HAUTE AUTORITÉ DE SANTÉ

EVALUATING

HEALTH TECHNOLOGY

METHODOLOGICAL GUIDANCE

Choices in methods for economic evaluation – HAS

Validated by the CEESP on 6 April 2020

Contents

This document and its bibliographic reference are available to download at www.has-sante.fr

Haute Autorité de Santé – Communication and Information Department

5 avenue du Stade de France – 93218 Saint-Denis la Plaine Cedex – France. Tel.: +33 (0)1 55 93 70 00

Foreword

As part of its mandate, HAS (Haute Autorité de Santé – the French Health Authority) "defines and issues guidelines and medico-economic opinions on prevention, healthcare, prescription and best care strategies, and contributes to their comparison or ranking to support public health and optimise health insurance spending" (Article R161-71 of the French Social Security Code).

Using economic evaluation tools, HAS' analyses and opinions aim to guide public decision-making by estimating the incremental costs and benefits of the various healthcare products, services or programmes. The economic evaluation thus contributes to a better allocation of public healthcare spending, and greater transparency with respect to citizens.

Ten years after the adoption of the Social Security Financing Act that entrusted HAS with the task of issuing economic opinions, it had become necessary to update the methodological guidance used by HAS to perform this task.

The 2020 version of the methodological guidance for economic evaluation at HAS has been enhanced with the experience gained over the past ten years. It has also provided an opportunity for HAS to stress the importance of interpreting the evaluations, which are often perceived as highly technical. In this perspective, several guidelines call for more reasoned reflection on the objectives of the evaluation upon its conception, along with a constant effort to justify the methodological choices made, and an extensive interpretation of the results produced.

These are the conditions to ensure the usefulness of the economic evaluation for decision-making.

Dominique Le Guludec

Chairwoman of the HAS Board.

Summary of Guidelines

STRUCTURAL CHOICES IN EVALUATION METHODS

Guideline 1: Defining the objective

The context should be specified (clinical context, regulatory context, impacts expected or observed).

The objective of the economic evaluation should be clearly set out and justified with respect to the context.

Within the framework of HAS' mandate, the prime objective of an economic evaluation is to guide public decision-making regarding the allocation of collective resources, in particular by documenting the cost-effectiveness criteria.

Guideline 2: Choice of the evaluation method

The reference case analysis uses the cost-utility analysis and the cost-effectiveness analysis as economic evaluation methods. The choice of the method first depends on the nature of the health-related consequences – whether expected or observed – of the intervention being evaluated.

- If health-related quality of life is not a major consequence, a cost-effectiveness analysis is used for the reference case analysis and the health outcome evaluated is the length of life.
- If health-related quality of life is a major consequence, a cost-utility analysis is used for the reference case analysis and the health outcome is evaluated in terms of quality-adjusted life years. The cost-utility analysis should always be accompanied by a cost-effectiveness analysis based on a life expectancy evaluation.

Second, the choice of the method depends on data availability.

Guideline 3: Choice of the perspective

The reference case analysis should be based on a collective perspective, covering all of the people or institutions affected (in terms of health effects or cost) by the production of an intervention within the scope of the overall patient care.

Otherwise, the choice of a healthcare system perspective in the reference case analysis should be duly justified.

These two perspectives involve:

- identifying and measuring health effects from the point of view of the populations affected by the intervention being evaluated. The health outcomes should be evaluated from the point of view of the population as a whole;
- identifying, measuring and valuing the production costs of the evaluated intervention and its comparators, independently from their sources of funding.

Guideline 4: Choice of the population analysed

The population analysed should consist of all individuals whose health is affected –whether directly or indirectly – by the intervention being evaluated. Any inability to include certain affected individuals in the reference case analysis should be duly justified.

The economic approach may call for the analysis of specific population subgroups when differences in care imply different comparators or when variability is expected in terms of health effects or costs.

Guideline 5: Choice of interventions to be compared

The reference case analysis should identify all clinically relevant interventions in the population analysed.

The arguments supporting the inclusion or exclusion of an intervention in the reference case analysis should be clearly set out.

The interpretation of the results of the evaluation should consider the degree of exhaustiveness of the comparators used. Given that all comparators need to be taken into account to evaluate cost-effectiveness, the risk of excluding an intervention likely to be on the cost-effectiveness frontier, and its consequences, should be discussed.

Guideline 6: Choice of the time horizon

The time horizon of the evaluation may be defined as the entire lifetime or a specific duration.

The choice of the time horizon should be based on an trade-off between the information produced by taking into account the expected or observed consequences of an intervention over the long term, and the uncertainty generated by extrapolation over time. This trade-off should be clearly set out and the choice should be justified.

The time horizon used should be the same for all interventions.

Guideline 7: Discounting method

Future costs and health outcomes should be discounted to present values whenever the time horizon exceeds 12 months.

The reference case analysis should use the public discount rate applicable at the time of the evaluation. On the date of publication of this guidance, that rate was 2.5% for time horizons of less than 30 years. Beyond that, the rate gradually decreases to a floor level set at 1.5%.

In the reference case analysis, costs and results should be discounted using the same rate.

CHOICE OF METHODS FOR THE EVALUATION OF OUTCOMES

Guideline 8: Evaluation of health outcomes

The consequences to be considered in the economic evaluation are the effects of the interventions on the health of the population analysed.

Guideline 9: Choice of the criteria in cost-effectiveness analyses

If the reference case analysis is a cost-effectiveness analysis, the health outcome criterion to be used should be that of life-years. The mortality indicator should be the all-cause mortality rate.

If the data required for the measurement of life-years are unavailable, the use of a predictive criterion of expected survival time may be acceptable, but only if there is strong, established evidence of the predictive nature of this surrogate endpoint.

As a supplemental analysis, a cost-effectiveness analysis may be conducted on the basis of other health outcome criteria, with arguments supporting the choice of the criterion used.

Guideline 10: Choice of the criteria in cost-utility analyses

If the reference case analysis is a cost-utility analysis, the health outcome criterion to be used should be that of quality-adjusted life years (QALYs).

Health-related quality of life is valued using health state preferences, measured through utility scores.

Guideline 11: Evaluation of effectiveness and tolerability

To conduct an economic evaluation, evidence-based comparative clinical data is essential, in particular when the cost-effectiveness of an intervention needs to be established.

Guideline 12: Clinical data sources

All clinical data sources for the evaluated intervention and its comparators should be identified through a systematic, reproducible methodology, and then presented in accordance with international standards and HAS guidelines.

Available data sources to document the effectiveness and tolerability of each intervention should undergo a stringent critical appraisal. The sources with the highest level of evidence should be selected.

For each selected source, the transposability of experimental health outcomes to standard French practices should be discussed.

Guideline 13: Comparative effectiveness evaluation methods

Direct comparisons in randomised controlled trials and meta-analyses are the best source of evidence of comparative effectiveness.

If the interventions studied have not been directly compared in the same study, an indirect comparison method should be used to estimate the differences in effectiveness between the interventions studied and a "reference" intervention.

- The choice of the reference intervention should be justified.
- The method used to estimate the relative effect with respect to the reference intervention should be the same for all interventions.

The choice of the comparative effectiveness evaluation method should be justified, and the method should be presented in a clear and detailed way.

- A network meta-analysis is recommended, subject to the validation of the underlying feasibility assumptions (consistency of populations, protocols and related bias risks, and verification of the transitivity assumption).
- The use of an indirect comparison method adjusted on the basis of individual data may be put forward if justified.

The clinical assumptions and methodology choices made to estimate comparative effectiveness should be described, justified and tested in a sensitivity analysis.

The level of evidence of the estimated comparative effectiveness should be clearly set out.

Guideline 14: Tolerability evaluation methods

For the collection and processing of tolerability data, the same methodological rigour should be applied as for the collection and processing of effectiveness data.

