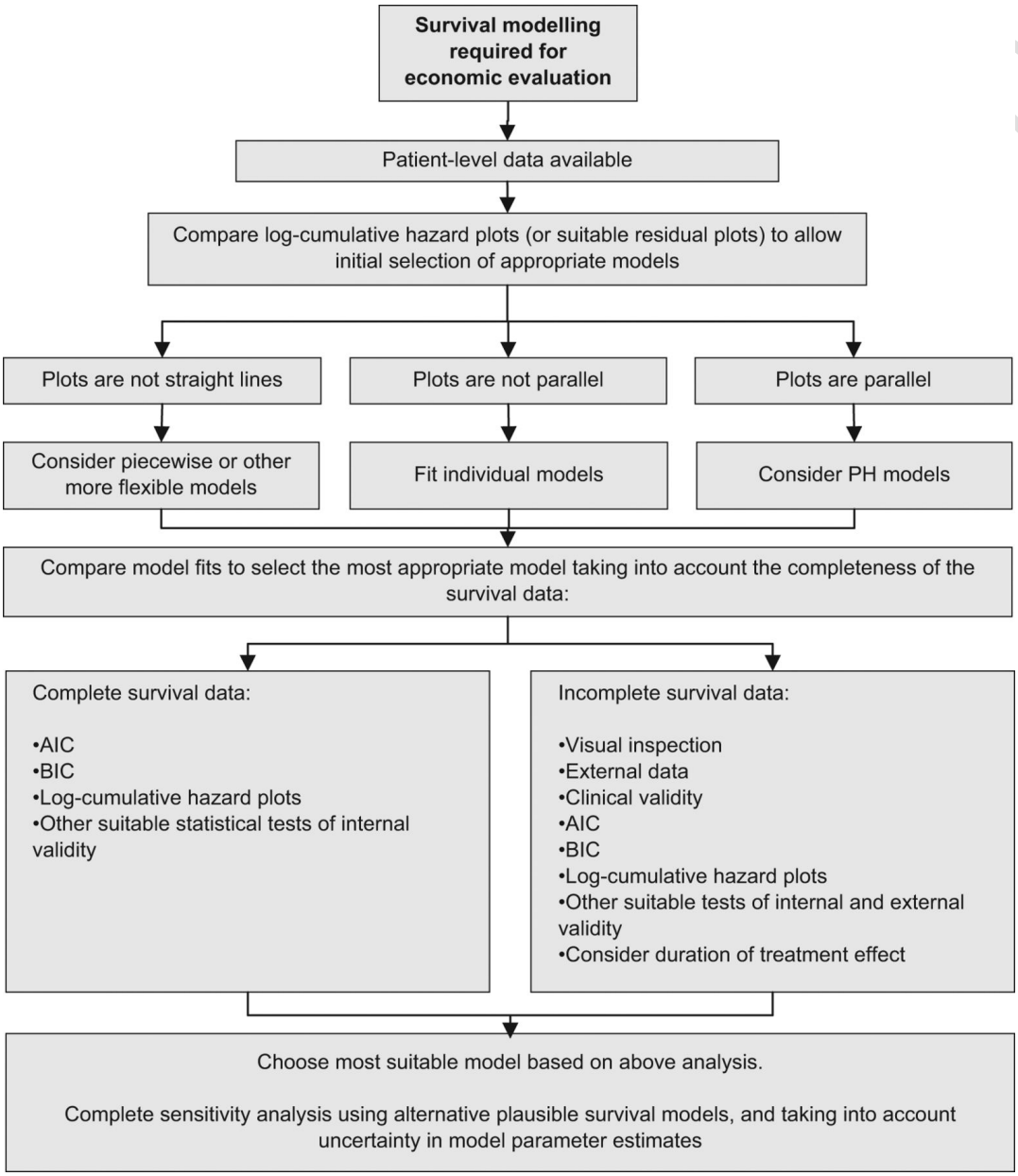
External data will, most likely, only be available for the comparator arm, and will therefore be most useful for evaluating the plausibility of projecting the comparator arm. Therefore, clinically valid assumptions about the duration of treatment effect will be necessary for extrapolating the effect 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. From Latimer 2013 (50)



