Norwegian Medical Products Agency

Submission guidelines

For Single Technology Assessment of Medical Devices

Valid from 2021-11-01

Updated 2025-07-01

Valid from 2021-11-1
Last update 2025-7-1

Foreword

The Norwegian Medical Products Agency (NoMA) is the sole health technology assessment (HTA) governmental body for single technology assessments (STA) of medical devices in Norway. NoMA took over this role from the Norwegian Institute of Public Health in January 2024.

NoMA is an agency under the Ministry of Health and Care Services of Norway. This document and its content reflect the principles described in White Paper 34 (2015-2016) (1), White paper 38 (2020-2021) (2) and White Paper 21 (2024-2025) (3). White Paper 21 is hereafter referred to as the priority-setting White Paper (*Prioriteringsmeldingen*). The purpose of conducting STAs in Norway is to address the three prioritisation criteria outlined here: benefit, resource use and severity, as well as uncertainty and budgetary consequences. Depending on the scope of the submission the prioritisation criteria are assessed in a quantitative (cost-utility analysis, CUA) or qualitative manner.

The Norwegian health care service consists of the primary and the specialist health service. NoMA performs STAs of medical devices that might be suitable for public funding by the regional health authorities (RHF) for specialist health care.

For health technologies that may be eligible for public funding through the regional health authorities, NoMA only conducts assessments following a commission from the Commissioning Forum for "New Methods", The National System of Managed Introduction of New Methods in the Specialist Health Care Service in Norway (*Bestillerforum for Nye metoder)*, a forum consisting of the medical directors from each of the four regional health authorities and two delegates from the Norwegian Directorate of Health.

Scope

This guideline details principles and requirements for submission of documentation for assessments of medical devices. We aim to ensure that assessments are timely, reliable, consistent, and relevant to the needs of decision-makers and key stakeholders. This guideline does not detail how assessments are conducted by NoMA.

This guideline is furthermore intended to assist health technology developers (HTD) in preparing documentation for economic evaluations. It should be used only for STA of medical devices by HTDs requesting public financing in the specialist health services. The guideline is not applicable for clinical decision-making on an individual patient level.

This document includes guidance on the type of documentation requested, as well as methodological requirements in calculating cost-effectiveness and performing budget impact analysis. Whereas the submission template specifies the requirements of the STA, the guideline serves as a reference for individual sections of the template as needed. The guideline is a living document and will therefore be reviewed and revised regularly.

Contents

Glossary of terms

Abbreviation	Definition
AFT	Accelerated failure time model
AIC	Akaike's Information Criteria
CER	Clinical Evaluation Report
CEAR	Clinical Evaluation Assessment Report
BIC	Bayesian Information Criteria
CCTR	The Cochrane Controlled Trials Register
CUA	Cost-utility analysis
DSU	Decision Support Unit
EQ-5D	EuroQol-5 dimensions
HR	Hazard ratio
HRQoL	Health-related quality of life
HSUV	Health state utility value (also referred to as QALY weight)
HTA	Health technology assessment
HTAR	Health Technology Assessment Regulation (EU) 2021/2282
HTD	Health technology developer
ICER	Incremental cost-effectiveness ratio
IFU	Instructions for Use
IPD	Individual patient data
ITC	Indirect treatment comparison
ITT	Intention to treat
IVDR	In Vitro Diagnostic Regulation (EU) 2017/746
KM	Kaplan-Meier
LYG	Life years gained
MAIC	Matching Adjusted Indirect Comparisons
MDR	Medical Device Regulation (EU) 2017/745
MeSH	Medical Subject Headings

Abbreviation	Definition
NICE	National Institute for Health and Care Excellence
NoMA	Norwegian Medical Products Agency
ML-NMR	Multi-level network meta regression
NMA	Network meta-analysis
OS	Overall survival
PER	Performance Evaluation Report
PFS	Progression free survival
PH	Proportional hazard
PICO	Patient population, intervention, comparator and outcome measures
PROs	Patient-Reported Outcomes
QALY	Quality Adjusted Life Year
RCT	Randomised controlled trial
RWD	Real World Data
SSB (KOSTRA)	Statistics Norway (Municipality-State-Reporting)
STA	Single Technology Assessment
STC	Simulated Treatment Comparisons
SSCP	Summary of Safety and Clinical Performance
VBA	Visual Basic for Applications

1. General submission requirements and information

The scope of the submission (completed template with accompanying documentation) should adhere to the commission from the Commissioning Forum for "New Methods", The national system of managed introduction of new methods in the specialist health care service in Norway (*Bestillerforum for Nye metoder*).

1.1 Pre-submission consultations

Pre-submission meetings between the HTD and NoMA are not mandatory but may be arranged upon request and are in most instances highly recommended. Guidelines for pre-submission meetings are available on the NoMA website. Health Technology Developers should notify NOMA, by email, of the planned submission date 2-3 months prior to submitting documentation for STA. This will enable NOMA to plan and allocate resources accordingly.

Once the documentation package has been submitted by the HTD, NoMA will review it within 15 working days of receipt. NoMA will subsequently inform the HTD whether the documentation meets formal requirements for submission. From the time a submission is deemed to meet formal requirements, NoMA has 180 calendar days to conduct the assessment. The fact that a submission is considered valid, does not however rule out the possibility of clarifications, supplementary and/or revised information being sought by NoMA at a later stage. If the HTD does not respond to such requests within a reasonable timeframe, NoMA may ask the Commissioning Forum to cancel the Commission.

A detailed log of the timelines in the process will be kept and presented in NoMA's report.

In addition, NoMA will consult clinical experts within the field relevant to the STA regarding the PICO, current clinical practice and any experience the clinicians might have with respect to the intervention or comparator. NoMA will also contact a patient representative via relevant patient organizations to seek information about their experience of living with the disease/conditions/impairment, as well as clinical outcomes that are relevant to the patients.

NoMA's report will be sent to the HTD for review to check for factual errors or information that should be exempt from public disclosure. The report will then be forwarded to the Commissioning Forum for New Methods for final approval before the report is published in its final form on the NoMA website.

1.2 Template

A submission template is available on the NoMA website. The HTD must use the template to prepare and submit documentation. The HTD is expected to submit appendices and/or additional information to supplement the template.

1.3 Radiation

The Norwegian Radiation and Nuclear Safety Authority (DSA) is the national authority and expert body in matters concerning nuclear security, radiation use, natural radiation and radioactive contamination in the environment. A medical device or method involving ionising or non-ionising radiation must comply with requirements stated in the Radiation Protection legislation as well as the Pollution Control legislation, if applicable. The DSA is part of the system of "New Methods" to ensure that the justification regarding the evaluation of new methods and applications in medical use of radiation on a general basis are documented before such methods and applications become available for general use. A method is justified if the total diagnostic or therapeutic benefit, for the individual and society, is higher than the

disadvantages related to the use of radiation. The radiation harm should therefore be included in the evaluation and the methods safety where this is relevant.

1.4 Responsibility

The entity responsible for preparing the submitted documentation, in addition to others who have been involved, must be stated.

1.5 Language

Documentation for Single Technology Assessment (STA) of medical devices must be submitted in one of the following languages: Norwegian, Swedish, Danish, or English.

1.6 Accessibility and universal design

In keeping with the European Union's Web Accessibility Directive, all public documents, websites and mobile applications must now be compliant with Web Accessibility Directive's requirement to ensure access for people who experience disabilities. MS Word (and other Microsoft software) provides an Accessibility Checker for identifying and repairing accessibility issues, which is located under the Review tab and Check Accessibility sub-tab.

1.7 Confidentiality

NoMA acts in compliance with the Public Administration Act and the Freedom of Information Act. Guidelines on handling confidentiality in relation to HTAs are available on the NoMA website.

1.8 Small patient groups

When preparing STA submissions for medical devices intended for very small patient groups with extremely severe conditions, the guidelines outlined in the "Arrangements for single technology assessment of pharmaceuticals for very small patient groups with extremely severe conditions" should be consulted as a supplement to these guidelines, as NoMA will use similar considerations for medical devices.

2. Population & intervention

2.1 Description of the disease/condition/impairment and Norwegian clinical practice

Briefly describe the relevant disease or condition targeted by the proposed intervention and how patients are currently diagnosed and treated in Norway. Ideally refer to national guidelines and to current Norwegian clinical practice. The HTD should specify any clinicians or key opinion leaders that it has consulted to confirm clinical practice.

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

2.2 Patient population

Describe as precisely as possible, the patient population in Norway for which the intervention is intended. Specify if the analysis covers only a portion of the intervention's indications/areas of use, or specific subgroups of the population. Identify the age group that is most affected by the disease or condition and state the mean age (or median age) of the relevant patient group in Norway. Confirm the mean (median) age by citing clinical experts, registry data or other relevant sources. If the intervention is a diagnostic test, these figures should, in general, reflect the population at the time of testing (see Chapter 13 for more details). If 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 predefined in clinical studies. Describe relevant diagnostic tests and methods used to select patients who are likely to be subject to the intervention.

2.3 The intervention

In this context, the intervention refers to any medical device that may be used to promote health, detect, prevent, or treat a disease/condition. Describe the following in table (see the submission template, Chapter 2):

- The name of the relevant device /equipment, diagnostic test, software version, and an overview of models approved for the European market, if applicable.
- Product type according to the Medical Device Regulation (MDR (EU) 2017/745) or In Vitro Diagnostic Regulation (IVDR (EU) 2017/746).
- The Unique Device Identifier (UDI-DI; according to MDR /IVDR).
- Medical device (MD) risk class according to MDR or IVDR.
- Name, number, and country of the notified body that issued the CE marking, with the date of first approval of the device/equipment for commercial use in the EU (CE marking) and the expiry date. (Copy of the CE certificate must be attached).
- Additional documentation to be submitted Summary of Safety and Clinical Performance (SSCP), Instructions for Use (IFU) in English and Norwegian, Clinical Evaluation Report (CER) for MD or Performance Evaluation Report (PER) for IVD, Clinical Evaluation Assessment Report, and if available Expert Panel Opinion or View.

Provide a description of the following:

- A product description including composition, technologies involved and technical characteristics.

- The expected lifetime of the device and how disposal will be handled.
- Any installation and/or maintenance services that may be required (how often and by whom?)
- A description of the development of the relevant device/equipment, if applicable (e.g., the time of first access to previous versions and the most significant changes made in the current version).
- Other equipment or consumables that must be used with the relevant device (e.g., biomarker testing, companion diagnostics, quantity and type of biological material required for IVD).

For medical devices based on decision systems incorporating machine learning processes (technologies within the scope of algorithms and/or artificial intelligence), the functions built or developed using these technologies should be described, and the questions in Chapter 10 of this guideline should be answered.

If the use of the device can result in dangerous or potentially dangerous levels of ionising and/or non-ionising radiation, this must be explained. The HTD should clarify what documentation is required in dialogue with the Norwegian Radiation and Nuclear Safety Authority (DSA).

For implantable equipment that can cause misinterpretations of MRI examinations, the HTD must describe the potential consequences.

For active implantable medical devices (AIMDs), specify the limits for compatibility with MRI procedures, and describe the main precautions that need to be taken. Where applicable, describe AIMD deactivation measures for conducting the tests.

If the clinical documentation for conformity assessment (CE marking) of a medical device is based on an equivalent device, this must be documented.

Include a description of planned Post Marketing Clinical Follow-up studies (PMCF) and other studies describing the effect of the relevant device (as detailed in Chapter 5.)

2.4 The intervention’s application

The documentation should include the following:

- Approved indication or application area, and any disease/condition relevant for the use of the method.
- Expected application area if the intervention is introduced into Norwegian clinical practice.
- If the intervention is already in use in clinical practice, this should be described (population, extent and time of introduction).
- The intervention’s placement in the treatment or diagnostic sequence.
- Describe the current patient flow in the application area.
- Expected patient flow if the intervention is introduced (preferably in a flow diagram).
- Proposed user competence and training of users, described in accordance with the Instructions for Use (IFU) and Summary of Safety and Clinical Performance (SSCP).

3. Comparator(s)

A health economic analysis in an STA, aims to quantify differences in efficacy/benefit and resource utilisation between the intervention and the relevant comparator(s).

3.1 Choice of comparator

The comparator should represent the alternative intervention most likely to be exclusively or partially replaced by the intervention under assessment.

In most cases, the comparator will reflect current standard of care (for example, according to national guidelines), or the treatment or health technology that is most extensively used in terms of number of patients. The comparator may be best supportive care, medical devices or medicinal treatments. In rare cases a relevant comparator could be no treatment, or a treatment or test sequence, provided robust data exists.

3.2 Multiple comparators

If there are several commonly used comparator treatments, or health technologies then include all relevant comparators in the assessment. Include as single comparators and do not merge two or more alternatives using, for example, average effects or costs. Such combinations will not reveal the cost-effectiveness of the intervention relative to each comparator.

3.3 Comparator(s) not previously assessed

If NoMA has not previously established the cost-effectiveness of a chosen comparator, an analysis against that comparator alone will, in most cases, not be sufficient to demonstrate cost-effectiveness of the intervention. Always provide an additional analysis against placebo/sham, best supportive care, or an alternative that can reasonably be assumed to be cost-effective.

If the comparator has been considered as established practice for an extended period of time, has documented efficacy for the population relevant to the STA, and has low associated costs, NoMA may accept it as the only comparator in the analysis.

3.4 Comparator(s) previously evaluated not cost-effective

If NoMA has previously concluded that the comparator is not cost-effective, but is still in use in clinical practice, always include an additional analysis as described in 3.3.

4. Outcomes

4.1 Efficacy outcomes

Outcomes applied in the health economic model should be clearly described, and any assumptions with respect to relative efficacy and duration of efficacy, and how this is modelled must be justified. Any lack of agreement between clinical documentation, health economic model input, and Norwegian clinical practice, must be highlighted, and the choices made for model inputs must be justified.

The effect estimate based on the primary outcome or the "hard" outcomes in the studies should be applied in the model. A hard outcome is an endpoint that is patient-important, well defined, and can be measured directly and unambiguously. Any reasons for not adhering to this must be justified in detail.

4.1.1 Surrogate outcomes

Intermediate outcomes[1], surrogate outcomes[2] or biomarkers, are used to predict clinical benefit in clinical trials and HTA processes in situations where direct measurement of clinically meaningful (final) outcomes are unavailable. A surrogate outcome may be a biomarker, or it may be an intermediate outcome.

If studies with surrogate outcomes have been submitted, it is important, but not sufficient, that the HTD establish that a surrogate lies on a causal pathway and is correlated with a clinical outcome in order to validate it. In addition, the HTD should demonstrate that modification of a surrogate without and with therapeutic intervention reliably modifies the clinical outcome (6). If possible, surrogate outcomes should be validated using multivariate meta-analytic methods, with an appropriate justification if not.