The frequency and severity of adverse effects should be reported in detail for all interventions, along with their duration and time to onset.

Guideline 15: Utility score evaluation methods

The utility scores used to weight life-years should be estimated using a multi-attribute approach, comprising the collection of health state data from patients via a generic questionnaire and the valuation of health states according to the preferences of the general population.

Among available classification systems, EQ-5D-5L is recommended (French EQ-5D-5L value set and EQ-5D-5L questionnaire). The French value set which prevails at the time of the evaluation should be used.

Failing that, as a transitional measure, the EQ-5D-3L classification system (French EQ-5D-3L value set and EQ-5D-3L questionnaire) should be used.

In the absence of a utility score from an EQ-5D classification system, a mapping approach should be opted for, in order to arrive at an EQ-5D utility score, provided that a function based on methodological quality standards is available and has been validated.

Other approaches should not be used for base-case analyses (see the specific questionnaire with the valuation of preferences in the general population, the standard gamble and time trade-off methods that directly assess, with the patients, the utility associated with their health state, and the ordinal approach). These can be subjected to a sensitivity analysis.

The estimation of utility scores through an approach revealing preferences for a hypothetical health state via vignettes or through a visual analogue scale is not acceptable, even within the framework of a sensitivity analysis.

Guideline 16: Data sources for the estimation of utility scores

Utility scores are either derived from an *ad-hoc* study specifically designed for the collection of the required quality-of-life data or drawn from a systematic literature review. They should not be based on the opinions of experts.

For the collection and processing of quality-of-life data for the estimation of a utility score, the same methodological rigour should be applied as for the collection and processing of effectiveness and tolerability data.

Guideline 17: Methods for estimating utility scores in specific populations

From the age of 16 years, the EQ-5D classification system should be used. Before the age of 16 years, a paediatric classification system should be used. In the absence of a value set for the valuation of French preferences, foreign value sets may be used.

In cases of impaired cognition, the use of the EQ-5D classification system via a version of the questionnaire filled in by a patient proxy is recommended.

METHODOLOGICAL CHOICES FOR THE EVALUATION OF COSTS

Guideline 18: Evaluation of costs

The evaluation of the total cost of an intervention is based on the intervention's production costs. This requires the identification, measurement and valuation of the resources consumed.

The scope of the cost items evaluated depends on the perspective adopted.

Under a collective perspective, all resources consumed in the production of the overall patient care are taken into consideration. They cover the domestic sphere (e.g. informal care), the healthcare sphere (e.g. stays, procedures, and health products) and the medico-social sphere (e.g. stays, personal care services).

Under the healthcare system perspective, the resources considered are those involved in the production of care (stays, procedures, and healthcare products).

Only direct costs should be considered in the reference case analysis.

A direct cost analysis may be presented as a supplemental analysis.

Guideline 19: Direct cost evaluation method

The evaluation of direct costs involves three stages: the identification, measurement and

evaluation of the resources associated with the intervention.

All resources associated with the intervention should be identified over the relevant time horizon. The impossibility of taking a resource into account should be duly justified.

The quantities of resources consumed should be measured through an appropriate method and clearly referenced and validated sources.

The valuation of the resources should be based on their unit production cost in France.

When valuation using the production cost is not possible, French tariffs may be used.

Guideline 20: Indirect cost evaluation method

The impact of an intervention on the time dedicated to a professional or leisure activity should be measured in terms of specific lost time per category of activities affected.

The method used to appraise this lost time is to be chosen by the evaluator and justified.

METHODOLOGICAL CHOICES CONCERNING MODELLING

Guideline 21: Modelling principles

The development of a model should comply with three requirements: justification, validation and exploration of uncertainty.

If there is no solid argument to clearly justify a particular methodological choice over several credible options, the one which is the least favourable to the intervention evaluated, in terms of cost or health outcome differential, should be opted for.

Guideline 22: Choice of model type and structure

The type of model and its structure should be defined so as to represent the patients' clinical progression and care, without introducing any needless complexity.

The technical characteristics of the model should be suited to the specificities of the evaluation (mode of progression over time, degree of heterogeneity and interaction among individuals, and degree of randomness) and comply with applicable recommendations.

The structure of the model (statuses, events, links) should be defined so as to capture the costs and health outcomes associated with the progression of the disease and comparative care.

complementary approaches: a probabilistic sensitivity analysis based on a second-order Monte Carlo simulation, and deterministic

Guideline 23: Estimation of the value of model parameters

Observed parameter values and distributions should be used to document the model.

When there is no observed value for a parameter, it should be estimated through an *ad hoc* calculation or calibration method. The estimation method used should be detailed and justified.

The uncertainty associated with the estimation of the value of the model parameters should be explored and quantified.

When an extrapolation technique is used, all assumptions should be detailed and justified, in particular those relating to the treatment duration and effect size.

– Sensitivity analyses should quantify the impact of methodological choices and modelling assumptions (e.g. model structure, data sources, calculation methods or assumptions to estimate the value of parameters not directly observed). The impact of the assumptions used for the extrapolation of treatment effects should be systematically explored.

The uncertainty associated with the estimation of the model parameters should be systematically explored using two

– the model and the simulated results are intuitively consistent (face validity);

sensitivity analyses identifying the parameters (or combinations of parameters) which have the greatest influence on the results of the evaluation.

For all of the sensitivity analyses presented, the credibility of the options tested should be justified, along with an interpretation of their results and their contribution to the comprehension of the evaluation.

When a scenario which is fundamentally different to that used in the reference case analysis is put forward, the presentation of the results should include a thorough exploration of the uncertainty through deterministic and probabilistic sensitivity analyses.

Guideline 24: Validation of the model

The model's ability to produce consistent and credible simulations should systematically be explored through the technical verification of the model and its internal validation, and through a procedure for the validation of simulated results aimed at ensuring that:

- the simulated results are consistent with external data not used for the configuration of the model (external or predictive validity);
- the model generates results which are comparable to those of other models whose validity has been recognised (cross-validity).

All significant differences and all inconsistencies should be examined, and their origin should be sought

Guideline 25: Exploration of uncertainty in the model

A systematic exploration of the sources of uncertainty associated with the evaluation's structural choices, the modelling choices and the model parameter estimations should be presented according to an appropriate methodology.

Sensitivity analyses should quantify the impact of a different structural choice in the reference case analysis (e.g. perspective, time horizon, population analysed, comparators, discount rate).

PRESENTATION AND INTERPRETATION OF CONCLUSIONS

Guideline 26: Interpretation of the economic evaluation

Quantitative results should be presented and interpreted in keeping with the objective of the economic evaluation.

The evaluation of cost-effectiveness requires the identification of the interventions on the cost-effectiveness frontier and the presentation of the results based on the incremental cost-effectiveness ratio (ICER) or net benefit (NB)

The degree of confidence associated with the results should be detailed.

Guideline 27: Presentation of the evaluation

The economic evaluation should be presented is a well-structured, clear and detailed way. The methodology should be transparent. The data and sources used should be clearly reported.

All of the relevant economic information to aid public decision-making should be extracted from the evaluation.

A clear, justified discussion should make it possible to estimate the robustness of the conclusion of the evaluation and define the conditions under which the conclusion would be altered. This discussion should rest on a critical appraisal of the methods and data used, and on the sensitivity analyses conducted.

For each of the interventions, non-discounted values should be presented for each major health outcome and cost component. The total costs and health outcomes obtained on the main criteria should then be calculated and discounted.

HAS frame of reference

The reference case analysis* – composed of a base-case analysis* and a comprehensive exploration of uncertainty – rests on the methodological choices defining HAS' frame of reference.