2.3 Studies in which patients can switch to new intervention

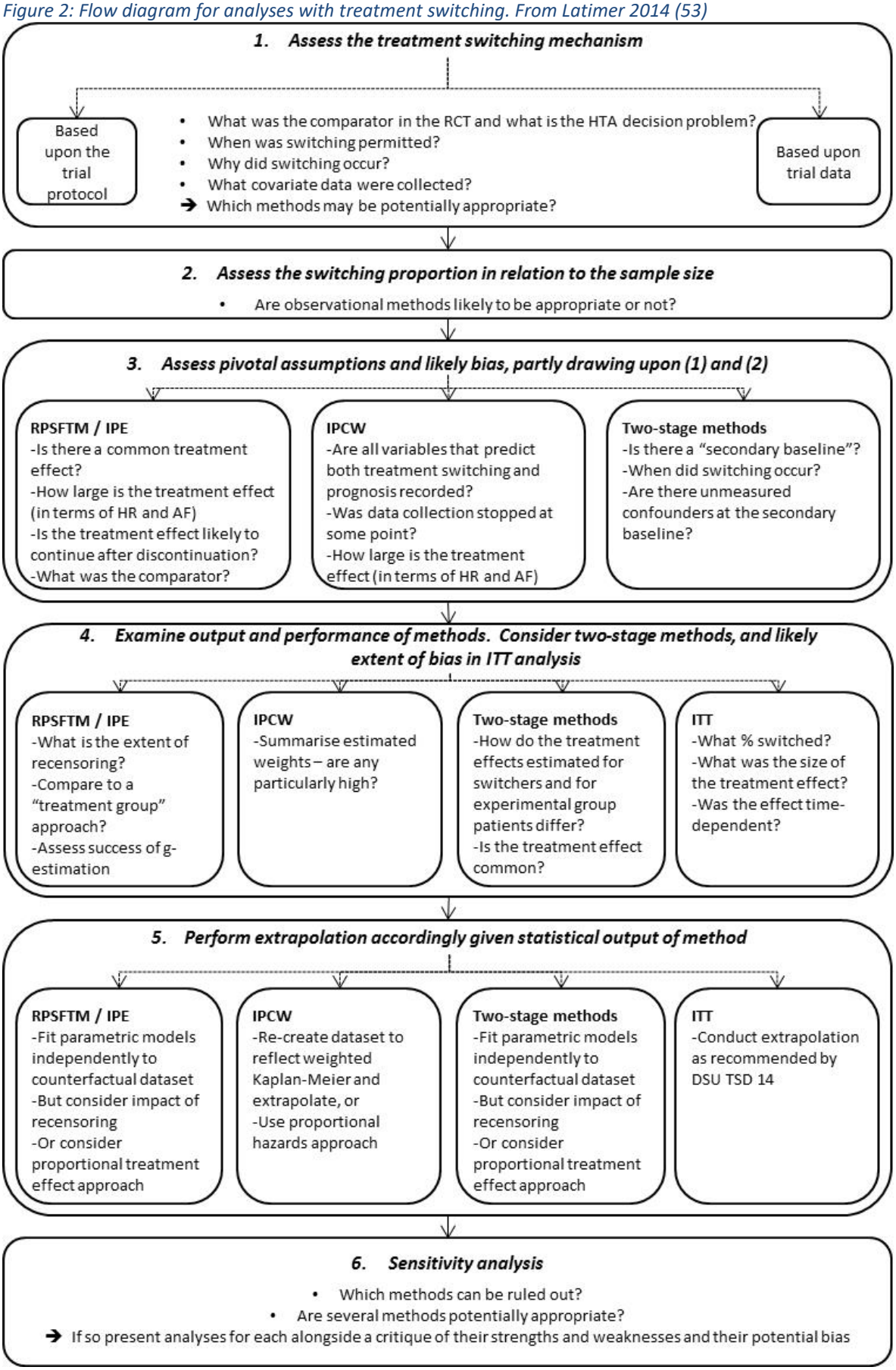
For ethical reasons, many controlled clinical studies allow patients in the control arm to switch over to the intervention arm or another active treatment at a given time point, often at progression of disease (treatment switching, crossover). The submitted documentation must justify permitting crossover, and indicate when the patient changed treatment.

For ethical reasons, treatment switching is relatively common in cancer studies. In such cases, switching will have an effect on the estimate for overall survival. Several correction methods can be used to estimate 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 particular 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 not adjusted for treatment switch, must always be submitted.

Analyses that have been corrected for treatment switch can be submitted. In such cases, 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 (53).