If intermediate outcomes, surrogate outcomes or biomarkers are used in the submission, the level of evidence should be established and supported by appropriate literature (Table 1). A surrogate outcome originating from an RCT (hierarchy level 1) is always the preferred source of evidence with the level of uncertainty increasing downwards in the hierarchy, from hierarchy level 1 to 3.

Table 1. Hierarchy of evidence for surrogate end point validity (7).

Hierarchy level	Requirement	Source of evidence
1	Treatment effect on surrogate corresponds to treatment effect on final outcome	Randomised controlled trials showing that changes in the surrogate are associated with commensurate changes in the final outcome
2	Consistent association between surrogate and final outcome	Epidemiological/observational studies
3	Biological plausibility of relation between surrogate and final outcome	Pathophysiological studies and understanding of the disease process

[1] An intermediate outcome is a surrogate outcome such as a measure of a function or of a symptom (disease-free survival, angina frequency, exercise tolerance) but is not the final outcome of the disease, such as survival or the rate of irreversible morbid events (stroke, myocardial infarction)4. EUnetHTA. D4.4 – OUTCOMES (ENDPOINTS). EUnetHTA; 2023.. Intermediate outcomes should be clinically meaningful for the patient.

[2] A surrogate endpoint is an endpoint that is intended to replace a clinical endpoint of interest that cannot be observed in a study - it is a variable that provides an indirect measurement of effect in situations where direct measurement of clinical effect is not feasible or practical5. ICH Expert Working Group. ICH Guideline E9: ICH HARMONISED TRIPARTITE GUIDELINE. International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use; 1998..

Literature to support/document the level of evidence should be acquired through a systematic literature review. Documentation of the literature searches and the relevant publications should be included in the dossier (see Chapter).

The uncertainties associated with the evidence should be described and if possible, quantified.

When a health economic model is submitted, the uncertainty associated with the relationship between the surrogate outcome and health-related quality of life and/or survival should be explored and quantified. Whether a surrogate endpoint is accepted as part of the STA will be assessed by NoMA in a case-by-case basis.

4.1.2 The learning curve

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 the learning curve. This curve reflects the extent to which the skills and experience of the operator of the medical device or diagnostic intervention will influence the total effectiveness. If the device 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 learning curve might include both user and community/institutional experience.

Describe potential factors affecting learning and how, if at all, the impact of the learning curve was accounted for in the clinical evidence.

The effect of the learning curve may be investigated through subgroup analysis or meta-regression, comparing studies with experienced and inexperienced operators (8). It is recommended that the HTD should clarify the effect of the learning curve in a scenario analysis in the health economic model, where the learning period is excluded (9). If there is a minimum patient volume required to maintain the operator's skills over time, the HTD should discuss this.

Incremental innovation

Frequent product modifications and upgrades may limit HTDs' ability to identify a "steady state" period during which it is appropriate to evaluate a medical device in a randomized controlled trial (10). It is acceptable to include evidence to support earlier versions of the device or diagnostic intervention considered, together with descriptions of any upgrades or changes regarding specific properties. The significance of these changes should also be discussed. Where relevant registry or real-world data (RWD) for the device under assessment should be included. A similar assessment, with respect to any anticipated further future changes, for devices currently under assessment in ongoing studies should be described.

4.2 Safety outcomes

An adverse event is defined for medical devices in MDR Art. 2(57): *"any untoward medical occurrence, unintended disease or injury or any untoward clinical signs, including an abnormal laboratory finding, in subjects, users or other persons, in the context of a clinical investigation, whether or not related to the investigational device*". Art.2 (58) further defines a 'serious adverse event' for medical devices and Art. 2 (59) device deficiency.

For IVD devices as defined in IVDR Art. 2 (60): "'a*dverse event' means any untoward medical occurrence, inappropriate patient management decision, unintended disease or injury or any untoward*

clinical signs, including an abnormal laboratory finding, in subjects, users or other persons, in the context of a performance study, whether or not related to the device for performance study."

Art.2 (61) further defines a 'serious adverse event' for IVD devices and Art. 2 (62) device deficiency.

Harms or adverse events may have an impact on patients' adherence, mortality, quality of life and resource use (11). There may also be risks to the operator associated with use of a device (12). Therefore, in accordance with the submission template:

- An overview of harms/adverse events as reported for both intervention and comparator in the studies must be submitted, as well as the studies' time frame (follow-up time).
- Moreover, the harms/adverse events of significance to the HTA in terms of frequency (absolute number of events and absolute number of patients and/or operators associated with each event), severity, and duration must be described.
- Further, describe the treatment of harms/adverse events in clinical practice and the health economic model (monitoring, follow-up, resource use and costs/disutilities).
- State why harms/adverse events have been included or excluded in the model. If there are critical levels of exposure or accumulated risk e.g. radiation, these should also be specified. The risks to the patient or operator associated with misuse of the device should be noted.

If the device/diagnostic intervention under assessment involves exposure to radiation, seek advice from the Norwegian Radiation and Nuclear Safety Authority (http://www.dsa.no/en/) for further guidance before submitting the assessment.

4.3 Quality of life outcomes

See Chapter 11.

4.4 Organisational implications

Describe, when relevant, additional changes to the health system necessitated by the proposed intervention. Examples could include a need for additional staff training, changes in management processes, the introduction of new patient administration or referral routines, changes to patient flow once the intervention is in use and changes in storage capacity. Describe any monitoring and/or quality assurance systems that need to be in place. Specify anticipated organisational implications as precisely as possible. If relevant, the impact on (particularly personnel) capacity should be described.

5. Information retrieval and selection of relevant documentation

The aim of the information retrieval and selection process should be to identify all studies relevant for the STA in question. To achieve comprehensive information retrieval, systematic literature searches in several databases and further information sources and search techniques are required.

5.1 Selection criteria

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

- A precise formulation of the research question (including specified PICOs)
- An a priori definition of the inclusion and exclusion criteria in table format, and reasoning behind their choice.
- Justification for the chose timeframe (i.e. the date limits if any are chosen)

5.2 Systematic literature search

Information retrieval requires systematic literature searches in multiple information sources and following good practice principles (13-16).

The submitted literature search must not be more than six months old at the time of submission.

The search strategy should be reported in such detail that it allows the search in all information sources to be replicated. Ideally, information pertaining to all 16 items on the PRISMA-S checklist (below) should be reported.

Information sources and methods:
1. Database name
2. Multi-database searching
3. Study registries
4. Online resources and browsing
5. Citation searching
6. Contacts
7. Other methods

Search strategies:
8. Full search strategies
9. Limits and restrictions
10. Search filters
11. Prior work
12. Updates
13. Dates of searches

Peer review:
14. Peer review

Managing records:
15. Total records
16. Deduplication

Below are a few special notes relating to some of the items on the check list. Please refer to Rethlefsen 2021 (17) for further details.

PRISMA-S items 1 and 3: mandatory and optional information sources for submission of documentation for technology assessments of medical devices to NoMA are listed below.

Mandatory sources:

- MEDLINE/PubMed
- EMBASE (if subscribed to)*
- US National Library of Medicine ClinicalTrials.gov
- WHO International Clinical Trials Registry Platform (ICTRP)

* If the HTD does not have a subscription to Embase, steps must be taken to compensate for the possibility of missing relevant records. Additional sources (check the list of optional sources) or one or more of the following supplementary search methods should be explored (PRISMA-S items 4-7):

- checking the reference lists of included publications, and maybe also selected excluded, almost relevant ones
- forward citation searching (using a tool for finding articles citing works included in the literature review)
- seeking advice from content experts
- searching PubMed/Embase using the identifier (registry numbers) of clinical trials in PubMed and/or a web browser

Optional sources to find secondary research:

- Epistemonikos
- International HTA database (INAHTA)
- Web pages of HTA organizations not registering their publications in the International HTA database

Optional sources to find primary and secondary research:

- Reference lists of included publications
- Cochrane Central Register of Controlled Trials
- Subject specific databases relevant to the condition or intervention (PsycINFO, Pedro, etc.)
- Lens.org
- Abstracts and proceedings from relevant conferences or annual meetings. Ask content experts for advice.

PRISMA-S item 6: NoMA recommends collating included publications known beforehand in a list and make sure they are all retrieved by the search strategy.

PRISMA-S item 8: Search strategies for ALL information sources should be copied and pasted exactly as run and made available in an appendix.

PRISMA-S item 12: If the original literature search is more than six months old, it must be updated in all information sources. Relevant studies identified in the updated search but not included in the analyses, should be listed in the appendix.

PRISMA-S item 14: NoMA does not require peer review of the electronic search strategies for submissions. However, we strongly recommend requesting NoMA's preliminary opinion on the suggested information retrieval process before moving on to screening/study selection.

5.3 Selection process

Briefly describe the steps in the selection (screening) process including information on whether one or more reviewers were involved and how disagreements were resolved.

Provide a flow chart, preferably aligned with PRISMA reporting guidelines (18, 19), summarizing the selection process (number of search hits in total and per database / study registry, number of duplicates, number of records excluded based on title and abstract screening, number of documents assessed in full-text, number of studies and study reports included).

A list of excluded studies after full text-screening detailing the reasons for exclusion should be given in the appendix.

5.4 Studies performed or sponsored by the HTD

Published and pivotal studies performed or sponsored by the HTD for the population in question must be submitted. Other documentation relevant for the STA, i.e., studies, data, analyses, other evidence and study reports, other data reported outside the study reports for which the HTD was a sponsor, must be included in the submission.

Information on ongoing or discontinued studies performed or sponsored by the HTD in the relevant population must be made available.

If unpublished data are submitted, the data should be described and appended (12). Provide the protocol and any intermediate findings for studies in progress. Provide the protocol and full study report (dated and signed) for completed studies, as well as any text submitted for publication (12).

5.5 Risk of bias assessment

The HTD should perform risk of bias assessment using validated instruments, for instance ROB2 and ROBINS-I (20, 21). Inclusion of such an assessment is encouraged.

5.6 Supplementary information used in the health economic analysis

The HTD should specify any studies not identified in the systematic literature search that have been used in the health economic model. This includes studies on types and distributions of subsequent treatment, studies used for transition probabilities, relative risks, odds ratios, hazard ratios or time to treatment discontinuation (TTD).

5.7 Patient/user experience

If qualitative data or other sources of information are available that can provide insight into user experiences associated with the intervention, NoMA recommends that the HTD describe these and provide references.

6. Documentation of relative efficacy by direct comparative studies

6.1 General considerations

This chapter describes requirements in cases where relative efficacy is established through direct comparative studies. The clinical evidence should be identified using a systematic literature search (information retrieval see Chapter 5 for details and requirements). When evidence synthesis is necessary, refer to Chapter 7 for guidance. If time- to- event data are employed, requirements regarding parametrisation and extrapolation of outcomes are outlined in Appendix 1.

To establish the relative efficacy and safety profile of a new intervention, the gold standard evidence is a RCT with low risk of bias. However, NoMA recognizes that, for various reasons (e.g. blinding or ethical issues), RCTs may be limited or not available. In such cases, observational studies may complement or even be the primary source of evidence for effectiveness and safety. Whether or not NoMA will accept observational data as a source of evidence will depend on the methodological quality of the study and the risk of bias assessment and will be determined on a case-by-case basis (22).

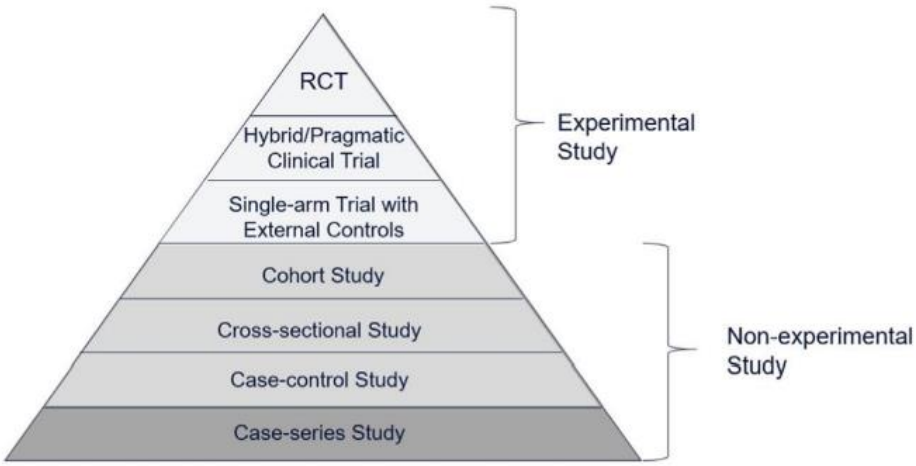

Figure 1. Pyramid level of evidence. There are other levels of evidence; for example, systematic review or meta-analysis could be on the top of the RCTs (23).

6.2 Reporting of pivotal studies and direct evidence

A detailed description of the pivotal study and other relevant studies for establishing direct relative efficacy, must be presented irrespective of design.

Study follow up time, treatment duration and dates for subsequent data-cuts must always be included in the submission. Always include data from the last available data-cut. Supportive and relevant ongoing studies should also be tabulated, as described in Chapter5.

7. Documentation of relative efficacy through evidence synthesis

In cases where evidence synthesis is the basis for relative efficacy, the PICO should still be described as outlined in Chapter 5.

7.1 Definitions

The following paragraphs describe terms used in this chapter (24, 25).

Exchangeability refers to the fundamental assumption for evidence syntheses, which implies that if patients from one treatment group were substituted into another, the same treatment effect is expected. The term encompasses similarity, homogeneity, and consistency.

Indirect comparison refers to any evidence synthesis in which treatment groups from different studies are compared.

Network meta-analysis (NMA) is a generalisation of meta-analysis to include more complex evidence networks and can include both direct evidence and indirect evidence.

Population-adjusted method for indirect comparisons refers to methods for indirect comparisons where individual patient data (IPD) from one or more studies are used to adjust for relevant population characteristics that differ between studies with the aim to estimate relative treatment effect.

7.2 General considerations

This section outlines the requirements for evidence synthesis methods used to combine multiple sources of evidence to estimate comparative effectiveness and/or safety (e.g., a pairwise meta-analysis [direct comparison], indirect comparison or network meta-analysis) between the intervention and comparator(s).

Efficacy and safety data from RCTs are preferred although data from studies of other designs may be accepted. All evidence included in the evidence synthesis must be selected following a systematic review, conducted, and reported in accordance with PRISMA guidelines (26). The analysis, reporting, and interpretation of the results from evidence synthesis analyses should adhere to good practice principles

A full systematic literature search must be carried out and the search described in detail as outlined in Chapter 5. For literature searches intended to support the documentation of relative efficacy, PICO must be considered. The studies included in the evidence synthesis should reflect the established PICO.