	The chosen methodological options are the responsibility of the author, who must justify their choice.
Objective	Within the framework of HAS' missions, the objective of an economic evaluation is to guide public decision-making regarding the allocation of collective resources, in particular by documenting the cost-effectiveness criteria.
Evaluation method	Cost-effectiveness or cost-utility analysis depending on the nature of the interventions' health effects and data availability. – CEA if health-related quality of life is not an important consequence. – CUA if health-related quality of life is an important consequence.
Perspective	Collective perspective or, failing that, healthcare system perspective. – Population whose health is affected (identification and measure of health effects) and general population (valuation of health preferences) – All of the resources involved in the production of care, irrespective of their source of funding.
Population analysed	All of the individuals concerned, either directly or indirectly, by the intervention evaluated.
Comparator interventions	All options in the population analysed should be identified. The selection of the comparator interventions should be duly justified.
Time horizon	The choice of a time horizon spanning the entire lifetime, or a specific period should be based on an trade-off between the information produced and the uncertainty generated by extrapolation over time.
Discounting	Beyond 12 months, discounting is based on the public discount rate applicable at the time of the evaluation (set at 2.5% at the time of publication of this guidance). After 30 years, the discount rate gradually decreases to 1.5%.
Health outcome criteria	CEA: lifetime (indicator: life years / all-cause mortality). CUA: quality-adjusted life years (indicator: QALY)
Cost criteria	Direct costs based on production costs or, failing that, on their tariff/price.
Conclusion of the evaluation	In terms of cost-effectiveness, the results presented identify the interventions on the cost-effectiveness frontier and provide an estimate of the ICER or NB on the cost-effectiveness frontier. Exploration of uncertainty through deterministic and probabilistic approaches Analysis of expenditure transfers between funders
Critical appraisal of the evaluation	Analysis of the validity of the method and uncertainty of the results Discussion on the conclusions and limitations of the evaluation

Introduction

In the field of health, economic evaluation provides indispensable guidance to public decision-makers to find a better balance in terms of improving people's state of health, equitable access to healthcare, the quality of healthcare services, the capacity to integrate innovation, and control of public spending. The scope of the interventions and actions involved is broad: health products and services, diagnostic/curative/preventive programmes or strategies, the issuing of guidelines, negotiation of the price of healthcare goods, etc.

As early as 2008, HAS started to include economic evaluation studies in its work programme. The Social Security Financing Act for 2012 broadened HAS' economic evaluation mandate – and that of the CEESP within it – thus confirming the place of cost-effectiveness as an aid to public decision-making. In its first prospective report (2018), HAS stressed that the change in the conditions underpinning the emergence of innovations reinforced the need to further increase the practice and quality of economic evaluation (Haute Autorité de Santé 2018).

HAS gives paramount importance to the evaluation of the cost-effectiveness criteria.

The two main economic approaches implemented and promoted by HAS are economic evaluation and budget impact analysis (BIA).

These two complementary analyses have different objectives.

Economic evaluation examines the estimated cost of the production of an intervention in relation to expected or observed health effects, in comparison with other relevant medical options, while exploring the uncertainty associated with these estimates.

On the other hand, the budget impact analysis estimates the annual financial impact of the adoption of an intervention. The issue of expense sustainability for the national health insurance system only makes sense when the economic interest of the intervention for the community has been demonstrated.

The writing of the guidelines for this revised version of the methodological guidance confirms HAS' will to base its economic evaluation on the cost-effectiveness criteria. This reflects the importance given to the procedures for building an cost-effectiveness frontier and estimating an incremental cost-effectiveness ratio (ICER) or incremental net benefit (INB).

The methodological quality of any evaluation is a prerequisite to aid decision-making.

An evaluation's high-quality methodology is not the only requirement. Its quantitative results must be discussed and interpreted – irrespective of whether or not reference values exist – in order to precisely define the nature of the information produced and determine its limits.

The in-depth exploration of uncertainty – which applies to any estimation due to its underlying assumptions and source data – is indispensable for the good methodological quality of an economic evaluation. There are various sources of uncertainty, stemming from the assumptions used in the model, the incompleteness of the data, etc. Technically speaking, sensitivity analyses make it possible to quantify the uncertainty and better understand its origins and consequences. They form an integral part of the results of the evaluation. Their contribution should be interpreted by the evaluator through a discussion of the exploratory or demonstrative scope of the results and conclusions of the evaluation.

HAS promotes the full use of all of the information provided by the economic evaluation.

While the main objective of the evaluation promoted by HAS is the estimation of cost-effectiveness, the economic evaluation also provides other important economic information that can be used by public decision-makers. For example, it makes it possible to estimate the cost of producing an additional health gain unit by replacing an intervention with another, which is not necessarily the comparator used for the evaluation of cost-effectiveness.

Thus, when the economic evaluation is faced with difficulties that make it impossible to build an cost-effectiveness frontier (e.g. non-exhaustiveness of the options compared), this invalidates the demonstration of cost-effectiveness but does not invalidate all of the other information provided by the evaluation.

It is fundamental to extract all of the relevant economic information from the evaluation to guide public decision-making.

The implementation of HAS' methodological guidelines is the responsibility of the evaluator.

The objective of this methodological guidance is to specify the methodological principles adopted by HAS for the production and analysis of an economic evaluation to estimate the cost-effectiveness of an intervention, defined as a health product, service or programme.

The principles set out in this guidance document do not solely apply to the regulatory field of advice concerning health products. They consist of general principles – many of which are in force in other countries – which are applicable irrespective of the economic evaluation context. Their implementation is part of a practice that combines justified scientific rigour and pragmatism. Accordingly, any evaluation is the responsibility of its author. The justification of their methodological choices is thus fundamental.

In line with this position, HAS did not want to produce a technical manual or lay down its guidelines in methodological decision trees.

The aim of HAS' methodological guidance is to improve the methodology and interpretation of economic evaluations, by prompting their authors to define a clear objective, justify the structural and technical choices made and, most importantly, provide an in-depth interpretation of the results produced in order to guide public decision-making.

Main modifications in relation to the previous version

Improving the interpretation of results and uncertainty for better use of the results of the economic evaluations in public decision-making.

- Guideline added on the defining of the objective of the evaluation
- Explanation of the specificities linked to the cost-effectiveness criteria, in particular concerning the selection of comparators
- Guideline in favour of conservative choices in the absence of solid arguments
- Prompting the interpretation of the results obtained during the exploration of uncertainty
- Prompting the interpretation of the results in keeping with the objective of the evaluation

Providing details concerning technical aspects to improve the quality of the evaluations.

- Introduction of a new section dedicated to the evaluation of comparative effectiveness and tolerance (7 new guidelines)
- Update of technical data sheets, in particular those dedicated to costing
- Details on the reasons behind the choice of the evaluation method (nature of expected consequences and data availability)
- Details concerning perspectives (collective and health system)
- Details concerning the trade-off underlying the choice of the time horizon (information produced over the long term versus uncertainty of extrapolation over time)
- Details concerning the use of predictive criteria of expected survival time
- Details concerning the evaluation of utility scores in paediatrics, impaired cognitive states and rare events
- Emphasis on the justification, validation and exploration of uncertainty in the modelling process
- Details concerning the data extrapolation stage and the importance of clearly explaining and justifying underlying assumptions.

Updating the methods used to remain in step with current scientific knowledge.

- Update of the applicable discount rate (Commissariat général à la stratégie et à la prospective, 2013)
- Update of the recommended method to estimate utility scores (Andrade, Ludwig, & Goni, 2020)
- Update of cost-effectiveness estimation metrics with the adoption of the net benefit (NB) criteria as an alternative to the incremental cost-effectiveness ratio (ICER)

1. Structural choices in evaluation methods

1.1. Objective of the evaluation

Guideline 1

The context should be specified (clinical context, regulatory context, impacts expected or observed).

The objective of the economic evaluation should be clearly set out and justified with respect to the context.