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

Figure 2: Flow diagram for analyses with treatment switching. From Latimer 2014 (53)

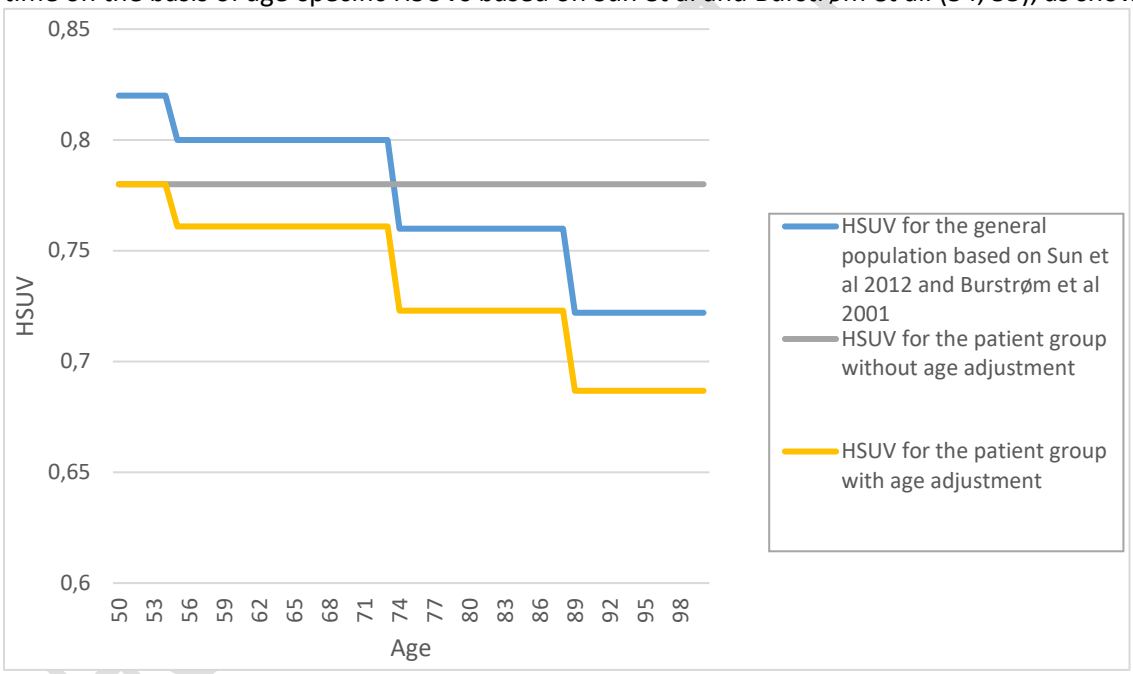


Appendix 3. Quality of life data

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 0.78 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 HSUVs based on the study by Sun et al (54), indicate a HSUV of 0.82 at age 50 years in the general population. Similarly, the HSUV at age 80 is 0.76 in the general population (see appendix 4.6). Without age adjustment one would then be using a higher health-related quality of life for a patient population over 80 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 0.78 multiplied by an adjustment index that is set to 1 at the start of the model. In this example, the index is reduced over time on the basis of age-specific HSUVs based on Sun et al and Burstrøm et al. (54, 55), as shown in



(refer also to appendix. 4.6). This is illustrated in figure 3, where HSUV for the general population based on data from Sun et al and Burstrøm et al. (54, 55) are represented by the blue line. The yellow and the grey line show HSUV for patients with and without age adjustment respectively.