A table outlining the following must be submitted along with an evidence synthesis for each of the studies discussed:

- Overview of study design (phase, randomisation, blinding, condition)
- Definition of endpoints
- Statistical analysis (including estimand[3] (27), and how intercurrent events and missing data were addressed)

[3] An estimand is a precise description of the treatment effect reflecting the clinical question posed by a given clinical trial objective.

- Dates of the study recruitment period
- Duration of follow-up
- Treatment duration
- Reasons for and proportion of censored observations
- Countries covered by the studies
- Types and distributions of any subsequent treatment received in the studies

Furthermore, any other factors that might differ between the studies and that can affect the treatment effect must be tabulated.

Risk of bias assessment using validated tools (e.g., ROB2, ROBINS-I) shall be included irrespective of the statistical approach used for the evidence synthesis (20, 21).

7.3 Assumption of exchangeability

Exchangeability (similarity, homogeneity, consistency) is the fundamental assumption required for evidence synthesis based on meta-analysis and network meta-analysis (28-30).

The assumption of similarity requires studies to be comparable with regards to possible effect modifiers across all treatments included in the evidence synthesis. The following must be evaluated to identify potential effect modifiers:

- List of potential effect modifiers, defined a priori in the protocol, related to study and patient characteristics, intervention and comparator characteristics, outcome characteristics, and observed values of relevant outcomes at baseline, especially in the comparator arms.
- Describe the methodology applied to identify effect modifiers, which should include a comprehensive review of the literature and consultation of healthcare professionals in the disease area.
- Describe the magnitude and direction of the interaction effect of the identified effect modifiers.
- Describe likely missing effect modifiers in one or more of the included studies and their potential effect.
- Conclude whether the assumption of similarity is expected to hold and provide the reasoning that supports the conclusions.

The assumption of homogeneity requires that there is no meaningful heterogeneity in the effect estimates from the studies included in each direct comparison, requiring at least two direct comparisons and at least five for reliable assessment (31, 32). It is important to note that heterogeneity can be caused by effect modifiers which are unknown, and factors initially considered similar or not initially considered as effect modifiers. Include the following the for assessment of homogeneity:

- Forest plots of the comparisons.
- Statistical tests for heterogeneity.
- Conclusion with reasoning on whether the assumption of homogeneity is expected to hold, and the studies can be pooled.
- Decision on the choice of model for inference, i.e., random vs. fixed, should be guided by the sampling frame and not solely rely upon statistical tests. Both random-effects and fixed-effect models should be submitted.

The consistency assumption can only be tested in networks of evidence where both direct and indirect evidence exists for a particular contrast (i.e., a loop structure), and it is thus an assessment of the between-study variation of different treatment contrasts (28, 33, 34). Include the following for the assessment of consistency:

- Identify loops in the network where consistency can be evaluated.

- Provide the method used to test for inconsistency with a justification and the criteria used for the conclusion on whether the consistency assumption can be considered violated.
- Results of the tests and the models used with interpretation and uncertainty in the results.
- For effect modifiers that are possible sources of inconsistency, a sensitivity analysis should be conducted to assess the magnitude of effect modification, including the direction of modification where applicable.
- Conclusion with reasoning on whether the assumption of consistency is expected to hold.

If any one of the assumptions concerning exchangeability is violated, a description of how the violation was handled including complete results of all analyses, must be submitted (34-38).

7.4 Statistical methods for evidence synthesis

Requirements for the statistical methods applied in evidence synthesis are described in the following paragraphs. Submit a description of the software used, along with codes and scripts.

7.4.1 Direct comparisons

Direct comparison also means one single comparison within a single study (see Chapter 6). Direct comparisons of treatments described in this section include two or more studies comparing the same intervention against the same comparator in a pair-wise meta-analysis and the HTD should include the following:

- Description with justification of the chosen statistical model (frequentist or Bayesian) and method for the meta-analysis and heterogeneity parameter.
- Tabular presentation of individual study results, including the number of patients per study.
- Forest plots with point estimates and confidence intervals. If a Bayesian analysis has been performed, credible and predictive intervals should be reported.
- In the case of Bayesian analysis, information on the nature of the prior distribution for all relevant parameters and evaluation of the impact of the prior distribution on the results of the analyses must be reported.

Selected methodological references (24, 25, 39-52).

7.4.2 Indirect comparisons

Methods for indirect comparisons include the Bucher method, Bayesian or frequentist network meta-analysis, and anchored population-adjusted methods.

Provided that the assessment of exchangeability is deemed to be fulfilled (Section 7.3), include the following:

- Full statistical description of the chosen model.
- Description of the appropriateness of the chosen method according to the evidence base.
- Graphical and tabular presentation of the evidence network with number of studies per comparison.
- Tabular presentation of number of patients and individual study results for the endpoints included in the evidence base.
- Forest plots and tabular presentation of the relative effect estimates with measure of uncertainty for the intervention versus all comparators included.
- Presentation and assessment of the results from both direct and indirect evidence if present in the network.
- In the case of Bayesian analysis, the number of iterations for burn-in and estimation, the number and convergence of Markov chains, and information and justification for the choice of prior distributions for all relevant parameters.

- In the case of inconsistency models within the Bayesian framework, plots of posterior mean deviance for both the model and the inconsistency model.
  If the submission includes indirect comparisons of time to event data, please refer to NoMA's guideline for submission of STAs for pharmaceuticals Section 7.4.3 (53).

Selected methodological references (24, 25, 28, 54-61).

7.4.3 Population-adjusted methods

The most common population-adjusted methods are matching-adjusted indirect comparisons (MAIC) and simulated treatment comparisons (STC) (62). Multi-level network meta regression (ML-NMR) is currently restricted to data that are not time-to-event (63). Population-adjusted methods can be used when the similarity assumption is considered violated.

In cases of anchored comparisons (connected evidence), the HTD should provide a justification that bias in the analyses will be reduced by using population-adjustment methods. Include the following:

- Complete description of the method and model for population adjustment and treatment effect estimation.
- Evidence that one or more of the observed patient characteristics are effect modifiers, and that the distribution of those effect modifiers is imbalanced enough to result in bias in the relative treatment effect.
- Method for identification and selection of effect-modifying covariates which are adjusted for, and the potential for un-measured effect modifiers.
- Potential for residual bias in the results and the size and direction of the bias.

Regression-based methods (STC, ML-NMR) fit an outcome regression model applied to the available individual patient data. Include the following:

- Model fit and justification of the appropriateness of the chosen model.
- Description of the degree of covariate overlap between the studies included.
- For STC: the method used to estimate outcome and the treatment effect that is targeted by the chosen approach.
- For ML-NMR: whether the data are sufficient to handle treatment-covariate interactions, if the shared effect modifier assumption is required, and whether any other assumptions have been made for model estimation.

In a MAIC, the patient population for the intervention is weighted to the aggregated data for the comparator (target) population. Include the following:

- Distribution of effect modifiers before and after weighting in the intervention population versus the comparator population and description of the impact of any residual imbalance.
- Distribution of weights and effective sample size after matching.
- Method for confidence interval estimation; the robust or bootstrapping methods are preferred.

Unanchored MAICs and STC scan be applied in disconnected networks. These analyses compare absolute values of outcome, and to obtain an unbiased estimate all factors affecting the outcome must be adjusted for. The points described above also apply for unanchored analyses and, in addition, the following must be submitted:

- Method for identification of all relevant prognostic variables.
- Description of whether the unanchored indirect comparison is appropriate with regards to which data are available for adjustment, outcome definition, and comparability of study characteristics.

- The extent of missing prognostic variables and the direction and size of bias from the lack of adjustment for prognostic factors.

In unanchored analyses where a prognostic factor for the intervention is of unknown distribution in the comparator study and hence unadjusted for, the results have a bias of unknown size and direction (24, 25, 62, 64). The results of such analyses might be rejected by NoMA as a source for quantification of relative efficacy in the STA.

Selected methodological references (25, 62-66).

7.4.4 Non-randomised evidence

Comparisons can be based upon non-randomised evidence including single-arm studies, cohort studies, case-control studies, other observational studies, and historical controls from different types of registries. In non-randomised evidence, there is high risk of confounding bias which can violate the underlying assumption of exchangeability. The following must be submitted for all the included evidence:

- Comparison of the inclusion and exclusion criteria.
- Comparison of baseline patient characteristics.
- Method used to identify prognostic and effect modifying factors and method used to adjust for confounding.
- Plausible direction of bias which arises from missing prognostic and effect modifying variables.
- Risk-of-bias assessment using an appropriate tool.

Propensity score methods can be applied to adjust for observed confounding in comparisons based on non-randomised data. These analyses, however, require complete access to individual patient data, and whenever used, the following assumptions must be addressed and reported:

- The positivity assumption (having individuals in all treatment groups for every combination of covariates) by assessing patient eligibility/contraindications and propensity score distribution.
- The overlap assumption (a practical manifestation of positivity) using histograms or density plots in the whole, matched, and/or weighted populations.
- The balance assumption before and after matching, weighting, or stratification by the absolute standardised difference.

Multivariate regression methods can also be used to adjust for observed confounders, but propensity score methods and multivariate regression can only adjust for observed variables, not unobserved confounders.

Results should be presented for crude and adjusted analyses with appropriate sensitivity analyses for evaluation of robustness.

Selected methodological references (20, 24, 25, 62, 67-74).

7.4.5 Extrapolation of data

Extrapolation of time-to-event data is covered in Appendix 1. For extrapolation of other endpoints, the HTD should justify the assumptions made and sources used to model long-term efficacy beyond the study period and the transitions between health states in the model.

7.5 Treatment switching

For ethical reasons, RCTs may allow patients in the control arm to switch to the intervention arm or another active treatment at a given time point. This will often occur at a time of disease progression, it

is referred to as treatment switching, and the analyses that has been corrected for the treatment switch can be submitted. In such a case, provide:

- A description of the switching mechanism, i.e., timepoint when the patients switched treatment.
- The number and proportion of patients who switched treatments.
- Baseline patient characteristics for both patients who switched and patients who did not switch treatment.
- Response status of patients before switching (response/no response/partial response etc)
- An intention to treat (ITT) analysis (or the relevant primary analysis if there is no ITT) with an estimate without adjustment for the treatment switch.
- A detailed description of the applied adjustment method with justification of why the underlying method-specific assumptions are fulfilled or not.
- An explanation for why other adjustment methods were not applied.
- A discussion of the assumptions, strengths, and weaknesses of the various methods (83).

8. Diagnostic interventions

8.1 Introduction

Diagnostic tests are used to inform clinical decision making, for example by predicting which individuals may benefit from a specific medical treatment. When the use of a diagnostic test is linked to a treatment decision, the evaluation of the diagnostic test should ideally be supported by studies that follow patients from testing via treatment to final clinical outcome, so-called end-to-end studies (75). If end-to-end studies are available, they must be submitted to NoMA then clinical and relative efficacy reported according to Chapter 6 and Chapter 7 in this guideline. If end-to-end studies are not available, it is acceptable to provide separate evidence for test characteristics, analytical validity, clinical validity, and treatment outcomes and to demonstrate how they are linked and estimated in the health economic model (see also Section 12.9).

8.2 Diagnostic test accuracy

If the intervention is a diagnostic test, the submission should include information on:

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

The analytical validity, clinical validity, and clinical treatment outcomes (clinical utility) of a test should be demonstrated (76).

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 turnaround time, amount of biological material needed (if applicable), ease of interpretation of the test, if the test is qualitative or quantitative, training and equipment needed to perform the test. Also describe characteristics that may be important to the patient or operator but are not captured by the test outcomes (e.g. feasibility, risk of adverse events, comfort).

Provide information on the decision rules/algorithms and whether they are in the public domain. Specify parameters such as the reference standard, prevalence, and test results in terms of sensitivity and specificity for each study. The internal validity and applicability of included diagnostic accuracy studies should be critically appraised using an appropriate instrument such as QUADAS-2 (77). If the diagnostic intervention involves artificial intelligence, also refer to the questions listed in Chapter10.

Special considerations regarding resource use and health economic modelling on diagnostic intervention are mentioned in Section 12.9 in this guideline.

9. Use of Real-World Data (RWD)

Real-world data (RWD) is a comprehensive term that refers to various types of data generated within the context of routine healthcare (e.g., electronic health records, administrative healthcare databases, clinical and disease registries, patient-generated data, and data produced from various other sources) that can inform on health status (78).

Real-world evidence (RWE) refers to the clinical evidence of a health technology or medical condition derived from the analysis of RWD for a specific research question (78).

RWD can be employed to generate RWE for various purposes within HTA such as epidemiology, demographics, treatment duration, usage of a health technology in real-world setting, and assessing the effectiveness and/or safety of health technologies (e.g., for new indications of already-used technologies or for documenting long-term follow-up).

In cases where RWD have been used by the HTD to quantify relative efficacy, data should be identified, analysed, and evidence developed through transparent and reproducible approaches. The HTD should characterize potential bias and uncertainties in the data source and describe the relevance for clinical practice in Norway. NoMA recommends using a framework or template for this purpose, such as the NICE Real-World Evidence Framework (79). The Framework provides comprehensive guidelines for the reporting and use of real-world data: detailed reporting is essential to enable independent replication and critical appraisal. This includes documenting the study design, analytical methods, data curation processes, patient characteristics, follow-up details, and handling of biases. Reporting should also cover the assessment of the generalizability of findings to the target population. The framework highlights the need for analytical methods that minimize bias, including selection, information, confounding, and external validity biases. It also suggests using sensitivity analyses to assess the robustness of findings. The Framework outlines several methods to address bias in studies using real-world data:

1. **Selection Bias**:
    - **At Study Entry**: Use methods like matching, restriction, or stratification to ensure comparable groups. Address prevalent-user bias, lead time bias, immortal time bias, and depletion of susceptible.
    - **At Study Exit**: Address informative censoring through techniques such as inverse probability of censoring weights.
2. **Confounding**:
    - **Identification and Selection of Confounders**: Use systematic approaches like directed acyclic graphs (DAGs) and expert opinion. Avoid overadjustment and control for relevant time-varying confounders.
    - **Analytical Methods**: Employ methods like multivariable regression, propensity score matching, stratification, or weighting. Use advanced methods when necessary to handle complex confounding scenarios.
3. **Information Bias**:
    - **Measurement Error and Misclassification**: Evaluate and adjust for measurement errors, especially differential misclassification, to avoid biased estimates. Use methods like calibration to correct errors.
    - **Missing Data**: Address missing data using methods like multiple imputation or sensitivity analysis to explore different missing data mechanisms.