Within the framework of HAS' mandate, the prime objective of an economic evaluation is to guide public decision-making regarding the allocation of collective resources, in particular by documenting the cost-effectiveness criteria.

The generic objective of an economic evaluation is to examine the difference in the cost and health outcome of an intervention in relation to one or more competing interventions according to the particular evaluation issue.

Within the framework of HAS' mandate, the prime objective of an economic evaluation is to guide public decision-making regarding the allocation of collective resources, in particular by documenting the cost-effectiveness criteria.

To meet this objective, it is necessary to: a) identify the options that offer the best combination of production factors and constitute the cost-effectiveness frontier (technical cost-effectiveness*), and b) identify the options that allow optimal allocation of available resources (allocative cost-effectiveness*).

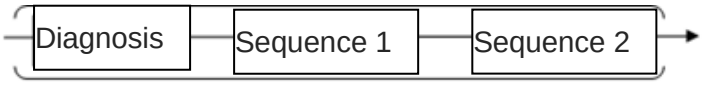
The type of method to be used depends on the evaluation context. This context should be specified and the questions it raises on the economic level – particularly in terms of cost-effectiveness – should be detailed in order to define the objective of the evaluation.

The choice of the evaluation method particularly depends on expected or observed impacts, the clinical context of the intervention being evaluated, and the regulatory context of the evaluation. These three aspects should be given special attention in the presentation of the objective.

Clinical context: healthcare strategy concept

The intervention evaluated is part of a healthcare strategy, i.e. a coordinated set of intervention sequences. The objective should specify whether the evaluation covers all sequences or only one (or some) of them.

Figure 1. Example of a healthcare strategy



The healthcare strategy and the place of the evaluated intervention in this strategy should be detailed and help to define the objective of the evaluation. The following two points should be taken into consideration.

- The evaluated intervention supplements or replaces other existing interventions.
- The evaluated intervention modifies the healthcare strategy (e.g. impact on several lines of treatment) or its impact is restricted to the sequence in which it fits.

Regulatory context: specificity of an evaluation produced within the framework of Article R. 161-71-3 of the French Social Security Code

In this specific context, the main objective of the evaluation is to assess the expected or observed cost-effectiveness of the health product or technology for the indication concerned, based on the knowledge available at the time of the request, in order to substantiate a price claim.

In cases of requests for the registration of a new product (initial registration) or for the extension of an existing product's indication to a new therapeutic area, only the main objective stated earlier is to be considered for the evaluation.

In cases of requests for re-registration or extension of the indication within the same therapeutic area[1], a second objective should be defined for the evaluation in order to substantiate the request to maintain or change the existing price.

The objective is then to justify, based on the cost-effectiveness criteria, the request to maintain or adjust the price, in view of the development of knowledge since the initial registration or previous evaluation. High-quality data should be provided to document this change (trial monitoring data and real-life data), in particular the data requested by HAS for the previous evaluation. This justification rests on the production of a quantitative evaluation or, failing that, a justified discussion.

- Re-registrations: these should systematically involve the presentation and analysis of the elements confirming the cost-effectiveness expected upon first registration at the negotiated face price.
- Concerning indication extensions within the same therapeutic area (see line change or new target population*): these should systematically involve the presentation and analysis of the elements making it possible to document the impact of the indication extension in terms of cost-effectiveness in the therapeutic area (e.g. first-line cost-effectiveness of the product in relation to the prior second-line indication).

Expected or observed impacts

When the impacts on health or costs are not the only effects claimed for an intervention, it is necessary to identify, describe, and possibly evaluate the other impacts[2].

These impacts should be assessed in a supplemental analysis* to the reference case analysis*, using an appropriate methodology.

1.2. Choice of the evaluation method

Guideline 2

The reference case analysis uses the cost-utility analysis and the cost-effectiveness analysis as economic evaluation methods. The choice of the method first depends on the nature of the health-related consequences – whether expected or observed – of the intervention being evaluated.

- If health-related quality of life is not a major consequence, a cost-effectiveness analysis is used for the reference case analysis and the health outcome evaluated is the length of life.
- If health-related quality of life is a major consequence, a cost-utility analysis is used for the reference case analysis and the health outcome is evaluated in terms of quality-adjusted life years.

[1] For example, this could be a second-line product claiming a first-line indication, or a product seeking an extension of indication to a wider population for the same pathology.

[2] At the time of publication of this document, methodological guidance is being prepared for the factoring-in of the organisational impact of health technology in its evaluation.

The cost-utility analysis should always be accompanied by a cost-effectiveness analysis based on a life expectancy evaluation.

Second, the choice of the method depends on data availability.

The two methods adopted by HAS are the cost-effectiveness analysis and the cost-utility analysis.

The choice of a cost-outcome analysis* (COA) first depends on the nature of the consequences of the evaluated intervention on health.

- If health-related quality of life* is not a major consequence of the intervention[3], the reference case analysis* is a cost-effectiveness analysis* (CEA) and the health outcome evaluated is the entire lifetime (see page 25, Health outcome criteria in cost-effectiveness analyses).
- If health-related quality of life* is a major consequence of the intervention, the reference case analysis* is a cost-utility analysis (CUA) and the health outcome is evaluated in terms of quality-adjusted life years (see page 26, Health outcome criteria in cost-utility analyses). The cost-utility analysis* should always be accompanied by a cost-effectiveness analysis*.

Second, the evaluator should determine whether the required data for the choice and implementation of the method recommended under the methodological requirements (see page 25, Methodological choices for the evaluation of health outcomes) are available and, if not, if such data can be produced at a reasonable cost and within a reasonable time frame.

The cost-benefit analysis is not recommended for the reference case analysis.

The cost-benefit analysis* (CBA) is the most suitable approach to evaluate the allocation of collective resources, since it makes it possible to assess the social added value of a public expenditure. Nevertheless, the methods used to carry out this type of analysis – particularly in the health sector – are subject to debate. Given the current state of this debate, HAS does not wish to promote such an analysis in the performance of its tasks, especially since the scarceness of cost-benefit evaluations in health – compared with CUAs and CEAs – limits the comparability of studies. Nevertheless, a CBA conducted in accordance with applicable methodological standards may be presented as a supplemental analysis* (Drummond & al., 2015) (Johannesson & al., 1996) (Nocera, Telser, & Bonato, 2003).

Analyses other than cost-effectiveness analyses can be used to aid decision-making in the field of health.

Other economic approaches may be used to aid decision-making, in particular when the evaluation issue is broader than the measurement of cost-effectiveness.

Cost-consequence analysis

The cost-consequence analysis* (CCA) is a type of cost-outcome analysis which makes it possible to compare interventions across several dimensions without an aggregate measure (Drummond, Sculpher, & Torrance, 2005) (Brazier, Ratcliffe, Salomon, & Tsuchiya, 2017). The costs and other evaluation criteria of the interventions compared are presented separately in a detailed manner.

This type of analysis is appropriate when the goal is to provide a precise and detailed account of the diverse impacts of a complex intervention, for instance for a public health evaluation. The range of

[3] The evaluator is required to set out the reasons behind his/her judgement that health-related quality should (or should not) be taken into account in the main outcome criterion, based on a claimed impact on quality of life.

expected impacts that can be taken into account is then as wide as necessary: health effects (events avoided, improvement of biological or radiological criteria, response rate, etc.), satisfaction of patients or professionals, caregiver well-being, reduction of inequalities among the population, or any other relevant consideration. Each impact is assessed according to a specific methodology, whose choice is to be justified.

Since the CCA cannot be used to assess the opportunity cost* of an intervention, it is not suitable when the objective of the economic evaluation is to estimate cost-effectiveness.