Figure 3: Development of HSUV over time

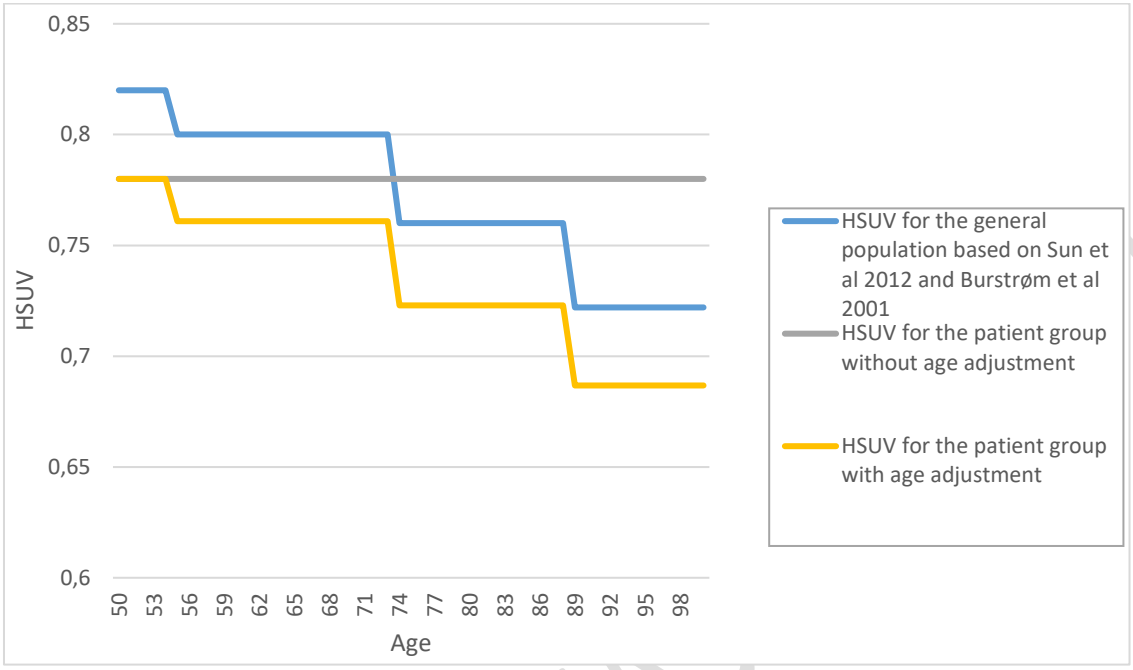


Table 4 ***Calculating age-adjusted HSUV***

Baseline HSUV for the patient group	Age	HSUV for the general population based on Sun et al 2012 and Burstrøm et al 2001	Adjustment index	HSUV for the patient group with age adjustment	HSUV for the patient group without age adjustment
0.78	50	0.82	1.00	0.78	0.78
	51	0.82	1.00	0.78	0.78
	52	0.82	1.00	0.78	0.78
	53	0.82	1.00	0.78	0.78
	54	0.82	1.00	0.78	0.78
	55	0.80	0.98	0.76	0.78
	56	0.80	0.98	0.76	0.78
	57	0.80	0.98	0.76	0.78
	58	0.80	0.98	0.76	0.78
	59	0.80	0.98	0.76	0.78
	60	0.80	0.98	0.76	0.78
	61	0.80	0.98	0.76	0.78
	62	0.80	0.98	0.76	0.78
	63	0.80	0.98	0.76	0.78
	64	0.80	0.98	0.76	0.78
	65	0.80	0.98	0.76	0.78
	66	0.80	0.98	0.76	0.78
	67	0.80	0.98	0.76	0.78
	68	0.80	0.98	0.76	0.78
	69	0.80	0.98	0.76	0.78
	70	0.80	0.98	0.76	0.78
	71	0.80	0.98	0.76	0.78
	72	0.80	0.98	0.76	0.78
	73	0.80	0.98	0.76	0.78
	74	0.76	0.93	0.72	0.78
	75	0.76	0.93	0.72	0.78
	76	0.76	0.93	0.72	0.78
	77	0.76	0.93	0.72	0.78
	78	0.76	0.93	0.72	0.78
	79	0.76	0.93	0.72	0.78
	80	0.76	0.93	0.72	0.78
	81	0.76	0.93	0.72	0.78
	82	0.76	0.93	0.72	0.78
	83	0.76	0.93	0.72	0.78
	84	0.76	0.93	0.72	0.78
	85	0.76	0.93	0.72	0.78
	86	0.76	0.93	0.72	0.78
	87	0.76	0.93	0.72	0.78
	88	0.76	0.93	0.72	0.78
	89	0.72	0.88	0.69	0.78
	90	0.72	0.88	0.69	0.78
	91	0.72	0.88	0.69	0.78
	92	0.72	0.88	0.69	0.78
	93	0.72	0.88	0.69	0.78
	94	0.72	0.88	0.69	0.78
	95	0.72	0.88	0.69	0.78
	96	0.72	0.88	0.69	0.78
	97	0.72	0.88	0.69	0.78
	98	0.72	0.88	0.69	0.78
	99	0.72	0.88	0.69	0.78
	100	0.72	0.88	0.69	0.78

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 use an 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 clinical experts. 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 QALYs 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 at the relevant age. We use the term $QALYs_A$ – short for remaining QALYs at age A. Use the remaining QALYs of the total population in the calculations, not 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 (56). This is combined with age-specific quality of life data to calculate quality adjusted remaining lifetime for different ages. Pending reliable Norwegian figures, the use of Swedish age-specific quality of life data is recommended, with value sets based on UK general population available for EQ-5D, based on Sun et al and Burstrøm et al (54, 55). We recommend using **Error! Reference source not found.** in appendix 4.6. The table shows the expected remaining quality adjusted life years according to age in the average population and is based on the sources listed above.