4. **External Validity Bias**:
    - **Assessment and Adjustment**: Compare the study sample to the target population using methods like matching and weighting to ensure representativeness. Use statistical models to standardize predictions to the target population.
5. **Sensitivity and Bias Analysis**:
    - Conduct sensitivity analyses to test the robustness of results against various assumptions and potential biases. This includes exploring the impact of unmeasured confounders and alternative data assumptions.
6. **Quantitative Bias Analysis**:
    - Use methods to estimate the potential impact of biases on study outcomes. Techniques like the E-value can assess how strong a confounder would need to be to nullify the observed effect.

These methods are critical for minimizing the risk of bias and ensuring the validity and reliability of findings in real-world evidence studies.

An alternative source is the STaRT-RWE (Structured Template and Reporting Tool for real World Evidence (74). The template includes sections for administrative information, study design, study population, data sources, and analytical methods. It encourages thorough documentation of all relevant parameters, including data linkage processes, operational definitions, and the handling of confounding factors.

10. Artificial intelligence (AI)

A separate section has been dedicated to artificial intelligence in these guidelines as the literature suggests that HTA of devices incorporating AI may differ from that of other devices. For example, Alami 2019 (80), Hendrix 2022 (80), and Belisle Pippon 2021 (81) note that issues such as complexity, generalizability across various settings, autonomy, continuous learning and the black box systems associated with AI mean that it warrants specific attention.

Haverinen (2019) (82) has developed a framework known as digiHTA employed by the Finnish HTA agency FINCCHTA. The following questions have been drawn from the digiHTA framework (82), and should be answered by the HTD if the medical device incorporates artificial intelligence/machine learning processes:

1. Which platforms and platform versions of the product are available?
2. Does the use of the product require registration or login?
3. What kind of product support does the company offer?
4. What are the intended user groups?
5. What kind of support does the end user need to use the product?
6. If users need training, who organizes it? When? What is the language of training?
7. Does the company have instructions (e.g., a project plan) for healthcare service providers to ensure fluent introduction of the product?
8. What is the company's testing process?
9. What is the company's process for handling error messages?
10. Does the company have the capacity to roll back to previous versions of the product?
11. Does the company have a process to proactively monitor the running of systems and system components to automatically identify faults and technical issues?
12. Does the company have a plan for decommissioning the product?
13. Has there been any downtime or impairment time in the use of the product during the last six months?
14. How often must devices or software versions related to the product be renewed?
15. Does the product provide benefits to the end users by improving their behaviour related to their own health? How so?
16. Does the product provide benefits to the organization (like improving care processes)? How so?
17. What is the company's process to handle adverse events?
18. Has the product undergone a risk analysis?
19. Are there any undesirable effects associated with misuse of the product?
20. Have all user groups been taken into account in product design, like people with visual or hearing impairments?
21. Has the product been tested with real user groups?
22. What kind of accessibility testing has been performed on the product?
23. Does the product have interfaces into the website or other software?
24. Are proprietary formats used to store and transfer data?
25. Are the definitions of the original proprietary formats openly available?
26. Does the product have interfaces for other companies' services?
27. Can the data contained in the product be exported in a commonly used or standard format?
28. Does the product use data from other systems via interfaces?
29. If yes, can the data produced by others be separated in the system?
30. What is the classification of AI? Visualization only, AI-assisted (e.g., diagnosis / classification / decision), or solely AI-controlled?
31. Could the problem be solved without the AI solution?
32. Is the solution based on machine learning or a neural network?
33. Do the staff have sufficient capacity to understand the operational logic of AI (e.g., do they need additional training)?
34. Are the conclusions and decisions of the AI solution transparent, i.e., can medical staff understand what the decisions are based on?
35. Is the AI solution validated in the environment in which it will be used?
36. What are the data sources for the AI solution?

37. Are the data sources used in the training of AI solutions relevant to a final use case (e.g. are the age and gender composition of training groups comparable to that of real user groups)?
38. Are the access rights required for the use of the data in order, and have data protection (e.g., GDPR) and security issues been taken into account?
39. When it comes to classifier teaching, are there enough data relative to the size of the smallest class?
40. Can the AI solution use incomplete data?
41. Can the AI solution use noisy data?
42. Is retraining possible for the AI solution?
43. What are the data sources for retraining?
44. How is it ensured that the system is not taught with irrelevant data?
45. What performance criteria are used?
46. Does the AI solution change care processes? How?
47. When does the AI solution propose an action? How, and who will actually implement it?
48. Is staff's approval needed for action proposed by the AI?

11. Documentation of health-related quality of life (HRQoL)

The health benefit in STAs is quantified using quality adjusted life years (QALYs) and should ideally be based on patient-reported EQ-5D measurements which are used to estimate health state utility values (HSUV) in the health economic model. Whenever EQ-5D measurements of health-related quality-of-life (HRQoL) are collected in the pivotal clinical studies for the intervention in question, these must be submitted and included as an option in the model. Use of HRQoL data from the literature must be supported by a systematic literature search (Chapter 5), and the choice of sources must be justified and discussed.

11.1 Instruments

HRQoL data for use in health economic analyses must be measured using generic preference-based measuring instruments, preferably EQ-5D-5L to enable comparison between therapeutic areas and STAs (83-85). Any disease-specific instruments used should be reported as supplementary information.

The EQ-5D-5L should be used to measure HRQoL in patients 12 years or older (83).

For measurement of HRQoL in children 8 years or older, both EQ-5D-Youth-3L and ED-5D-Y-5L (86-88) can be used, for which tariffs are currently being developed (89). Average age, age distribution, and age range of the respondents, must be submitted (90).

Where EQ-5D or other generic preference-based measuring instruments are deemed inappropriate for capturing certain aspects of a disease or the patients' health-related quality of life, for example when coping/adaptation ("mestring") is affected, this must be supported by empirical data as well as by a qualitative discussion. For guidance consult NICE DSU Technical Support Document 8 (91).

Figure 2 below depicts the hierarchy of preferred sources for health-related quality of life data (92).

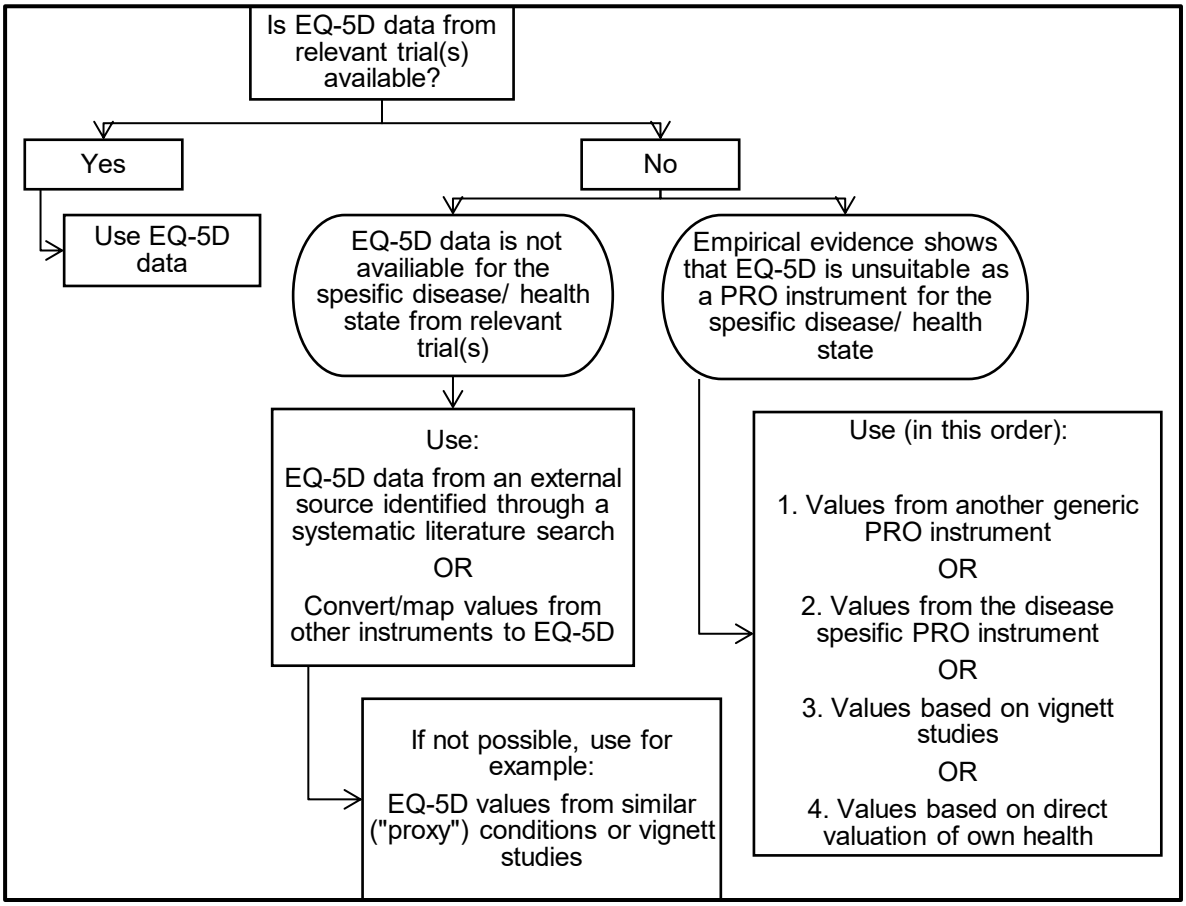

Figure 2. Hierarchy of preferred sources for health-related quality of life data (modified from (92))

11.2 Valuing

To ensure consistency across STAs and between measures of severity and health economic analyses in a STA, HRQoL are valued using tariffs derived from the general population and relevant to the Norwegian adult population. The Norwegian population-based EQ-5D-5L tariff (93) must be applied.

Use of an experienced-based tariff should be justified providing explanation of the differences between this tariff and a general population-based tariff.

If HRQoL values have been collected using the 3L version of EQ-5D questionnaire, use the EQ-5D-3L tariff from the UK general population from Dolan based on 3L for valuation (94), and use the Norwegian 3L population norms from Stavem et al. (95) for severity calculation, available on NoMAs website.

11.3 Mapping

Other generic preference-based instruments such as SF-6D, 15D, HUI, AQoL, and QWB can be used in instances where patient-reported EQ-5D data are not submitted, this should be justified and discussed with Noma.

If only disease-specific instruments are available, mapping must be used to predict EQ-5D values accompanied by a transparent description of the method. For details, refer to NICE DSU Technical Support Document 10 (96).

11.4 Age adjustment

Age-adjusted HSUV are used in the calculations of expected remaining QALYs for the general population (Chapter 13), hence such adjustments ensure consistency with calculation of severity.

Lack of age adjustments must be justified.

11.5 Level adjustment

When the best health state exceeds the utility weight of the general population (see online Excel spreadsheet), utility values should be adjusted using a multiplicative method. This approach preserves the relative differences between stages while preventing utility values from exceeding those of the general population. The method corresponds to the adjustment DMP applies to utility values for age over time in models.

Without adjustment:

- Health state A from study: 0.91
- Health state B from study: 0.73
- General population: 0.87 (age 50)

With adjustment:

- Health state A from study: 0.91 * (0.87 / 0.91) = 0.87
- Health state B from study: 0.73 * (0.87 / 0.91) = 0.69
- General population: 0.87 (age 50)

The adjustment is applied only to utility values derived from the same source. If external sources are used for subsequent or worse health states, these are not adjusted using the same value. If the external sources also report utility values exceeding those of the general population, they should be adjusted in a similar manner but using their own baseline value.

The adjustment, the inclusion of external utility values, and the associated uncertainty (internal/external validity) should be discussed.

If there are very small differences (≤0.01) between utility values for health states from the study and the general population, one may consider not making adjustments.

The calibration of absolute shortfall (AS), as previously practiced when utility values exceeded normative values, is obsolete.

11.6 Treatment specific health state utility value (HSUV)

If treatment specific HSUV are used, these must be comprehensively justified and documented. Treatment specific HSUV can only be considered when differences in HRQoL are documented in clinical studies and have a clinical rationale.

11.7 HRQoL for caregivers

If an intervention impacts HRQoL for caregivers, the effects for the caregiver may be included in health economic analyses quantified as QALYs. The inclusion of HRQoL for caregivers require the same level of evidence as that required for HRQoL for patients. Include the following:

- Description of the condition in question.
- Description of the condition's specific effect on caregiver's HRQoL.
- Description of the intervention and the comparator's effects on caregivers' HRQoL.
- Analysis with and without effect on caregivers' HRQoL.

For ethical and methodological reasons, any effects of increased life expectancy for patients on caregivers' HRQoL should not be included.

12. Health economic analyses

The health economic model must be designed such that all relevant scenarios can be explored. The selection of the time horizon, population (or sub-groups), parametric model for time-to-event data, and all other variables used in the model must be modifiable and not restricted to a single analysis.

12.1 General aspects

The choice of health economic model must be well-justified, and should aim to be as simple, straightforward, and transparent as possible, while comprehensive enough to capture all the relevant factors that could impact a decision. The model should provide an adequate picture of the condition being modelled and the course of the condition. The model's construction, assumptions, potential restrictions, and how input data were modelled must be fully documented and described. Increasing model complexity entails stronger requirements for model transparency.

The model should ideally be consistent with both the clinical documentation that establishes relative efficacy, and Norwegian clinical practice. The model framework must be both internally and externally validated. Internal validity refers to the precision and consistency of the model's calculations. External validity entails checking the model's results against independent sources, such as epidemiological studies or other relevant data.

If global models are used, they must be adapted to reflect Norwegian clinical practice, epidemiological data, costs, and other potentially relevant aspects. It must be clearly stated how the model has been adjusted to fit a Norwegian setting. Any lack of adaptation must be justified with a description of its potential impact on the results of the analysis.

12.2 Specific requirements

- Microsoft Excel or TreeAge is the preferred software for HTA submission. Contact NoMA prior to submission if alternative software is used.
- A description of function and purpose of each sheet in the spreadsheet must be provided.
- If the Excel-model uses macros, provide a description of the macros used. Visual Basic for Applications (VBA) code should include brief description for all procedures. Make sure all macros run successfully before submitting.
- Spreadsheets for Excel-model must be transparent and fully user modifiable.
- Spreadsheets must not contain password protected sheets or cells, or hidden cells, or use proprietary or non-transparent programmes and/or programming language.
- All sheets for Excel-model should include visible headers by default. Do not hide rows, columns, or sheets that are not used. If sheets (or parts of sheets) are not used but cannot be easily deleted (due to offset functions, macros, etc.), this should be clearly stated/highlighted.
- Provide Markov-traces and /or other plots that can aid reviewers to critically appraise the modelling assumptions.
- When time-to-event analysis techniques are applied, present KM-curves and all extrapolations within the same graph, within the health economic model.

12.3 Analysis methods

12.3.1 Cost-utility analysis (CUA)

The recommended method for health economic evaluations submitted to NoMA is CUA. The result of such an analysis is an incremental cost-effectiveness ratio (ICER) expressed as cost per QALY gained.