Multiple Criteria Decision Analysis

The multiple criteria decision analysis (MCDA) is a decision-making aid suited to complex decision-making contexts, characterised by multiple potentially competing objectives. The multiple criteria decision analysis refers to a set of techniques aimed at documenting deliberations in a well-structured, explicit way on the basis of multiple, precisely defined criteria (Thokala, Devlin et Marsh 2016) (Josselin et Le Maux 2017) (Zawodnik et Niewada 2018), for instance for the overall evaluation of a health programme or health technology (see full health technology assessment).

The evaluation issue covered by this type of analysis is thus broader than that of the economic evaluation covered in this methodological guidance.

Multiple criteria decision analyses conducted within the framework of a public health evaluation or health technology evaluation follow internationally applicable methodological guidelines (Marsh, Ijzerman et Thokala 2016) (Muhlbacher et Kaczynski 2016) (Angelis et Kanavos 2016).

1.3. Choice of perspective

Guideline 3

The reference case analysis uses a collective perspective, covering all individuals or institutions affected by the production of an intervention – whether in terms of health effects or cost – within the scope of the overall care provided. Otherwise, the choice of a healthcare system perspective in the reference case analysis should be duly justified.

These two perspectives involve:

- identifying and measuring health effects from the point of view of the populations affected by the intervention being evaluated. Health outcomes are assessed from the point of view of the general population;
- identifying, measuring and valuing the production costs of the evaluated intervention and its comparators, independently from their sources of funding.

The perspective of the economic evaluation determines the persons or institutions for whom the effects on heath and costs will be considered, in keeping with the objective of the evaluation.

Within the framework of HAS' missions, the objective of an economic evaluation is to guide public decision-making regarding the allocation of collective resources, in particular by documenting the cost-effectiveness criteria.

The perspective adopted to document cost-effectiveness is the collective perspective. It makes it possible to consider the consequences for all of the individuals and institutions involved in the production

of the overall care, whether in the domestic sphere (users and their informal caregivers), the healthcare sphere (care providers) or the medico-social sphere (providers of medico-social aid[4]).

When the collective perspective cannot be documented in the reference case analysis*, a healthcare system perspective may be opted for. It differs from the collective perspective, in that it focuses solely on health care production[5], while the collective perspective covers the overall care provided to the person.

The evaluation of health outcomes and costs depends on the chosen perspective.

- The evaluation of health outcomes identifies the relevant health effects from the point of view of the populations concerned (i.e. patients, healthcare system users and informal caregivers under a collective perspective; only patients and users under a healthcare system perspective). The health effects are then measured according to physical units (life years), which may be valued on the basis of preferences. In the latter case, the preferences considered are those of the general population (see page 32, Measurement and valuation of health-related quality of life in cost-utility analyses).
- The cost evaluation identifies, measures and values all of the resources consumed in the production of care, i.e. from an overall viewpoint under a collective perspective, and from a restricted viewpoint under a healthcare system perspective.

The social perspective is not recommended in the reference case analysis* as it involves the use of resources which are not directly included in the care production process. The social perspective requires the factoring-in of aspects such as the value of production losses resulting from the illness.

Any impossibility to conduct the evaluation according to a collective perspective should be duly justified, as well as the choice of the perspective adopted[6].

1.4. Choice of the population analysed

Guideline 4

The population analysed should consist of all individuals whose health is affected –whether directly or indirectly – by the intervention being evaluated. Any inability to include certain affected individuals in the reference case analysis should be duly justified.

The economic approach may call for the analysis of specific population subgroups when differences in care imply different comparators or when variability is expected in terms of health effects or costs.

1.4.1. Specification of the population analysed

The population analysed* consists of all individuals whose health is affected by the intervention evaluated, whether directly (sick persons, the vaccinated population, the population to be screened, etc.) or indirectly (the general population in the case of vaccination, caregivers, etc.).

The individuals whose health is directly affected by the intervention evaluated are systematically included in the population studied in the reference case analysis*.

[4] Medico-social institutions and services dedicated to long-term care, disabled persons, etc.

[5] Under the healthcare system perspective, informal care is no longer considered as a resource if it mainly involves personal care (helping the patient to eat, get dressed, etc.), daily tasks (shopping, preparing meals, etc.) or patient surveillance. Likewise, consequences in the medico-social sphere are no longer considered.

[6] As a reminder, the scope of the economic evaluation is not limited to the efficiency opinions issued by HAS. Some economic evaluations may require a specific perspective to address a specific objective, such as a healthcare institution perspective when a decision needs to be taken on a hospital investment, or a national health insurance perspective.

When the evaluated intervention has consequences on the health of other individuals, the population analysed* may be extended to those individuals. Examples include the impact on the health of care-givers, the protection of non-vaccinated persons, or the negative effect of antibiotic therapies in the event of the spread of resistance to antibiotics.

Any inability to include in the evaluation certain individuals whose health is likely to be affected by the intervention evaluated should be duly justified.

Specificity of an evaluation produced within the framework of Article R. 161-71-3 of the French Social Security Code

Within the specific framework of CEESP opinions on health products, the population analysed* is defined by Article R. 161-71-3 of the French Social Security Code, which states that CEESP shall issue an opinion on the expected or observed cost-effectiveness of the care provided or technology covered by the national health insurance system. In such opinions, the population analysed* corresponds to the population of the indication for which the French Transparency Committee (CT) or National Committee for the Assessment of Healthcare Devices and Technologies (CNEDIMTS) have advised reimbursement by the national health insurance system.

1.4.2. Identification of the population subgroups analysed

The economic evaluation requires the defining of specific population subgroups if differences in health effects[7] or costs are expected, or if the comparators differ in certain subgroups.

When the population is divided into subgroups because of the difference in expected health effects, the analysis of the difference in health outcomes for each subgroup rests on randomised controlled trials or other types of studies that comprise an analysis with these subgroups in their protocol. An exploratory analysis may be conducted on the basis of data with a lower level of evidence.

When the analysis of certain subgroups is required for other reasons (e.g. different costs or comparators) and no trial has demonstrated different health effects between these subgroups, an assumption of invariance of the health effect of the intervention evaluated is to be used. The treatment effect applied to each subgroup is the effect assessed for the wider population included in the clinical trials. This assumption is subject to a sensitivity analysis.

Specificity of an evaluation produced within the framework of Article R. 161-71-3 of the French Social Security Code

Within the specific framework of CEESP opinions on health products, the cost-effectiveness of the product is assessed for each subgroup analysed, corresponding to each level of clinical added value (CAV) claimed.

1.5. Choice of interventions to be compared

Guideline 5

The reference case analysis should identify all clinically relevant interventions in the population analysed.

The arguments supporting the inclusion or exclusion of an intervention in the reference case analysis should be clearly set out.

[7] On the clinical level, the relevance of subgroup analyses is supported by the likelihood of the variation of the treatment effect in a particular subgroup in relation to the total trial population (based on pharmacological, biological and clinical arguments).

The interpretation of the results of the evaluation should consider the degree of exhaustiveness of the comparators used. Given that all comparators need to be taken into account to evaluate cost-effectiveness, the risk of excluding an intervention likely to be on the cost-effectiveness frontier, and its consequences, should be discussed.

All interventions should be identified.

The reference case analysis* should identify all clinically relevant interventions in the population analysed[8]. They may consist of different types of interventions: medical treatment, medical device, surgery, palliative care, therapeutic abstention, prevention, non-medicinal therapy, etc.

When certain interventions are specific to a subgroup, the results are to be estimated per subgroup in the reference case analysis, or in any supplemental analysis if the subgroup data are exploratory. The possibility of interpreting an estimated result in terms cost-effectiveness for the population analysed is to be discussed[9].

The exclusion of an intervention should be justified.