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 had the new treatment not been used, i.e., 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 0. 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 (e.g.,

Markov), it is normally useful to have a model-based estimate to ensure consistency between the different 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 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.

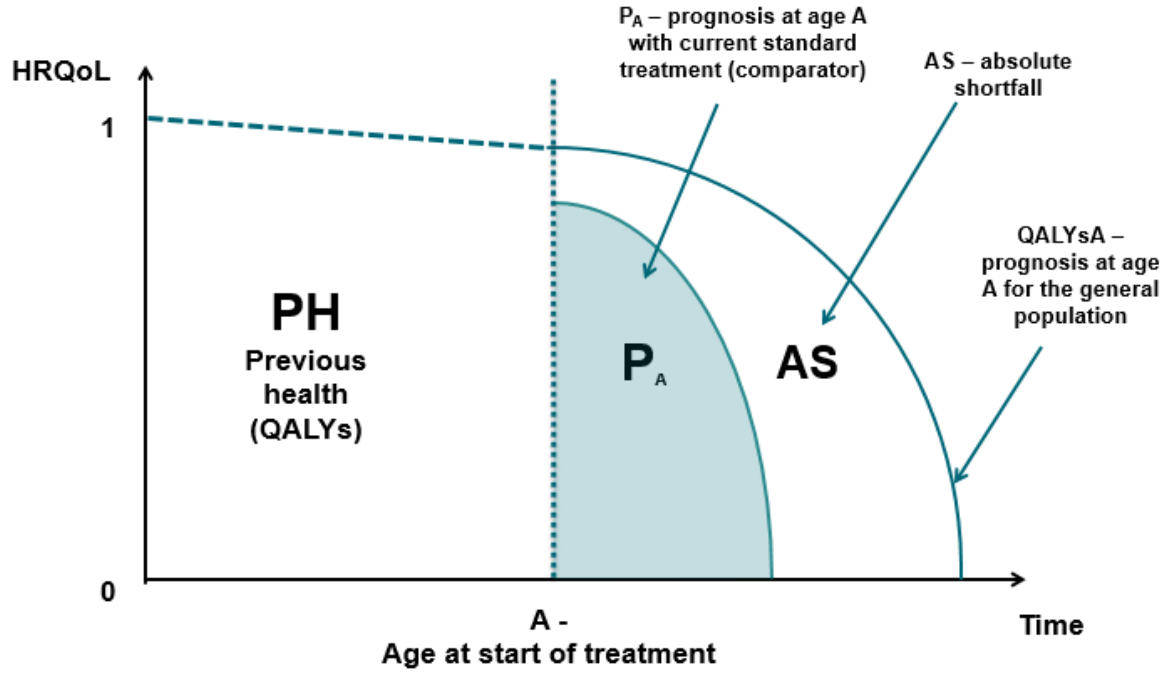
1. Age. The mean age at treatment start in the patient group relevant for treatment is 57 years, based on estimates by clinical experts. This is supplemented by data from national registries.
2. For a 57-year-old the calculated, expected remaining healthy life years ($QALYs_{57}$) is 21.4 QALYs. This is based on mortality data for the Norwegian population from Statistics Norway (56) and Swedish age-specific quality of life data, with British population-based value-setting tariffs (54, 55). See appendix. 4.6.
3. Prognosis. Patients have an expected remaining lifetime of 2.5 years, corresponding to 1.5 QALYs (undiscounted) with the current standard treatment (the comparator). This is based on simulations from the health economic model included in the submitter's documentation, after the Norwegian Medicines Agency has evaluated the documentation.
4. The absolute shortfall (AS) will then be 21.4 QALYs – 1.5 QALYs = 19.9 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, future health would be given by the area under the solid blue line from time point 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 considering future health loss in quantifying severity.

Note that the efficacy of the new treatment under evaluation is *not* included in the calculation of severity. Efficacy is included, however, 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, e.g., with a one-year perspective. This could be a chronic, non-fatal condition.