12.3.2 Cost-minimisation analysis

Cost-minimisation analysis is appropriate when the HTD can provide documentation to support that the efficacy and safety profiles for the intervention and the comparator are, or can be assumed to be, equal or non-inferior.

12.4 Treatment sequence modelling

Sometimes it may be appropriate to model treatments as part of a sequence. For example, the order in which diagnostic tests are carried out could have a significant impact on clinical outcomes. High-quality documentation on the relative differences in efficacy between different treatment sequences is necessary.

12.5 Analysis perspective

Benefits and costs that must/must not be included in the analysis, in accordance with the priority-setting White Paper guidance, is outlined below. Benefits and costs related to the medical device under evaluation can either result from the intervention, or be expected to change, due to the intervention. The following guidance implies an extended health-service perspective.

The following benefits must be included (if applicable):

Effects on:

- The patient's lifespan.
- The patient's HRQoL.

The following benefits may be included (if applicable):

Effects on:

- The caregiver's health-related quality of life. Results of the analyses must be presented both with and without inclusion of this effect.

The following costs must be included (if applicable):

- Treatment or prevention costs, paid by the health service or by the patient/caregiver.
- Transport costs related to travelling to and from treatment, whether paid by the health service or by the patient/caregiver.
- Patients' and their caregivers' use of time during patient treatment.

In accordance with the Priority-setting White Paper, the following is not to be included:

- Productivity changes resulting from the intervention.
- Consequences of patients' future use of public services and receipt of public benefits/pensions.
- Unrelated health service costs and savings.
- Tax expenses for public financing.
- Public benefits, pension payments, value-added tax, and other transfer payments.

12.6 Time horizon

The time horizon must be sufficiently long to capture all important future differences in costs and health effects between two or more alternatives. If the medical device influences mortality, the time horizon should be based on the patient's expected lifetime. If there is no documentation for, or it is unlikely, that

the relative efficacy will be sustained, a shorter time horizon may be more appropriate but should be justified.

12.7 Resource use and costs

Resource use refers to the consumption of goods, services, time, and potential capital cost. Whenever possible, market prices in the private sector should serve as the basis when estimating unit costs or calculation prices (97). Unit costs and resource use are to be presented and justified separately. Norwegian unit costs must be used. If calculations are performed in currencies other than Norwegian kroner, the exchange rate used must be clearly stated.

All assumptions and justifications for included costs must be thoroughly documented and reported in detail to ensure transparency and facilitate assessment.

When calculating device costs, the expected retail price without value added tax (VAT) must be used. It must be possible to change the price of the medical device within the model to perform analyses using discounted prices.

If the model includes medicinal products, maximum pharmacy retail prices excluding VAT should be used and it should be possible to change the prices in the model. It is preferable to model all medicinal costs per package, e.g., the complete package collected from the pharmacy. If wastage and/or vial sharing is included in the health economic model, the calculations must be transparent and editable.

Transportation costs associated with travel to and from the treatment site must be included in the analysis if relevant, while necessary transportation costs for the caregiver can also be considered if relevant and well-documented.

The NoMA cost database includes relevant sources for unit costs and is intended to ensure a consistent approach for inclusion of unit costs in STAs and health economic models.

If unit costs are not calculated based on Norwegian cost databases/publications, they can be sourced from other cost studies/publications. If costs are estimated based on prices from another country in another year, they may be converted by first adjusting local costs to the price year in the model using the local consumer price index, and then applying the average exchange rate for that year.

Average costs can generally be used, and “standard” calculations for average cost per resource type (e.g., visits to doctors, hospital treatment, nursing home costs, laboratory services) can be undertaken for instance:

- **Hospital services:** The cost per hospital admission or outpatient clinic attendance can be calculated by multiplying the DRG-points by the corresponding unit price providing an estimate of the total costs per admission/attendance. Co-payments made by patients for outpatient consultations can be disregarded as it is already accounted for through the DRG weighting. If no information on the relevant DRG code is available, cost per day or per consultation can be obtained from the Norwegian Directorate of Health’s database (SAMDATA) for the specialised health services (covers somatic, mental health services and multi-disciplinary specialised addiction treatment). More information on activity-based financing, DRG-weighting, unit prices, and SAMDATA can be found on the Norwegian Directorate of Health’s website. Apply values from «Kostnadsvekt», not «DRG Basispoeng». The cost weights and the value of a cost weight are updated annually in relation to changes in medical practice and changes in operating conditions.

- **Nursing home, general practitioner, specialist services, imaging diagnostics, and laboratory tests:** The Norwegian Medical Association provides an overview of relevant tariffs, patient contributions, and subsidies on its website. A database (Unit cost database/ "*Enhetskostnadsdatabase*") published on the NoMA website contains most of the relevant costs incurring at nursing homes, in the general practitioner's office, specialist health care, imaging diagnostics and laboratory tests. The database should be used as the main source of unit costs.
- **Other costs:** If unit costs are included and they are not collected from the NoMA cost database, the source and a justification must be included in the submission. If the costs are meant to reflect costs incurred in the general practitioner's office or by a specialist, the cost per contact/consultation/unit is calculated by multiplying the remuneration by a factor of two. This is a pragmatic approach to account for both remuneration and public subsidy, as well as other overhead costs. The calculation gives a rough estimate, however multiplying the remuneration amount by two is considered to give a better cost estimate than using the tariff rate directly.

Average costs estimates will encompass both fixed and variable costs. However, in certain instances, one or more treatment options under consideration may result in additional capital costs. Such costs should be explicitly highlighted and included into the analysis. Alternatively, when treatment options are not anticipated to impact fixed costs, it may be more relevant to solely consider the variable costs.

12.8 Capital costs and fixed medical equipment

In order to incorporate capital assets, such as hospital-based medical equipment, in the model, it is necessary to calculate the cost of each session of patient use. This involves estimation of several variables, such as the equipment's expected lifespan depreciation rate, the patient volume over a given period, and the overhead cost rate. Maintenance and removal costs should also be included. The capital costs should be annuitized over the equipment's lifespan. The straight-line method of depreciation should be used.

NoMA 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 Drummond et al 2016 (98). 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 (75).

12.9 Considerations for diagnostic interventions

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

12.9.1 Linked evidence approach

The link between intermediate and final outcomes should be supported by documentation (75). Ideally, this should involve end-to-end studies that follow patients from testing to treatment outcomes. As such studies are rare it may however be necessary to employ a linked evidence approach (75). This involves creating a model that combines data from various sources, including:

- Diagnostic accuracy studies: These studies assess the performance of the diagnostic test itself, such as its sensitivity and specificity.

- Treatment effectiveness studies: These studies evaluate the impact of different treatment options on patient outcomes.
- Clinical guidelines and expert opinion: These sources provide information on the typical care pathway for patients with the condition being diagnosed, including the sequence of tests and treatments.

12.9.2 Costing

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, estimate the average costs of the indication for which the STA is submitted. Some fixed costs may be assigned to other indications given proper justification (75).

If the introduction of the technology requires additional investment in infrastructure or personnel, these costs should be incorporated in the analysis.

12.9.3 False results

The costs and outcomes associated with false diagnostic results should be included in the health economic model. The NICE Diagnostic Assessment Programme Manual (75) notes that "All health benefits (or harms) resulting directly or indirectly from the use of the diagnostic tests (including both the true and false results) should be included."

The NICE Manual notes that regarding false positives may have the following impacts:

- Delayed treatment of another cause: Patients with false-positive results may have another cause for their symptoms. Discovery of that cause may be delayed by the false-positive result with a reduction or delay in the benefits of treatment for that cause.
- Unnecessary treatment: Patients may receive treatment for a condition they do not have, leading to potential side effects, complications, and additional costs.
- Psychological impact: False-positive results can cause anxiety and stress in patients who believe they have a serious condition.
- Increased healthcare utilization: Patients may undergo further unnecessary tests and procedures to confirm or rule out the initial false-positive diagnosis.

The model should account for the potential negative consequences of false positives, such as the factors listed above, and the associated costs and impacts on health outcomes.

The NICE manual describes the impact of false negatives as follows:

- Some patients will not receive the treatment or have treatment delayed until further symptoms appear.
- Disease progression: The underlying condition may worsen due to the lack of timely treatment, potentially leading to more severe complications and poorer outcomes.
- Reduced quality of life: Patients may experience a decline in their quality of life due to the progression of the undiagnosed condition.
- Increased healthcare utilization: Patients may require more intensive and costly treatments in the future due to the delayed diagnosis.

The model should account for the factors listed above, and the associated costs and impacts on health outcomes.

12.10 Organisational implications

12.10.1 Organisational changes

Potential organisational implications are mentioned in Section 4.4 of these guidelines. Any organisational changes that can lead to variation in parameter estimates must be described, justified and incorporated into the health economic model. If relevant, describe and estimate investments in staff and/or patient training, and any additional required physical infrastructure If it is likely that there will be an implementation (transient) period where both the intervention and the comparator co-exist at the same time this should be accounted for in the model (99). State if the intervention is likely to be more cost-effective in one setting than another (31). If there are several organisational 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 organisational arrangements that are necessary for maintenance, (emergency) extraction and handling (at the end of the device's lifespan).

12.11 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.

12.12 Time and unit costs for patient and caregiver

If the intervention and comparator have different duration of treatment administration and/or travel time, differences must be accounted for in the health economic analysis.

NoMA recommends that the value of time for patients and caregivers is calculated as a common rate, regardless of employment situation, and assigned the value of leisure time. The value of increased/decreased leisure time is determined by the net real wages in Norway (100).

Any changes in time for work and/or other daily activities/leisure time owing to the health technology/medical device (productivity gains) must not be included.

12.13 Projection of unit costs

Unit costs are generally kept unchanged throughout the modelled time horizon, due to uncertainty surrounding future technological and market developments. However, if there are good/compelling reasons to project changes in unit costs, a description and justification must be provided.

12.14 Present value and discounting

Comparisons of benefits and costs that occur over different time periods in a CUA must be converted to present value by discounting both the annual benefits and costs by the applicable rate provided in the Priority-setting White Paper (1), which states that the discount rate should be equivalent to the rate set by the Ministry of Finance. As per Rundskriv R-109 (100), the discount rate should be 4% annually for the first 40 years of the time horizon (years 0-39), 3% annually for the next 35 years (years 40-74), and thereafter 2% annually (year 75 and onwards).

13. Severity

13.1 General aspects

In HTA, severity must be quantified using absolute shortfall (101). Absolute shortfall is the reduction in expected future healthy life years without the treatment under consideration (i.e., with the current standard treatment), expressed in terms of QALYs lost compared to the general population of the same age. Absolute shortfall is required in cost-utility analyses, but cost-minimisation analyses do not require a shortfall calculation.

The quantification of absolute shortfall is performed in stages and is described in detail below. Absolute shortfall differs between treatment interventions and preventive measures. Comorbidity is discussed separately below.

When presenting results, any uncertainties in the estimation of absolute shortfall, such as age or prognosis uncertainties, should be described.

An Excel spreadsheet ("
Tools for severity calculation and age adjustment") for severity calculation and age adjustment is available on NoMA's website.

13.2 Quantification of absolute shortfall

Step 1 - Age

Define the mean age at start of treatment for the relevant Norwegian patient group considered for the new treatment. If the group has a substantial age disparity, consider including the median age in addition. Ensure consistency between the age used in the severity calculations and the health economic model, as a general rule. If there is considerable uncertainty or conflicting age estimates from different sources, consider using an age range/interval, noting the likely position of the mean or median within the interval.

Sources for mean age estimation can be registry data, study data and/or clinical expert information. Utilise the source that best represents the relevant population in Norway.

Step 2 - Expected remaining QALYs for the general population

Estimate the remaining undiscounted QALYs for an average person from the general population with the age identified in Step 1. This is referred to as quality adjusted expected remaining lifetime from the relevant age ($QALYs_A$).

Mortality data for the Norwegian population from Statistics Norway is recommended for calculation of expected remaining lifetime at different ages (102, 103). This is combined with age-specific quality of life data to calculate quality adjusted remaining lifetime for different ages. Refer to the Excel spreadsheet, which displays the expected remaining quality-adjusted life years according to age in the average population.

Step 3 - Prognosis

Calculate the prognosis for the relevant patient population at treatment initiation. The prognosis refers to the average number of remaining healthy life years for the patient group with the current standard treatment (P_A = prognosis at age A). Hence, prognosis is calculated for the current standard treatment

(comparator) and is measured in QALYs. It is calculated for the remaining lifetime for the patients, taking into account the expected number of undiscounted QALYs patients can anticipate with the comparator treatment. If there is no active treatment, select a patient population for prognosis calculation according to the guidelines for comparator choice in health economic analyses, such as best supportive care or no treatment (Chapter 3). When health economic calculations are based on a lifecycle model, a model-based estimate is preferred to ensure consistency. Alternative sources may include relevant clinical studies, registry data, or data from systematic literature searches.

Step 4 - Absolute shortfall

Use undiscounted values for $QALYs_A$ and P_A to quantify absolute shortfall = $QALYs_A - P_A$

13.3 Calibration through level adjustment

The basis for calculation of prognosis is the HSUV. In cases where HSUV are higher than those of the average population, HSUV in the health economic analysis must be calibrated through level adjustment, see Appendix 2 for an example. The HTD should consider and justify the reasonableness of applying adjustment.

13.4 Preventive measures

Preventive measures target both diseases and conditions, but in the following text for simplicity the term diseases are exclusively applied.

Take into consideration that all individuals in the group will not contract the disease, and that there is a time lag between prevention and disease onset.

Quantification of severity for preventive measures must:

- be applicable for the disease that is prevented,
- be applicable for the patient group contracting the disease in the absence of the preventive measure,
- be measured from average disease onset.

A description of quantification of severity for preventive measures is provided below, with examples provided in Appendix 3.3.

13.4.1 Prevention of a single disease

Determine which individuals/patients from the relevant population to include in the calculation. Severity is calculated from the average prognosis and absolute shortfall for those expected to be affected by the disease, targeted by the preventive measure under the current standard prevention.

13.4.2 Prevention of multiple diseases

Calculate the absolute shortfall for each of the diseases for the relevant population with current preventive practice (the comparator in the health economic analysis), as described in 13.4.1 above.

Calculate weighted absolute shortfall for the diseases assigning largest weights to conditions with highest health benefit (gained QALYs).

If 90% of the benefit, measured in QALYs, can be attributed to the prevention of disease A, then the absolute shortfall for disease A should be weighted at 90% in the weighted absolute shortfall.

13.5 Interventions which treat multiple diseases

When an intervention impacts multiple conditions, the principle for quantifying severity equals the procedure for medical devices which have preventive effects on multiple conditions as described in Section 13.4.2.