The cost-effectiveness of an intervention can only be estimated if the evaluation includes all relevant interventions, as the omission of an intervention is likely to alter the cost-effectiveness frontier*, and possibly the ICER estimation in the case of the intervention that precedes the intervention evaluated on the cost-effectiveness frontier*. The interpretation of the results of the evaluation thus depends on the interventions included in the reference case analysis*. This involves the following:

- Among the relevant interventions identified, the evaluator needs to justify his/her choice of the interventions included in or excluded from the evaluation.
- The impact of the exclusions on the interpretation of the evaluation results in terms of cost-effectiveness should be discussed and, in particular, the risk of excluding an intervention likely to be on the cost-effectiveness frontier and its consequences.

The choice of the interventions to be used in the reference case analysis* should first be based on the good practice guidelines produced and published by scientific institutions under conditions excluding conflicts of interest.

- Comparators include interventions for which there is published clinical data, and health products for which there are published prices or maximum compensation amounts. The medicines under evaluation by EMA, and which meet those conditions, may be included if they are covered by a temporary usage authorisation (TUA), a post-TUA programme, or an early filing procedure with HAS.
- The inclusion of non-recommended interventions or the exclusion of recommended interventions should be justified in detail, on a case-by-case basis (e.g. context specificity, data availability, etc.). The interest of envisaging options such as non-medicinal therapeutic options or therapeutic abstention is reiterated.
- The low use of an intervention is not sufficient reason to justify its exclusion from the analysis, if it is medically relevant (e.g. emerging intervention; intervention from the past). There is no *a*

[8] In the specific case of an evaluation attached to a clinical trial (piggy-back study*), the economic evaluation measures the differences in cost and effectiveness between the interventions included in the trial. In general, the results cannot be interpreted in terms of efficiency if the options studied are not exhaustive right from the start of the care.

[9] In particular, the comparator used for the estimation of the ICER (prior option on the cost-effectiveness frontier) must be the same for the different subgroups composing the population analysed. If the ICER cannot be interpreted in efficiency terms, the economic information available should be presented in terms of differences in costs and health outcomes.

priori reason for considering that a little-used intervention in standard practice is not on the cost-effectiveness frontier*.

- Medicines used without a marketing authorisation (MA) may be used in the reference case analysis* if they are widely used in common practice. This in no way validates inappropriate use, but merely confirms a practice. Such a practice may cover a wide range of situations such as a manufacturing strategy, a therapeutic dead-end, or scientific publications suggesting its interest.

Second, the choice of comparators should be based on explicit and justified trade-off between the information generated by the inclusion of a comparator and the biases associated with available data. The impact of the exclusion of a comparator is to be tested in a sensitivity analysis.

The reference case analysis is time specific.

Any change in the healthcare strategy – in particular the arrival of a new comparator on the market – will invalidate the previous cost-effectiveness evaluation.

The evaluation takes account of the changes in the characteristics of technologies over time (performance, cost, etc.). Where the introduction of an intervention entails the withdrawal of another intervention (e.g. change in vaccination procedure, new version of a medical device[10] or vaccine, industrial development strategy, etc.), the intervention replaced is to be included in the analysis in order to inform the decision-maker of all the consequences of the substitution.

The anticipation of changes in practices, along with their learning, should also be discussed.

1.6. Choice of a time horizon

Guideline 6

The time horizon of the evaluation may be defined as the entire lifetime or a specific duration.

The choice of the time horizon should be based on an trade-off between the information produced by considering the expected or observed consequences of the intervention over the long term and the uncertainty generated by extrapolation over time. This trade-off should be clearly set out and the choice should be justified.

The time horizon used should be the same for all interventions.

The time horizon* is the period of time during which costs and health effects are considered for the evaluation. The choice of the time horizon* for the reference case analysis* should be explicit and justified.

The two recommended options are:

- a lifetime* horizon, i.e. until death;
- a specific time horizon* (e.g. to a defined age or over a defined period).

The choice of the time horizon* should be justified, in particular in relation to the natural history of the disease, the persistence of costs and health effects associated with the interventions, life expectancy, and the quality of the long-term data.

[10] In the absence of specific data on the device for which the request is made, and when the changes are incremental, extrapolation of the data available for prior versions is acceptable, provided that the incremental character of the changes is documented. The impact of this assumption on the relative effectiveness of the device and its adverse effects is to be tested in a sensitivity analysis.

The choice of the time horizon* should mainly be based on an trade-off between the information produced over a time horizon* which is long enough to reflect all differences in costs and health effects, as well as the uncertainty generated by extrapolation over time when this is necessary. This trade-off should be explained and justified.

- A lifetime horizon* is used if at least one of the interventions has an impact over the patient's entire life, either in terms of costs, length of life, quality of life, morbidity, deficiencies or incapacity (chronic disease or debilitating disease) and if the uncertainty associated with the lifetime extrapolation is acceptable. This acceptability should be duly justified.
- A specific time horizon*, defined over a shorter period than life expectancy, is appropriate if differences in costs and health outcomes are no longer observed beyond a certain period of time or if the uncertainty inherent in the extrapolation of observed lifetime data is not acceptable.

In certain cases (such as vaccination), a multigenerational time horizon may be necessary.

Only the health effects and costs occurring over the time horizon* should be considered in the evaluation. All of the interventions compared should be evaluated over the same time horizon.

1.7. Discounting method

Guideline 7

Future costs and health outcomes should be discounted to present values whenever the time horizon exceeds 12 months.

The reference case analysis should use the public discount rate applicable at the time of the evaluation. On the date of publication of this guidance, that rate was 2.5% for time horizons of less than 30 years. Beyond that, the rate gradually decreases to a floor level set at 1.5%.

In the reference case analysis, costs and health outcomes should be discounted using the same rate.

Discounting* makes it possible to compare interventions at different points in time by assessing future costs and health outcomes at their present values. Expected costs and health outcomes should be discounted whenever the time horizon* exceeds 12 months.

The discount rate gives a present value to future costs and health outcomes.

In accordance with HAS' mandate, the discount rate used in the reference case analysis* is the rate which applies to all public investment decisions.

In France, this rate has been defined by the group of experts chaired by E. Quinet (Commissariat général à la stratégie et à la prospective 2013). "It reflects the relative price in present terms and sets the limit we are willing to accept for the future"[11]. It has been set at 2.5% since 2013. It applies to amounts expressed in constant currency (without inflation). It is revised on a regular basis; the rate used in the evaluation is the rate applicable at the time of the evaluation.

The discount rate does not take account of the uncertainty associated with the interventions, which is dealt with separately.

[11] The discount rate depends on the pure preference rate for the present, the elasticity of the marginal utility of consumption, and the consumption growth rate per capita (French Planning Commission, 2005).

Discounting implies predicting the relative prices of the goods involved in the interventions evaluated.*

In the reference case analysis*, the relative price of the health outcome[12] for the community does not vary over time. Costs (expressed in monetary units) and health outcomes (expressed in their own measure unit) are thus discounted at the same rate.

The value of the discount rate may vary with the time horizon.

When the time horizon* is very long, like for vaccination, the discount rate gradually decreases after 30 years, down to a floor rate of 1.5% (Lebègue 2005).

The value of the discount rate is subject to a sensitivity analysis.

To appraise the robustness of the conclusions of the evaluation, sensitivity to the discount rate is tested using, as a minimum, an increased rate (4.5%) and a zero rate.

The discount rate sensitivity analysis is conducted on costs and health outcomes at the same time.

[12] The relative price of the health outcome is taken to be the reference value of a unit of that good, relative to that of the monetary unit deemed as constant.

2. Methodological choices for the evaluation of health outcomes

2.1. General principles

Guideline 8

The consequences to be considered in the economic evaluation are the effects of the interventions on the health of the population analysed.

The consequences of the interventions to be considered in the evaluation are their effects on health.

Certain developments in economic literature strive to consider effects beyond the health aspect[13]. At present, these approaches are not recommended in reference case analysis*. They may be presented as a supplemental analysis*.