1. Age. The estimated mean age for treatment start in the relevant patient group is 50 years.
2. For a 50-year-old the calculated, expected remaining healthy life years ($QALYs_{50}$) is 26.7 QALYs. This is based on mortality data for the Norwegian population from Statistics Norway and Swedish age-specific quality of life data, with British population-based value-setting tariffs. See appendix. 4.6
3. Prognosis. The prognosis (undiscounted) in the health economic model analysis is 0.75 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, e.g., Sun et al (54), the HSUV for a 50 year old is 0.82. Assume also that the relative QALY loss caused by the disease is constant for the rest of life: Then the prognosis can be estimated as: Prognosis= (0.75/0.82) * 26.7 QALYs = 24.4 QALYs. In the calculation 26.7 QALYs is the expected remaining QALYs for a 50-year-old.
4. AS will then be 26.7 QALYs – 24.4 QALYs = 2.3 QALYs.

4.3 Example of calibrating two data sources – level adjustment

In some cases, the HSUVs for symptom-free states in the health economic analyses that 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 for which 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 12.5 QALYs. The AS is 12.5 QALYs – 3 QALYs = 9.5 QALYs.

The HSUV in the prognosis calculation will come from clinical studies of the intervention under evaluation or from other studies where the quality of life for the disease/condition has been measured, while the HSUV included in the quality-adjusted life years table come from another source. This must be taken into account as shown in our example:

In the health economic analysis the condition has a "symptom-free" HSUV of 0.85. This weight is used in the prognosis calculation. In the calculation of remaining $QALYs_{70}$ however, the HSUV for an average 70 years old, taken from Sun et al (54), is lower at 0.80.

This should be adjusted for by multiplying the prognosis estimate by the factor 0.80/0.85.

Thus the adjusted prognosis will be 3 QALYs* 0.80/0.85 = 2.8 QALYs. The adjusted absolute shortfall will then be 12.5 QALYs – 2.8 QALYs = 9.7 QALYs.

In this example, the adjustment did not lead to major changes in the calculated absolute shortfall. In other cases, it the impact can be greater. In general, when this adjustment has been used, submitters should consider whether the adjustment is reasonable.

4.4 Examples – calculating degree of severity for preventive measures interventions

4.4.1 New intervention to predict or prevent one type of disease.

1. Age. The new predictive diagnostic test or preventive intervention 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 life years ($QALYs_{60}$) is calculated as 19.3 QALYs. This is based on mortality data for the Norwegian population from Statistics Norway (56) and Swedish age-specific quality of life data (54, 55) (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 19.3 QALYs – 7.3 QALYs = 12.0 QALYs.

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

- From the time point that the disease manifests (time point A), not from the time point 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 or predictive test. The figure will thus refer to the sub-group who get the disease at a later time point (time point A).
- with prognosis based on the current standard *treatment* of the disease.

4.4.2 New intervention to predict or prevent two types of disease, A and B – Calculation of 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 receiving the intervention. 1.8 QALYs, i.e., 90 %, of the benefit is linked to prevention of disease A. 0.2 QALYs, i.e., 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 interventions to 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 interventions. Benefit and use of resources are included directly in a cost-effectiveness analysis for calculating the cost-per-QALY ratio of the intervention. 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 intervention. Severity is a consideration of *distribution or fairness* that is considered in addition to effectiveness. If an intervention 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/pharmaceuticals, cf. point 1 above). Such a weighting can be used whether it is the same patient group which has/will get several diseases or different groups which each have/will get one of the diseases.

4.6 Expected remaining QALYs in the general population

Table 1 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 (56) and the age-specific HSUV in the right hand column.

Pending reliable Norwegian figures, the HSUV from two Swedish studies have been used (54, 55). In the studies, Swedish age-specific quality of life data is combined with British population-based EQ-5D value-setting tariffs (24).

HSUV for the age group 21-73 years are taken from Sun et al (54), which is the most recent of the two Swedish studies and has the greatest number of respondents. In this publication, HSUV for other age groups are not presented. For the age group 0-20 years, we have assumed that HSUV are somewhat higher than for the age group 20-33 years. We have set it at 0.89.