13.6 Diagnostic interventions

If the diagnostic intervention has one, or a few, indication(s), absolute shortfall should, as a rule, be calculated applying the treatment intervention principle described in Section 13.2. The relevant prognosis is thus associated with the population being given the test, not those who have the disease(s) tested for. If a diagnostic intervention may potentially be used for several indications, such as a CT-scanner, it may not be feasible to calculate absolute shortfall. The process of establishing an appropriate absolute shortfall in such cases should be performed through consultations with NoMA.

13.7 Sequelae and adverse reactions

For medical devices targeting a primary condition, the assessment and calculation should focus on the overall severity of the primary condition and symptoms directly related to the primary condition.

For medical devices aimed at symptoms resulting from the primary condition (known as sequelae), the severity of the sequelae must be evaluated and quantified, rather than the primary condition. For interventions targeting adverse reactions, it is the severity of the adverse reaction, not the primary condition, that must be evaluated and calculated. However, in certain instances, quantifications of severity for sequelae can also include the severity of the primary condition if (104):

- Sequela(e) and the primary condition are strongly related clinically.
- Wording in the indication of the medical device links it to the primary condition.
- Sequela(e) are specific for the primary condition in question.
- The mechanism of action for the treatment of sequela(e) is specific for patients with the condition in question.

For interventions targeting symptoms that are not related to the primary condition, the severity of the symptoms, not the primary condition, must be evaluated and calculated.

14. Uncertainty

Uncertainty in health economic analyses must be explored and discussed (98, 105).

14.1 Addressing uncertainty in the analyses

Uncertainty in health economic analyses must be examined and presented through sensitivity analyses. The impact of uncertainty on the analysis outcomes must be discussed, highlighting the following:

- factors that are drivers of uncertainty
- whether uncertainty can be reduced
- if and when additional data that may reduce uncertainty can be expected
- presence of any bias
- how changes in parameters or assumptions affect the results

14.2 Deterministic sensitivity analysis

Methodological and structural uncertainty, as well as uncertainty related to generalisability, must be analysed using deterministic sensitivity analyses.

In deterministic sensitivity analyses, selected variables are adjusted to examine the sensitivity of the model outcomes to these changes. This type of analysis is conducted in the form of one-way, two-way, or multiway sensitivity analyses and scenario analyses.

Anticipated changes in costs over time If the price of a technology or the cost of its implementation are likely to increase or decrease due to the wider distribution of the equipment or technical skills (acquired in the learning phase), these anticipated changes should be examined in a sensitivity analysis. Where the patent for a healthcare product is nearing expiry, the foreseeable price fall should be examined in a sensitivity analysis (12).

Scenario analyses must also be performed to assess the impact of alternative values for specific sets of parameters on the model outcomes. These analyses represent a base case, “worst case”, and “best case”, or other alternative plausible scenarios. See example in Appendix 1.3.

14.3 Probabilistic sensitivity analyses

Probabilistic sensitivity analyses may be employed to address the impact of joint parameter uncertainty. Justify the choice of the included variables and their probability distribution. Ideally, the probability distribution of each variable and its key moments (usually the expected value and standard error) should be based on empirical data. When empirical data are lacking, a plausible probability distribution must be selected for the variable. For each type of variable, there are usually only a few relevant probability distributions for use in probabilistic sensitivity analyses (98).

If probabilistic sensitivity analyses are submitted, the results must be presented as a scatter plot of the simulated incremental cost-effectiveness ratios (ICERs).

15. Budget impact analysis

The manner in which the budget impact analysis (BIA) should be carried out depends on the type of medical device under consideration, i.e. whether it is a single user (e.g. wearable or implantable) device or a multi-user device (equipment used by several patients over time, e.g. a CT scanner located in a hospital). Specific requirements are described in the following paragraphs.

Use one of the two provided budget impact templates for BIA, either single user or multi-user. The budget impact analyses must include:

- Expected number of patients treated with the intervention for each of the initial five years.
- Expected number of patients treated with the comparator(s) for each of the initial five years.
- Expected retail price including value added tax (VAT) and excluding discounting.
- A scenario where the intervention under review is approved for public financing.
- A scenario where the intervention is not approved for public financing.
- The difference between the two scenarios for each of the initial five years.
- All relevant costs affected by the approval of an intervention for public financing.
- Costs considered negligible and omitted must be discussed and justified.

The BIA must clearly present the expected budgetary consequences for the Regional Health Authorities.

The budget impact is split into three parts:

- Consequences for the Regional Health Authorities budget for medical devices/products.
- Budgetary consequences for the Regional Health Authorities overall.
- Budgetary consequences for the health care sector overall.

Both templates provided allow for BIA that are not based on a cost-effectiveness model. In addition, by providing reasonable justification (e.g. organisational complexity) the HTD can present the budget impact using other methods than those recommended here (contact NoMA for guidance). The default time horizon in the budget analyses is five years. However, it may vary depending on the economic lifespan and/or depreciation of the technology. Provide a justification if the time horizon deviates from five years.

The table from the BIA templates should be presented in the main dossier, along with a discussion of the uncertainty associated with the results.

15.1 Budget impact analysis for single-user medical devices

If the medical devices are intended for use by a single person only, which in most cases means that it is portable, wearable or implantable, use the template "bia template devices single user". Enter the expected number of patients eligible for treatment initiation each year, the expected retail price (including VAT) and market share of the intervention and comparator. Further, enter the expenditure per patient from the health economic model's Markov trace for devices, the specialist health service and the remainder of the health sector. The budget impact is calculated as the difference between a scenario in which the device is approved for public financing, and a scenario in which it is not.

15.2 Budget impact analysis for multi-user medical devices

The BIA template for multi-user or shared- devices is similar to the single user devices template with one exception: the costs associated with purchasing a device unit is to some extent independent of the number of patients and is entered in a separate sheet (Sheet 2.1.2). Costs for devices should normally

be amortized and the annual costs taken from the health economic model's Markov trace. The costs, i.e. non-device costs for the specialist health services and other costs are calculated per patient as in the "single user bia template". In Sheet 2.1.2, enter the number of shared devices likely to be purchased each year over the next five years. Also enter the number of new comparator devices, if the comparator is not a device, use the single user template Sheet 2.1 for comparator costs and adapt the multiuser Sheet 2.4 accordingly. Costs of comparator devices already purchased are not considered.

16. References

1. Helse- og omsorgsdepartementet. Meld. St. 34 Verdier i pasientens helsetjeneste. 2015-2016.
2. Helse- og omsorgsdepartementet. Meld. St. 38 (2020–2021) Nytte, ressurs og alvorlighet — Prioritering i helse- og omsorgstjenesten. 2021.
3. Helse- og omsorgsdepartementet. Meld. St. 21 (2024-2025) Rettferdig prioritering i vår felles helsetjeneste. 2021.
4. EUnetHTA. D4.4 – OUTCOMES (ENDPOINTS). EUnetHTA; 2023.
5. ICH Expert Working Group. ICH Guideline E9: ICH HARMONISED TRIPARTITE GUIDELINE. International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use; 1998.
6. EUnetHTA. EUnetHTA 21 – Individual Practical Guideline Document D4.4 – OUTCOMES (ENDPOINTS) 2023 [Available from: https://www.eunethta.eu/wp-content/uploads/2023/01/EUnetHTA-21-D4.4-practical-guideline-on-Endpoints-v1.0.pdf.
7. Ciani O, Buyse M, Drummond M, Rasi G, Saad ED, Taylor RS. Time to Review the Role of Surrogate End Points in Health Policy: State of the Art and the Way Forward. Value in health : the journal of the International Society for Pharmacoeconomics and Outcomes Research. 2017;20(3):487-95.
8. EUnetHTA. Therapeutic medical devices: Guideline. European Network for Health Technology Assessment; 2015.
9. National Health Care Institute. Guideline for economic evaluations in healthcare Diemen: National Health Care Institute; 2024 [updated 16.01.24. Available from: https://english.zorginstituutnederland.nl/about-us/working-methods-and-procedures/guideline-for-economic-evaluations-in-healthcare.
10. 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.
11. CADTH. Guidelines for the Economic Evaluation of Health Technologies: Canada. Canadian Agency for Drugs and Technology in Health; 2017.
12. HAS. LPPR: Dossier submission to the Medical Device and Health Technology Evaluation Committee (CNEDiMTS). Saint-Denis: Haute Autorité de Santé; 2020.
13. NICE. 5 Identifying the evidence: literature searching and evidence submission. Developing NICE guidelines: the manual (PMG20): National Institute for Health and Care Excellence; 2014/2024. p. 92-107.
14. Lefebvre C, J G, Briscoe S, Featherstone R, Littlewood A, Metzendorf M-I, et al. Chapter 4: Searching for and selecting studies. 2023. In: Cochrane Handbook for Systematic Reviews of Interventions [Internet]. Cochrane. 6.4. Available from: https://training.cochrane.org/handbook.
15. EUnetHTA JA3WP6B2-2 Authoring Team. Process of information retrieval for systematic reviews and health technology assessments on clinical effectiveness: Methodological Guidelines. Diemen (The Netherlands): European network for Health Technology Assessment (EUnetHTA); 2019.
16. Område for helsetjenester i Folkehelseinstituttet. Slik oppsummerer vi forskning: Metodeboka [Nettdokument]. Oslo: Område for helsetjenester i Folkehelseinstituttet; 2022 [Available from: https://www.fhi.no/nettpub/metodeboka/.
17. Rethlefsen ML, Kirtley S, Waffenschmidt S, Ayala AP, Moher D, Page MJ, et al. PRISMA-S: an extension to the PRISMA Statement for Reporting Literature Searches in Systematic Reviews. Syst Rev. 2021;10(1):39.
18. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ. 2021;372:n71.
19. PRISMA Flow Diagram - PRISMA Statement [Available from: https://www.prisma-statement.org/prisma-2020-flow-diagram.
20. Sterne JA, Hernán MA, Reeves BC, Savović J, Berkman ND, Viswanathan M, et al. ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. BMJ. 2016;355:i4919.
21. Sterne JAC, Savović J, Page MJ, Elbers RG, Blencowe NS, Boutron I, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. BMJ. 2019;366:l4898.

22. Danish Health Technology Council. The Danish Health Technology Council's methods guide for the evaluation of health technology Aalborg Ø: Danish Health Technology Council; 2021.
23. Fang Y, He W, Wang H, Wu M. Key considerations in the design of real-world studies. Contemp Clin Trials. 2020;96:106091.
24. EUnetHTA. EUnetHTA 21 - D4.3.1: DIRECT AND INDIRECT COMPARISONS 2022 [Available from: https://www.eunethta.eu/wp-content/uploads/2022/12/EUnetHTA-21-D4.3.1-Direct-and-indirect-comparisons-v1.0.pdf.
25. EUnetHTA. EUnetHTA 21 D4.3.2 DIRECT AND INDIRECT COMPARISONS 2022 [Available from: https://www.eunethta.eu/wp-content/uploads/2022/08/EUnetHTA-21-Deliverable-D4.3.2-Methodological-Guideline-on-Direct-and-indirect-comparisons-V1.0.pdf.
26. Page MJ, Moher D, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. PRISMA 2020 explanation and elaboration: updated guidance and exemplars for reporting systematic reviews. BMJ. 2021;372:n160.
27. ICH Expert Working Group. Addendum on estimands and sensitivity analysis in clinical trials to the guideline on statistical principles for clinical trials e9(r1) 2019 [Available from: https://database.ich.org/sites/default/files/E9-R1_Step4_Guideline_2019_1203.pdf.
28. Dias S, Ades A, Welton N. Network Meta-Analysis for Decision Making. Chichester, UK: Wiley; 2018.
29. 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.
30. 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.
31. 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.
32. Bender R, Friede T, Koch A, Kuss O, Schlattmann P, Schwarzer G, et al. Methods for evidence synthesis in the case of very few studies. Res Synth Methods. 2018;9(3):382-92.
33. Dias S, Welton NJ, Caldwell DM, Ades AE. Checking consistency in mixed treatment comparison meta-analysis. Stat Med. 2010;29(7-8):932-44.
34. 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.
35. Berlin JA, Santanna J, Schmid CH, Szczech LA, Feldman HI. Individual patient- versus group-level data meta-regressions for the investigation of treatment effect modifiers: ecological bias rears its ugly head. Stat Med. 2002;21(3):371-87.
36. Cooper NJ, Sutton AJ, Morris D, Ades AE, Welton NJ. Addressing between-study heterogeneity and inconsistency in mixed treatment comparisons: Application to stroke prevention treatments in individuals with non-rheumatic atrial fibrillation. Statistics in Medicine. 2009;28(14):1861-81.
37. Dias S, Sutton AJ, Welton NJ, Ades AE. Evidence synthesis for decision making 3: heterogeneity--subgroups, meta-regression, bias, and bias-adjustment. Medical decision making : an international journal of the Society for Medical Decision Making. 2013;33(5):618-40.
38. Higgins JP, Thompson SG. Controlling the risk of spurious findings from meta-regression. Stat Med. 2004;23(11):1663-82.
39. Deeks JJ, Higgins JP, Altman DG. Analysing data and undertaking meta-analyses. In: Cumpston M, editor. Cochrane Handbook for Systematic Reviews of Interventions, 2nd Edition: Wiley; 2019. p. 241-84.
40. DerSimonian R, Laird N. Meta-analysis in clinical trials. Controlled Clinical Trials. 1986;7(3):177-88.
41. Jackson D, Law M, Rücker G, Schwarzer G. The Hartung-Knapp modification for random-effects meta-analysis: A useful refinement but are there any residual concerns? Stat Med. 2017;36(25):3923-34.
42. Knapp G, Hartung J. Improved tests for a random effects meta-regression with a single covariate. Stat Med. 2003;22(17):2693-710.