The reference case analysis* identifies all of the consequences that may affect the health of the individuals included in the population analysed*. Special attention should be paid to the identification and analysis of the adverse effects associated with each intervention.

2.1.1. Health outcome criteria in cost-effectiveness analyses

Guideline 9

If the reference case analysis is a cost-effectiveness analysis, the health outcome criterion to be used should be that of life-years. The mortality indicator should be the all-cause mortality rate.

If the data required for the measurement of life-years are unavailable, the use of a predictive criterion of expected survival time may be acceptable, but only if there is strong, established evidence of the predictive nature of this surrogate endpoint.

A cost-effectiveness analysis may be based on other health outcome criteria in a supplemental analysis, with arguments supporting the choice of the criteria.

In the cost-effectiveness analysis, the preferred health outcome criterion is length of life.

When a cost-effectiveness analysis* is conducted, length of life is the criterion used to evaluate cost-effectiveness in the reference case analysis*. The indicator used to measure this criterion is the all-cause mortality rate.

If it is demonstrated that all interventions are equivalent on the length of life criterion, the reference case analysis* should be based on cost minimisation. A supplemental analysis* may be presented, for information purposes, using another health outcome criterion. The choice of this criterion should be duly justified.

[13]Three approaches have been put forward in published literature: measuring subjective wellbeing (Dolan and Metcalfe 2012) (P. Dolan 2013), the use of a monetary metric for health equivalent income that introduces an equity criterion (Fleurbaey, Luchini and Schokkaert 2013) (Samson et al. 2018), the use of a capability metric (Lorgelly 2015).

In the absence of survival data, cost-effectiveness is difficult to interpret.

Measuring a relative effect on the length of life is essential to estimate the cost-effectiveness of health interventions.

If the data required to measure life-years are not available, a survival prediction criterion may be used, but only if there is strong, established evidence of the predictive character of this surrogate endpoint*. The correlation factor should be presented and duly justified. The uncertainty generated by the predictive relationship should be explored through a sensitivity analysis.

In the absence of this scientific demonstration, the reference case analysis should use the assumption of the lack of difference in survival between the intervention evaluated and its comparators. Where relevant, the results of an evaluation based on an assumption of higher survival should be presented in a scenario analysis*.

A cost-effectiveness analysis* may be conducted using a health outcome criterion other than length of life in a supplemental analysis (e.g. analysis based on morbidity). The results of the evaluation should be presented in terms of differences in costs and health outcomes. In the absence of a metric based on survival, the construction of an cost-effectiveness frontier makes it possible to examine the technical cost-effectiveness* of the intervention evaluated. However, the magnitude of the estimated ICER or INB is difficult to interpret in terms of allocative cost-effectiveness*.

2.1.2. Health outcome criteria in cost-utility analyses

Guideline 10

If the reference case analysis is a cost-utility analysis, the health outcome criterion to be used should be that of quality-adjusted life years (QALYs).

The health-related quality of life is evaluated on the basis of health state preferences measured through utility scores.

The health-related quality of life is evaluated through health state preferences.

When the health-related quality of life is identified as a major consequence of the intervention evaluated, the health outcome criterion to be used is the quality-adjusted life years.

The health-related quality of life* is measured through a utility score, reflecting preferences for different health states (Weinstein, Torrance et McGuire 2009) (Drummond, Sculpher et Torrance 2005).

The preferred health outcome criterion is QALY (quality-adjusted life year).

The number of QALYs allocated to an intervention is calculated by weighting the time periods spent in the different health states characterising the development of the disease with the utility scores associated with these states.

The weighting of QALYs according to the individual characteristics of the persons involved in the intervention (socio-demographic factors, severity, etc.) is still a topic of debate. Numerous issues, in particular concerning methodology and ethics, have not yet been resolved (Baker, Bateman et Donaldson 2010). At present, this type of weighting is not recommended.

2.2. Quantification of comparative effectiveness and safety

Guideline 11

To conduct an economic evaluation, evidence-based comparative clinical data are essential, in particular when the cost-effectiveness of an intervention needs to be established.

2.2.1. Identification and selection of available data sources

Guideline 12

All clinical data sources for the intervention being evaluated and its comparators should be identified through a systematic, reproducible methodology, and then presented in accordance with international standards and HAS guidelines.

Available data sources to document the effectiveness and tolerability of each intervention should undergo a stringent critical appraisal. The sources with the highest level of evidence should be selected.

For each selected source, the transposability of experimental health outcomes to standard French practices should be discussed.

Identification of available clinical data sources for the intervention and its comparators

Available data are identified through a systematic literature review.

The data describing the clinical effectiveness and tolerability of the intervention and its comparators are identified through a systematic literature review conducted in accordance with the PRISMA-P guidelines (Moher, Shamseer et Clarke 2015) or Cochrane Handbook (Higgins, et al. 2019), and reported according to the PRISMA guidelines (Liberati, et al. 2009) (Gedda 2015)[14].

The document research field covers all comparators identified as relevant (see page 20, Choice of interventions to be compared).

The level of evidence of clinical studies should be assessed.

A study's level of evidence is determined by the adequacy of the protocol in relation to its objectives (study design, population, assessment criteria and statistical analysis method), the existence (or not) of major biases in its production, and its statistical power. In 2013, HAS published a review of the systems used to assess the level of evidence in scientific studies (Haute Autorité de Santé 2013), which did not call into question the criteria defined by ANAES in 2000 (see Annex 3, page 69).

The level of evidence of useful data for subgroup analysis depends on two characteristics of the trial protocol: the pre-specification of the subgroup analyses; the stratification of randomisation according to these subgroups, with sufficient power.

[14] The PRISMA guidelines for the writing and reading of systematic reviews and meta-analyses were first published in 2009 and are currently being updated.

Experts' opinions

The level of evidence associated with experts' opinions is low. The use of experts' opinions to make up for the absence or shortage of related treatment effect data is not acceptable in reference case analyses.

Experts' opinions may be used to justify the choice of the data or to justify the relevance of the data or assumptions tested in a sensitivity analysis, so long as the method used to obtain these opinions is detailed (criteria used to select the experts, number of experts approached and who responded, disclosures of potential interests, method used to record the opinions, questions asked, and identification of the data documented through experts' opinions).

For quantitative parameters, a formal method of elicitation is preferable (see Annex 13, page 100).

Selection of studies and presentation of health outcomes

The selection of the studies used to estimate the clinical effectiveness and tolerability parameters of the intervention and its comparators rests on the critical appraisal of the studies identified through the systematic literature review.

Where a study is not selected, its conclusions should be briefly presented, and the reason for its exclusion should be explained.

The selected studies should be presented in a detailed manner in a table (methodology, results, main elements required for their interpretation, level of evidence). When known, the sources of variability of the expected effect should be documented and described so they can be considered in the analysis (interaction variables such as age, gender, disease severity, etc.).

Ideally, the same sources should be used to estimate clinical effectiveness and tolerability. Justification is required when the sources used are different.

For each selected source, the transposability of experimental health outcomes to standard French practices should be discussed, in particular with regard to the methodological rigour of the selected trials and the identification of the sources of bias that can influence the effect of an intervention.

These various elements should be reviewed in order to discuss the available data's ability to document the evaluation's clinical parameters.

2.2.2. Estimation of comparative effectiveness

Guideline 13

Direct comparisons in randomised controlled trials and meta-analyses are the best source of evidence of comparative effectiveness.

If the interventions studied have not been directly compared in the same study, an indirect comparison method should be used to estimate the differences in effectiveness between the interventions and a "reference" intervention.

- The choice of the reference intervention should be justified.
- The method used to estimate the relative effect with respect to the reference intervention should be the same for all interventions.

The choice of the comparative effectiveness evaluation method should be justified, and the method should be presented in a clear and detailed way.