In order to obtain fairly even age ranges, we have established an age group 74-88 years based on data from Burstrøm et al (55). For this group, we have calculated a simplified weighted average which gives a HSUV of 0.76 (rounded). The calculation is based on the following: For the age group 74-79 years we assume a HSUV at 0.79 based on Burstrøm et al (55). For the age group 80-88 years we use a HSUV of 0.74 from Burstrøm et al (55).

This gives a drop from 0.80 to 0.76 from the age group 55-73 years to the age group 74-88 years. We assume a corresponding (relative) drop from the age group 74-88 years to the last age group 89-105 years, to which we give a HSUV of 0.72.

Table 1: 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	69.1	0.89	36	38.0	0.85	72	11.3	0.8
1	68.3	0.89	37	37.2	0.85	73	10.7	0.8
2	67.5	0.89	38	36.3	0.85	74	10.1	0.76
3	66.6	0.89	39	35.5	0.85	75	9.5	0.76
4	65.7	0.89	40	34.7	0.85	76	9.0	0.76
5	64.8	0.89	41	33.8	0.85	77	8.5	0.76
6	63.9	0.89	42	33.0	0.85	78	8.0	0.76
7	63.1	0.89	43	32.2	0.85	79	7.5	0.76
8	62.2	0.89	44	31.4	0.85	80	7.0	0.76
9	61.3	0.89	45	30.6	0.82	81	6.5	0.76
10	60.4	0.89	46	29.8	0.82	82	6.1	0.76
11	59.5	0.89	47	29.0	0.82	83	5.6	0.76
12	58.6	0.89	48	28.2	0.82	84	5.2	0.76
13	57.7	0.89	49	27.4	0.82	85	4.8	0.76
14	56.8	0.89	50	26.7	0.82	86	4.4	0.76
15	56.0	0.89	51	25.9	0.82	87	4.1	0.76
16	55.1	0.89	52	25.1	0.82	88	3.7	0.76
17	54.2	0.89	53	24.4	0.82	89	3.4	0.72
18	53.3	0.89	54	23.6	0.82	90	3.1	0.72
19	52.4	0.89	55	22.9	0.8	91	2.9	0.72
20	51.6	0.89	56	22.1	0.8	92	2.7	0.72
21	50.7	0.87	57	21.4	0.8	93	2.5	0.72
22	49.9	0.87	58	20.7	0.8	94	2.3	0.72
23	49.0	0.87	59	20.0	0.8	95	2.1	0.72
24	48.2	0.87	60	19.3	0.8	96	2.0	0.72
25	47.3	0.87	61	18.6	0.8	97	1.9	0.72
26	46.5	0.87	62	17.9	0.8	98	1.8	0.72
27	45.6	0.87	63	17.2	0.8	99	1.6	0.72
28	44.8	0.87	64	16.5	0.8	100	1.5	0.72
29	43.9	0.87	65	15.8	0.8	101	1.5	0.72
30	43.1	0.87	66	15.1	0.8	102	1.5	0.72
31	42.2	0.87	67	14.5	0.8	103	1.3	0.72
32	41.4	0.87	68	13.8	0.8	104	1.1	0.72
33	40.5	0.87	69	13.2	0.8	105	0.8	0.72
34	39.7	0.87	70	12.5	0.8			
35	38.8	0.85	71	11.9	0.8			

Appendix 5. Reference case - health economics

The table below sums up by key words *some* of the requirements for health economic analyses in these guidelines.

Table 5 Reference case

Element in the analysis	***Standard analysis***	***Chapter in the guidelines***
Comparator	The treatment alternative(s) the new intervention is likely to replace	0
Analysis perspective	A form of extended health service perspective	9.2
Time horizon	Long enough that all the important future differences in costs and benefits between the interventions are captured	10.4
Analysis method	Cost-utility analysis (CUA)	0
Measure of benefit	QALY	1
Method for measuring benefit	Generic preference-based instruments (mainly EQ-5D-3L)	8.1
Method for valuing benefit	Population-based tariffs (mainly UK tariffs)	8.2
Value added tax (VAT)	Should not be included	9.2
Productivity changes as a result of the new pharmaceutical	Should not be included	9.2 og 9.3
Unrelated, future health service costs and savings	Should not be included	9.2
Marginal costs of public funds	Should not be included	9.2
Discounting	4 % per year for costs, benefit and life years.	9.4
Methods for dealing with uncertainty	One-way sensitivity analyses (shown in tornado diagram), multiway sensitivity analyses (mainly scenario analyses) and PSA	12.2
Degree of severity	Absolute shortfall	11