43. Kuss O. Statistical methods for meta-analyses including information from studies without any events-add nothing to nothing and succeed nevertheless. Stat Med. 2015;34(7):1097-116.
44. Mathes T, Kuss O. A comparison of methods for meta-analysis of a small number of studies with binary outcomes. Res Synth Methods. 2018;9(3):366-81.
45. Röver C, Bender R, Dias S, Schmid CH, Schmidli H, Sturtz S, et al. On weakly informative prior distributions for the heterogeneity parameter in Bayesian random-effects meta-analysis. Res Synth Methods. 2021;12(4):448-74.
46. Schulz A, Schürmann C, Skipka G. Performing meta-analyses with very few studies. In: Evangelor E, Veroniki AA, editors. Meta-Research: Methods and Protocols. New York: Humana; 2022. p. 91-102.
47. Stijnen T, Hamza TH, Özdemir P. Random effects meta-analysis of event outcome in the framework of the generalized linear mixed model with applications in sparse data. Statistics in Medicine. 2010;29(29):3046-67.
48. Sutton AJ, Abrams KR. Bayesian methods in meta-analysis and evidence synthesis. Stat Methods Med Res. 2001;10(4):277-303.
49. Turner RM, Domínguez-Islas CP, Jackson D, Rhodes KM, White IR. Incorporating external evidence on between-trial heterogeneity in network meta-analysis. Stat Med. 2019;38(8):1321-35.
50. Turner RM, Jackson D, Wei Y, Thompson SG, Higgins JP. Predictive distributions for between-study heterogeneity and simple methods for their application in Bayesian meta-analysis. Stat Med. 2015;34(6):984-98.
51. Veroniki AA, Jackson D, Bender R, Kuss O, Langan D, Higgins JPT, et al. Methods to calculate uncertainty in the estimated overall effect size from a random-effects meta-analysis. Res Synth Methods. 2019;10(1):23-43.
52. Wiksten A, Rücker G, Schwarzer G. Hartung-Knapp method is not always conservative compared with fixed-effect meta-analysis. Stat Med. 2016;35(15):2503-15.
53. NoMA. Submission guidelines For Single Technology Assessment of Medicinal Products. Oslo: Norwegian Medical Products Agency; 2024.
54. Dias S, Ades AE. Absolute or relative effects? Arm-based synthesis of trial data. Res Synth Methods. 2016;7(1):23-8.
55. Hong H, Chu H, Zhang J, Carlin BP. A Bayesian missing data framework for generalized multiple outcome mixed treatment comparisons. Res Synth Methods. 2016;7(1):6-22.
56. Lumley T. Network meta-analysis for indirect treatment comparisons. Stat Med. 2002;21(16):2313-24.
57. Mbuagbaw L, Rochwerg B, Jaeschke R, Heels-Andsell D, Alhazzani W, Thabane L, et al. Approaches to interpreting and choosing the best treatments in network meta-analyses. Syst Rev. 2017;6(1):79.
58. Salanti G, Ades AE, Ioannidis JP. Graphical methods and numerical summaries for presenting results from multiple-treatment meta-analysis: an overview and tutorial. J Clin Epidemiol. 2011;64(2):163-71.
59. Salanti G, Higgins JP, Ades AE, Ioannidis JP. Evaluation of networks of randomized trials. Stat Methods Med Res. 2008;17(3):279-301.
60. White IR, Turner RM, Karahalios A, Salanti G. A comparison of arm-based and contrast-based models for network meta-analysis. Statistics in Medicine. 2019;38(27):5197-213.
61. 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.
62. 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.
63. Phillippo DM, Dias S, Ades AE, Belger M, Brnabic A, Schacht A, et al. Multilevel network meta-regression for population-adjusted treatment comparisons. J R Stat Soc Ser A Stat Soc. 2020;183(3):1189-210.
64. Phillippo DM, Ades AE, Dias S, Palmer S, Abrams KR, Welton NJ. Methods for Population-Adjusted Indirect Comparisons in Health Technology Appraisal. Medical decision making : an international journal of the Society for Medical Decision Making. 2018;38(2):200-11.

65. Phillippo DM, Dias S, Ades AE, Welton NJ. Assessing the performance of population adjustment methods for anchored indirect comparisons: A simulation study. Stat Med. 2020;39(30):4885-911.
66. Remiro-Azócar A, Heath A, Baio G. Methods for population adjustment with limited access to individual patient data: A review and simulation study. Res Synth Methods. 2021;12(6):750-75.
67. Ali MS, Prieto-Alhambra D, Lopes LC, Ramos D, Bispo N, Ichihara MY, et al. Propensity Score Methods in Health Technology Assessment: Principles, Extended Applications, and Recent Advances. Front Pharmacol. 2019;10:973.
68. Austin PC. Balance diagnostics for comparing the distribution of baseline covariates between treatment groups in propensity-score matched samples. Stat Med. 2009;28(25):3083-107.
69. Desai RJ, Franklin JM. Alternative approaches for confounding adjustment in observational studies using weighting based on the propensity score: a primer for practitioners. Bmj. 2019;367:l5657.
70. Faria R, Hernández Alava M, Manca A, Wailoo A. NICE DSU Technical Support Document 17: The use of observational data to inform estimates of treatment effectiveness for Technology Appraisal: Methods for comparative individual patient data 2015 [Available from: http://www.nicedsu.org.uk/.
71. Hernán MA, Robins JM. Using Big Data to Emulate a Target Trial When a Randomized Trial Is Not Available. Am J Epidemiol. 2016;183(8):758-64.
72. Stuart EA, Lee BK, Leacy FP. Prognostic score-based balance measures can be a useful diagnostic for propensity score methods in comparative effectiveness research. J Clin Epidemiol. 2013;66(8 Suppl):S84-S90.e1.
73. Victor N. On clinically relevant differences and shifted null hypotheses. Methods Inf Med. 1987;26(3):109-16.
74. Wang SV, Pinheiro S, Hua W, Arlett P, Uyama Y, Berlin JA, et al. STaRT-RWE: structured template for planning and reporting on the implementation of real world evidence studies. Bmj. 2021;372:m4856.
75. National Institute for Health and Clinical Excellence. Diagnostic Assessment Programme Manual. 2011.
76. Teutsch SM, Bradley LA, Palomaki GE, Haddow JE, Piper M, Calonge N, et al. The Evaluation of Genomic Applications in Practice and Prevention (EGAPP) Initiative: methods of the EGAPP Working Group. Genet Med. 2009;11(1):3-14.
77. Whiting PF, Rutjes AW, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. Ann Intern Med. 2011;155(8):529-36.
78. EUnetHTA. EUnetHTA 21 - D4.6 VALIDITY OF CLINICAL STUDIES 2022 [Available from: https://www.eunethta.eu/wp-content/uploads/2022/12/EUnetHTA-21-D4.6-Practical-Guideline-on-validity-of-clinical-studies-v1.0-1.pdf.
79. NICE. NICE real-world evidence framework. National Institute for Health and Care Excellence; 2022.
80. Hendrix N, Veenstra DL, Cheng M, Anderson NC, Verguet S. Assessing the Economic Value of Clinical Artificial Intelligence: Challenges and Opportunities. Value in health : the journal of the International Society for Pharmacoeconomics and Outcomes Research. 2022;25(3):331-9.
81. Bélisle-Pipon JC, Couture V, Roy MC, Ganache I, Goetghebeur M, Cohen IG. What Makes Artificial Intelligence Exceptional in Health Technology Assessment? Front Artif Intell. 2021;4:736697.
82. Haverinen J, Keränen N, Falkenbach P, Maijala A, Kolehmainen T, Reponen J. Digi-HTA: Health technology assessment framework for digital healthcare services. Finnish Journal of eHealth and eWelfare. 2019;11(4):326–41.
83. EuroQol. EQ-5D [updated 16.05.2020. Available from: https://euroqol.org/.
84. Gold MR, Siegel JE, Russell LB, Weinstein MC. Cost effectiveness in health and medicine: Oxford University Press, Oxford; 1996.
85. 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.
86. **EuroQol**. EQ-5D-Y-5L User Guide, 2024 [Available from: https://euroqol.org/information-and-support/documentation/user-guides.

87. Kreimeier S, Åström M, Burström K, Egmar AC, Gusi N, Herdman M, et al. EQ-5D-Y-5L: developing a revised EQ-5D-Y with increased response categories. Qual Life Res. 2019;28(7):1951-61.
88. Wille N, Badia X, Bonsel G, Burstrom K, Cavrini G, Devlin N, et al. Development of the EQ-5D-Y: a child-friendly version of the EQ-5D. Qual Life Res. 2010;19(6):875-86.
89. Ramos-Goni JM, Oppe M, Stolk E, Shah K, Kreimeier S, Rivero-Arias O, et al. International Valuation Protocol for the EQ-5D-Y-3L. Pharmacoeconomics. 2020;38(7):653-63.
90. Hill H, Rowen D, Pennington B, Wong R, Wailoo A. A review of the methods used to estimate and model utility values in nice technology appraisals for pediatric populations 2019 [Available from: http://nicedsu.org.uk/wp-content/uploads/2019/10/DSU-Utility-values-in-HTAs-Final-report-2019.pdf.
91. 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.
92. Dawoud D, Lamb A, Moore A, Bregman C, Rupniewska E, Paling T, et al. Capturing what matters: updating NICE methods guidance on measuring and valuing health. Quality of Life Research. 2022;31(7):2167-73.
93. Garratt AM, Stavem K, Shaw JW, Rand K. EQ-5D-5L value set for Norway: a hybrid model using cTTO and DCE data. Qual Life Res. 2025;34(2):417-27.
94. Dolan P. Modeling valuations for EuroQol health states. Medical care. 1997;35(11):1095-108.
95. Stavem K, Augestad LA, Kristiansen IS, Rand K. General population norms for the EQ-5D-3 L in Norway: comparison of postal and web surveys. Health Qual Life Outcomes. 2018;16(1):204.
96. 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.
97. Direktoratet for økonomistyring. Veileder i samfunnsøkonomiske analyser. 2014.
98. Drummond MF, Sculpher MJ, Claxton K, Stoddart GL, Torrance GW. Methods for the economic evaluation of health care programmes: Oxford university press; 2015.
99. HAS. Choices in methods for economic evaluation – HAS. Saint-Denis: Haute Autorité de Santé 2020.
100. Finansdepartementet. Rundskriv R-109 - Prinsipper og krav ved utarbeidelse av samfunnsøkonomiske analyser mv. 2021.
101. Magnussen J, Aaserud M, Granaas T, Magelssen M, Syse A, Celius EG, et al. På ramme alvor - alvorlighet og prioritering. In: Helse- og omsorgsdepartementet, editor. 2015.
102. Incerti D. Parametric survival modeling 2019 [Available from: https://devinincerti.com/2019/06/18/parametric_survival.html#weibull-distribution-ph.
103. Statistisk sentralbyrå. Dødelighetstabeller [Available from: https://www.ssb.no/befolkning/statistikker/dode.
104. Beslutningsforum for nye metoder. Sak 048-2020: Veiledende kriterier for bruk av skjønn ved vurdering av absolutt prognosetap ved behandling av plager/sykdommer som følger av hovedtilstand, ved hurtig metodevurderinger av legemidler, pp 9-17 2020 [Available from: https://nyemetoder.no/Documents/Beslutninger/Saksdokumenter%20Beslutningsforum%20for%20nye%20metoder%2025.05.2020.pdf.
105. 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.
106. 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.
107. 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.
108. 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.
109. 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.
110. 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.
111. Dohoo I, Martin W, Stryhn H. Methods in Epidemiologic Research. Charlottetown, Prince Edward Island: AVC. Inc; 2012.

112. 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.
113. 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.
114. Collet D. Modelling survival Data in Medical Research. 3 ed: Chapman and Hall/CRC Press; 2014.
115. Ishak KJ, Kreif N, Benedict A, Muszbek N. Overview of Parametric Survival Analysis for Health-Economic Applications. PharmacoEconomics. 2013;31(8):663-75.
116. Rutherford M, Lambert P, Sweeting M, Pennington B, Crowther MJ, Abrams KR, et al. NICE DSU TECHNICAL SUPPORT DOCUMENT 21: Flexible Methods for Survival Analysis 2020 [Available from: http://www.nicedsu.org.uk.
117. Kearns B, Stevens J, Ren S, Brennan A. How Uncertain is the Survival Extrapolation? A Study of the Impact of Different Parametric Survival Models on Extrapolated Uncertainty About Hazard Functions, Lifetime Mean Survival and Cost Effectiveness. Pharmacoeconomics. 2020;38(2):193-204.
118. Palmer S, Borget I, Friede T, Husereau D, Karnon J, Kearns B, et al. A Guide to Selecting Flexible Survival Models to Inform Economic Evaluations of Cancer Immunotherapies. Value in health : the journal of the International Society for Pharmacoeconomics and Outcomes Research. 2023;26(2):185-92.
119. Jackson C, Stevens J, Ren S, Latimer N, Bojke L, Manca A, et al. Extrapolating Survival from Randomized Trials Using External Data: A Review of Methods. Medical decision making : an international journal of the Society for Medical Decision Making. 2017;37(4):377-90.
120. 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.

Norwegian Medical Products Agency

Appendix

1. Extrapolation of time-to-event data

Time-to-event data, also known as survival or event history, encompasses instances such as time to progression in cancer (progression-free survival or PFS), time to death (overall survival or OS), or time to a cardiovascular event or treatment discontinuation.

Health economic analyses commonly involve parameterisation and extrapolation of clinical time-to-event data beyond the study period. This section outlines the requirements for parameterisation and extrapolation of survival data for health economic analyses, applicable to both direct and indirect comparisons of relative efficacy. See Chapter 12.6 for guidance on selecting a time horizon.

1.1 Parameterisation of data from clinical studies

Extrapolating data beyond the study follow-up period usually involves using a parametric function. These functions assume that the underlying event risk (baseline risk) adheres to a specific distribution, as opposed to non-parametric (e.g., Kaplan-Meier) or semi-parametric (e.g., Cox model) functions.

The selection of a parametric function relies on statistical analyses of best mathematical fit and biological criteria, which pertain to the expected development of risk for the current condition/disease and endpoint.

Parameterisation must be based on the actual data from clinical studies to emphasise the direct effect of the treatment under evaluation.

Systematic statistical tests and graphic evaluations must be carried out to allow the choice of the most accurate parametric function (106-113).

To achieve satisfactory fitting, a given function must fulfil the following criteria:

1. The function must align well with the observed efficacy data from the study or studies.
2. The extrapolated segment must be clinically and biologically plausible.

Provide a detailed justification for the chosen function, considering the two criteria above. Functions that do not meet both criteria are likely unsuitable.

When fitting a parametric survival model, one may include explanatory variables, typically treatment, but also other variables. When choosing from many potential explanatory variables, the variable selection process should be described. Automated stepwise approaches are prone to overfitting and can lead to a poor fit in out-of-sample data and in extrapolated data. The theoretical arguments or expert opinion should therefore always be considered and reported.

1.2 Curve fitting to observed study data

For optimal evaluation of the curve fit, an extensive description and analysis of any assumptions and properties regarding the parametric functions and relevant clinical data should be submitted to document the adjustment(s) to the observed study data all the points in the list below must be included as a minimum:

- The following parametric functions must be included in the health economic model: exponential, Weibull, Gompertz, gamma, log-logistic, log-normal and generalised gamma distributions.
- Statistical tests and graphical presentation for testing of proportional hazard (PH), accelerated failure time model (AFT) and for assessing the fit of standard parametric functions (114, 115):
    - Log-cumulative hazard plot for PH: $\log(-\log(S(t)))$ vs. $\log(\mathrm{t})$ with linear trendlines for the intervention and comparator
    - Plot based on Schoenfeld residuals for assessment of PH

  - Quantile-Quantile-plot for AFT of the times of survival percentile $t_0(p)$ vs $t_1(p)$ with a linear trendline, using the percentiles (p) of the inverse survival functions for the intervention and comparator:
    - $t_0(p) = S_0^{-1}\left(\frac{100-p}{100}\right)$, $t_1(p) = S_1^{-1}\left(\frac{100-p}{100}\right)$
  - $\log(S(t)/(1-S(t)))$ vs. $\log(t)$ with linear trendlines for the intervention and comparator (graphical test for the suitability of log-logistic, and joint AFT modelling)
  - Inverse normal $(1-S(t))$ vs. $\log(t)$ with linear trendlines for the intervention and comparator (graphical test for the suitability of log-normal, and joint AFT modelling)

- If neither PH nor AFT appear suitable, standard parametric models fitted to each treatment arm independently should be considered before other, more flexible functions, such as a piecewise function, Royston-Palmer models, and spline models (116).
- Smoothed and unsmoothed hazard plots for the observed data from the clinical study per treatment arm (102, 117).
- Smoothed hazard plots for the observed data from the clinical study with the hazard function of all the standard parametric functions plotted in the same figure, per treatment arm (103, 117)
- Akaike's Information Criteria (AIC), Bayesian Information Criteria (BIC) and/or other suitable tests for those functions which are relevant based on the criteria described above, per treatment arm.
- Graphical presentation of time to event data curves, where both Kaplan-Meier (KM) data and the parametric distribution are shown in the same figure. Similar graphical presentation should also be included in the health economic model (in the spreadsheet).
- In some cases, curves with KM data for the first part of the study period can be appropriate, followed by a parametric tail which shows the extrapolation beyond this point (transition point). The transition point must be evaluated in each individual case. As a minimum requirement an analysis must be presented where the tail is set at the time point where 50 % of the included population in each treatment arm is still "at risk".
- Parameterisation of survival data should be conducted in a transparent way that allows the analysis to be reproduced.

1.3 Plausibility of the extrapolated part of the curve

The plausibility of the extrapolated part of the survival curve must be documented and justified both biologically and clinically for the patient group in question. External data (e.g., survival curves, hazard functions) can be employed to evaluate the assumptions made in the extrapolation. Systematic and reproducible identification of evidence is essential.

External data may include data for the same medical device(s) used in the same indication, more mature data from the same medical device(s), used in a later line of treatment for the same disease, data for the same medical device(s) used in a similar indication or evidence from a medical device with a similar mechanism of action used in the same indication (118).

Consider external sources as indicative only. Use of external data necessitates a balanced assessment of the degree to which discrepancies between long-term survival projections and external data sources can be attributed to:

- Shortcomings/weaknesses in the chosen parametric function, and/or
- Limitations in the external data source

Typically, external data is available solely for the comparator arm, rendering it most suitable for assessing the plausibility of projecting the comparator arm. Therefore, the 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. Various assumptions must be tested in the scenario analyses. It is recommended to include three alternative scenarios (119) for the long-term effect of a treatment that so far has been shown effective in studies with limited follow-up time:

- pessimistic scenario where the effect is null beyond study follow-up time,

- optimistic scenario where the effect beyond study follow-up time is assumed constant and the same as in short-term, and
- compromise scenario where the effect beyond study follow-up time is diminishing.

1.4 Algorithm and implementation in the health economic model

Figure 3 below depicts the algorithm for selection of a parametric model in time to event data analysis for health economic analyses.

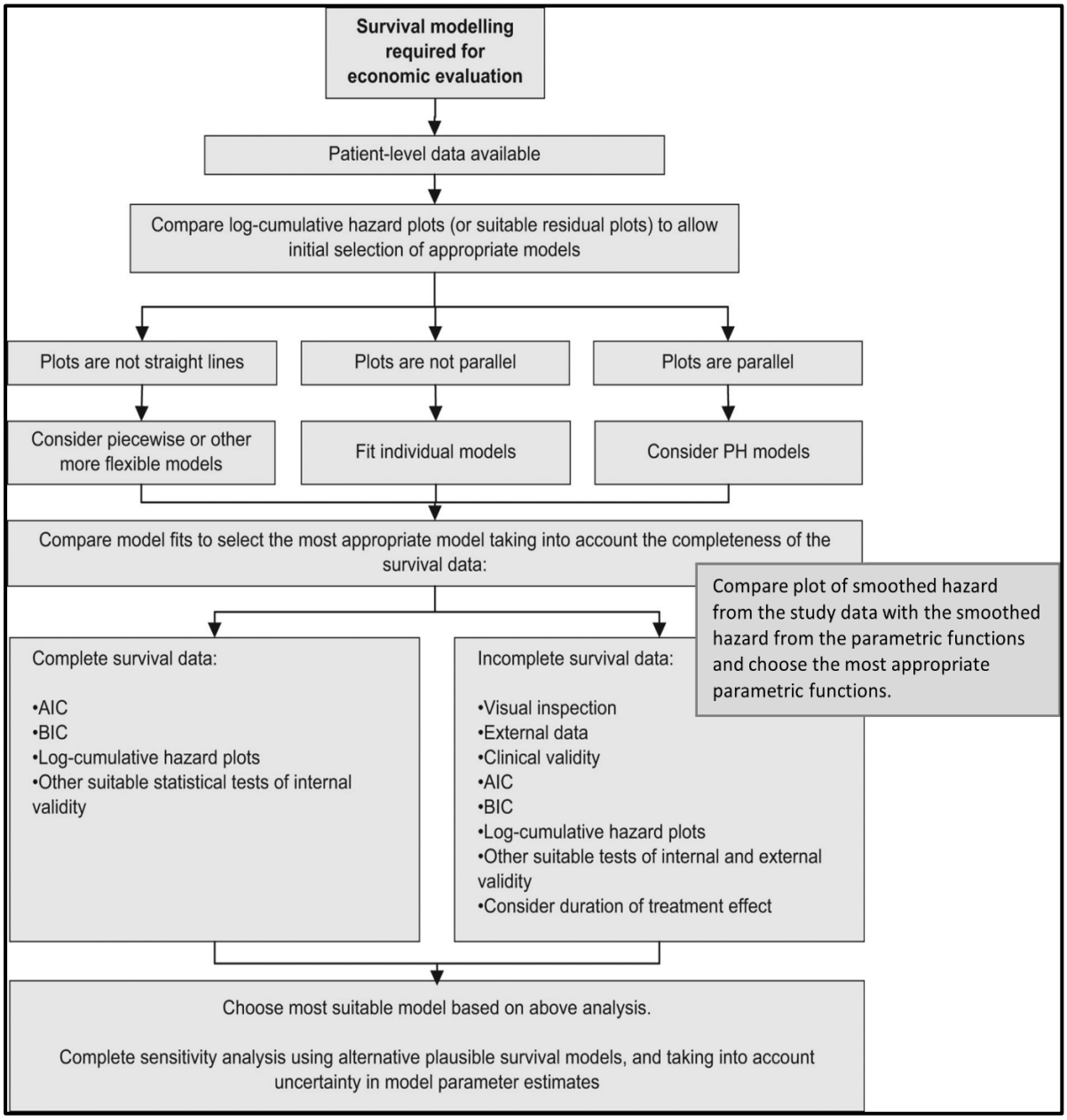

Figure 3. Algorithm for selection of a parametric model (modified from Latimer 2013 (120))

2. Age adjustment of future expected HSUV using relative utility decrements

Consider a chronic disease modelled with a HSUV of 0.820 for the patients' "best" expected health state. The mean age of the patients is 50 years, and the health economic model utilises a lifetime perspective. Without age adjustment, the HSUV remains constant for the proportion of patients reaching the "best" health status throughout their lifetime.

The utility values for the general population are generally not fixed across all ages. Without age adjustment, the health-related quality of life for a patient population over 81 years in the example above would be higher (0.820) than that assumed for the general population (0.811).

Age adjusted HSUV results from the initial HSUV of 0.820 multiplied by an adjustment index. This index is set at 1 at the start of the model and changes over time based on the age specific HSUV. An example of how this can be done is demonstrated in "*Tools for severity calculation and age adjustment*".

3. Quantification of severity

3.1 Quantification of severity/absolute shortfall in a model with lifetime perspective

Mean age at treatment initiation in the relevant patient group is 57 years, based on clinical experts' opinion supplemented by data from national registries.

Expected remaining healthy life years for a mean age of 57 years is estimated to 24 QALYs (see "*Tools for severity calculation and age adjustment*".

Based on the health economic model, the prognosis for the relevant patient group is an expected remaining lifetime of 2.5 undiscounted years or 1.5 undiscounted QALYs with the current standard treatment (comparator).

Absolute shortfall is:

24 QALYs – 1.5 QALYs = 22.5 undiscounted QALYs

Figure 4 illustrates graphically quantification of absolute shortfall for an intervention on group level. New treatment is considered at age A. In the absence of the disease, future health is represented by the area under the solid blue line ($QALYs_A$) from timepoint A. The disease shortens lifespan and reduces quality of life with the current standard treatment, as-shown by the shaded area (P_A). Absolute shortfall is the difference between $QALYs_A$ and P_A.

The figure does not take into account potential health loss related to the disease before treatment start, as the Priority-setting White Paper recommends considering only future health loss when quantifying severity.

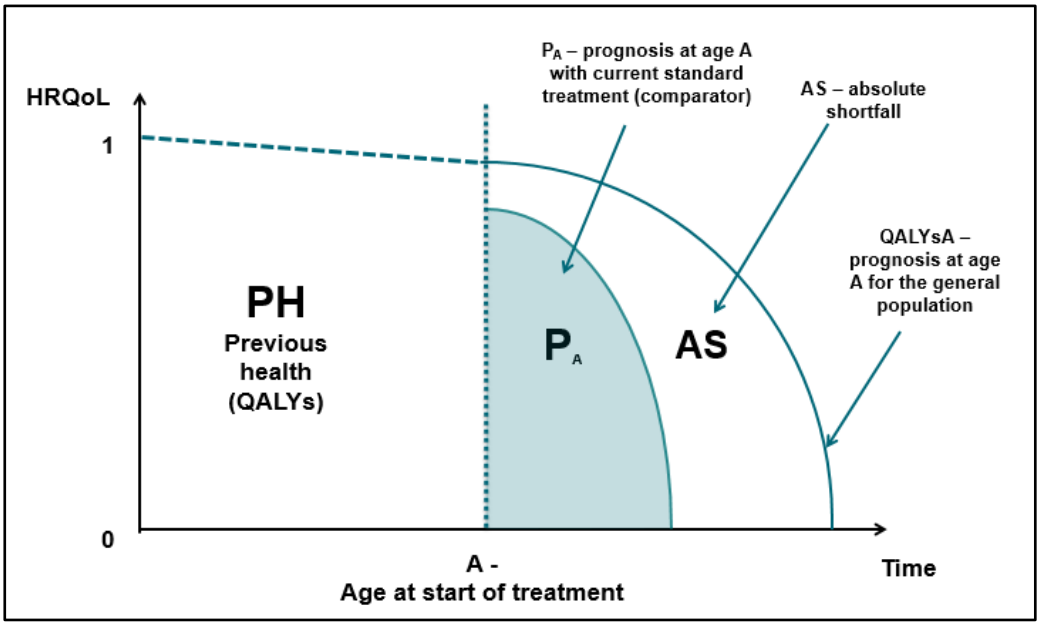

Figure 4. Graphical illustration of severity quantification, where health-related quality of life (HRQoL) on a scale from 0 (dead) to 1 (full health) is presented on the vertical axis, time is presented on the horizontal axis.

3.2 Quantification of severity/absolute shortfall in a model without lifetime time horizon

A chronic, non-fatal condition could be modelled with a perspective shorter than lifetime.

Mean age for treatment initiation in the relevant patient group is estimated to be 50 years.

For a 50-year-old, the expected remaining healthy life years ($QALYs_{50}$) is estimated to 29.6 QALYs.

The prognosis (undiscounted) from the health economic analysis is 0.750 undiscounted QALYs, applicable to a one-year time horizon. To calculate the lifetime prognosis, consider the disease and its progression given current standard treatment. Assuming relative QALY loss caused by the disease is constant over time, the prognosis is estimated as:

(0.750/0.870) * 27.0 QALYs = 23.3 undiscounted QALYs

where 27.0 QALYs is expected remaining QALYs, and 0.870 is HSUV for a 50-year-old (95).

Absolute shortfall is:

27.0 QALYs – 23.3 QALYs = 3.7 QALYs

3.3 Quantification of severity/absolute shortfall when preventing one disease

The new preventive measure is administered to the relevant population starting at a mean age of 40 years. With the current preventive practice (the comparator in the health economic analysis), the disease generally manifests at a mean age of 60 years. Thus, 60 years is used as input age in the calculation of absolute shortfall.

For a 60-year-old the expected number of remaining healthy life years ($QALYs_{60}$) is calculated as 21.6 QALYs.

For this disease in the relevant population, the prognosis with the current standard treatment is 7.3 QALYs. The prognosis takes into consideration that some individuals affected by the disease will die, but the majority will survive although with diminished quality of life to a certain extent (for instance myocardial infarction). The average prognosis is calculated as a weighted average of the prognosis for those who die and those who survive.

Absolute shortfall is:

21.6 QALYs – 7.3 QALYs = 14.3 QALYs

Calculation of absolute shortfall for preventive measures can also be shown as in Figure 4, the absolute shortfall is then calculated as:

- From timepoint A when the disease manifests (not the timepoint when the preventive measure is initiated).
- For a patient who develops the targeted condition, the subgroup who develops the disease at timepoint A (not for those who receives the preventive measure).
- With prognosis based on the current standard treatment of the condition.

3.4 Quantification of weighted severity/absolute shortfall when preventing two diseases

Quantification of weighted severity/absolute shortfall when preventing two diseases

Calculated absolute shortfall for disease A: 10 QALYs. Calculated absolutes shortfall for disease B: 6 QALYs.

In the health economic analysis, the benefit is estimated as 2.0 QALYs, representing the average incremental effect per person who receives the preventive measure. Of this, 1.8 QALYs (90%) are linked to prevention of disease A, and 0.2 QALYs (10%) to prevention of disease B.

The weighted absolute shortfall for disease A and B in this case is calculated as follows:

90% * 10 QALYs + 10 % * 6 QALYs = 9.6 QALYs

3.5 Quantification of expected remaining QALYs in the general population

Expected remaining QALYs and (health related) HSUV by age for the general Norwegian population are shown in "Tools for severity calculation and age adjustment", available at NoMA's website. Expected remaining QALYs are based on mortality data for the Norwegian population sourced from Statistics Norway (103) and age specific HSUV.