– A network meta-analysis is recommended, subject to the validation of the underlying feasibility assumptions (consistency of populations, protocols and related bias risks, and verification of the transitivity assumption).
– The use of an indirect comparison method adjusted on the basis of individual data may be put forward if justified.

The clinical assumptions and methodology choices made to estimate comparative effectiveness should be described, justified and tested in a sensitivity analysis.

The level of evidence of the estimated comparative effectiveness should be clearly set out.

Choice of the reference intervention

Comparative effectiveness*, or the difference in effectiveness between interventions, is estimated in relation to a reference intervention.

The choice of the reference intervention, as well as the method of estimation of the treatment effect associated with this reference intervention, should be justified (Cooper, Sutton et Achana 2015) (Dias, Welton et Sutton 2013).

In most cases, the reference intervention is the long-standing intervention which is the most widely used in common practice. If another intervention is chosen, this choice should be justified and tested in a sensitivity analysis.

Methods to estimate comparative effectiveness

The method used to estimate comparative effectiveness is also used to estimate the difference in effectiveness between the reference intervention and each of the interventions. If the same method cannot be used, the choice of the methods used should be justified.

Direct comparisons and pairwise meta-analyses

Randomised controlled trials directly comparing the intervention being evaluated with relevant comparators, along with pairwise meta-analyses, constitute the best source of evidence of comparative effectiveness*.

Data stemming from randomised controlled trials on a representative segment of the population analysed* are preferred.

When there are several studies comparing two interventions on the same criteria using comparable protocols, the treatment effect estimations will need to be summarised through a pairwise meta-analysis*. The meta-analysis should be conducted in accordance with existing guidelines (Moher, Shamseer et Clarke 2015).

When a study cannot be included in a pairwise meta-analysis due to data unavailability or a protocol which is not comparable to that of the other studies, its results should be presented and used to discuss the uncertainty of estimated relative effectiveness by excluding it.

Indirect comparisons stemming from network meta-analyses

If direct comparison* of the interventions is not possible, the reasons for this situation should be explained and an indirect comparison method should be used.

The network meta-analysis* should include all selected interventions and be conducted in accordance with existing guidelines (Jansen, et al. 2014) (Hutton, et al. 2015) (Cooper, Sutton et Achana 2015) (Ades, Caldwell et Reken 2012). As set out in the section "Selection of studies and presentation of

health outcomes", the methodology, the characteristics of the patients upon their inclusion, and the results of the trials included in the network meta-analysis* should be reported in a table. The level of evidence of each relative effect assessed should be specified (Puhan, Schünemann et Murad 2014). As much information as possible should be input into the network in order to reduce the uncertainty of estimations, while preserving the consistency of the network and avoiding the introduction of a bias likely to undermine the analysis.

The clinical assumptions and methodological choices underlying the meta-analysis should be explained and justified.

- Network: number of trials contributing to the network, number of trials for each comparison, justification of excluded trials
- Statistical model(s) used, and assumptions adopted (e.g. assumption of the proportionality of risks)
- Assumptions: grouping of interventions, assumptions of equivalence, existence of covariates interacting with the treatment effect
- Heterogeneity analysis
- Sensitivity analyses

Estimations stemming from a network meta-analysis* should be presented in a table and in graphs such as forest-plot graphs.

Bayesian and frequentist approaches are acceptable (Dias, Welton et Sutton 2013) (Dias, Sutton et Welton NJ 2013) (Song, Loke et Walsh 2009). The approach adopted should be clearly explained. Any heterogeneity of results stemming from direct comparisons* and inconsistencies between results stemming from direct comparisons* and indirect comparisons* should be discussed.

Matching-Adjusted Indirect Comparisons

The use of Matching-Adjusted Indirect Comparisons* (MAIC) and Simulated Treatment Comparisons (STC) is possible when a network meta-analysis* is not suitable and when the networks include at least one study with individual data.

The impossibility of conducting a network meta-analysis* should be duly justified. In particular, the use of matching-adjusted indirect comparisons* can be envisaged in two situations:

- The selected trials form a connected network but one or more interaction variables alter the effect of the treatment assessed in the studies, and in addition, the distribution of these interaction variables differs among the various selected trials (high heterogeneity of the trials composing the network).
- The selected trials form a disconnected network (lack of a common comparator or inclusion of a single-arm trial).

These studies should be conducted in accordance with current guidelines (Phillippo, Ades et Dias 2016).

The consistency between the pseudo-population stemming from the individual data adjustment and the population analysed in the evaluation should be discussed.

Comparisons stemming from observational studies (real-life data)

The methodology, patient characteristics and results of the selected observational studies should be reported in detail (Strobe - Strengthening the reporting of observational studies in epidemiology 2007).

Statistical methods recommended for the production of valid comparative data should be used (e.g. propensity scores) (Faria, Hernandez Alava et Manca 2015).

In particular, the selected observational studies' biases should be identified and quantified and their results should be adjusted accordingly.

Comparative effectiveness estimation not recommended

Naive indirect comparisons* are not acceptable to evaluate the relative effect of an intervention in a reference case analysis.

Where only naïve indirect comparisons can be made, the analysis is then exploratory and cannot come to any conclusion as to the cost-effectiveness of the intervention evaluated.

2.2.3. Consideration of adverse effects

Guideline 14

For the collection and processing of tolerability data, the same methodological rigour should be applied as for the collection and processing of effectiveness data.

The frequency and severity of adverse effects should be reported in detail for all interventions, along with their duration and time to onset.

General principle

Tolerability data are included in the evaluation by selecting the adverse effects likely to have a differential impact on costs and health outcomes.

To estimate tolerability, the same methodological rigour should be used as for the estimation of effectiveness. The available data should be identified and combined in a clear, systematic and robust way, with a description of the limitations associated with the data or the method used to estimate adverse effects (AE).

The systematic review of sources and data should be reported in accordance with the PRISMA guidelines and its specific check-list for the description of adverse effects (Zorzela, Loke, & Ioannidis, 2016).

Identification and selection of adverse effects

The identification of adverse events is based on the tolerability profile of the intervention evaluated and its comparators. For each of the interventions, a summary of the available data should be presented, providing up-to-date knowledge on the adverse effects (irrespective of the grade) associated with these interventions, and more specifically on the high-grade or serious adverse effects associated with the interventions.

The criteria used for the selection of the adverse effects to be considered in the evaluation should be clearly set out and justified. In particular, the adverse events associated with the intervention and its comparators should be defined with the same level of precision (ICH 2019). The selected adverse effects should be precisely described: adverse effects occurring a single time or in a recurring way; temporary or irreversible adverse effects; occurrence of adverse effects over time (duration and time of occurrence), etc.

It is important to ensure that the method used for the selection of adverse effects reflects the tolerability profile of the intervention evaluated and that of its comparators in a well-balanced way in terms of the frequencies reported in the studies.

Method used to estimate the probability of occurrence of adverse effects

The heterogeneity of sources should be discussed and, where appropriate, considered when it is likely to affect the estimation of the probability of occurrence of adverse effects in the interventions. For example, this may consist of differences between studies in terms of follow-up duration or the duration of exposure to the interventions.

When several studies are available to estimate the probability of occurrence of adverse effects in an intervention, the choice of aggregating all available data (or not) using an appropriate method should be discussed.

The evaluation of the impact of adverse events should take account of the recurrence of the events. In this respect, the estimated frequency used in the evaluation should be the total event frequency observed during the trial rather than the frequency of patients having experienced at least one event.

2.3. Measurement and valuation of health-related quality of life in cost-utility analyses

2.3.1. Utility score estimation method

Guideline 15

The utility scores used to weight life-years should be estimated using a multi-attribute approach, comprising the collection of health state data from patients via a generic questionnaire and the valuation of health states according to the preferences of the general population.

Among available classification systems, EQ-5D-5L is recommended (French EQ-5D-5L value set and EQ-5D-5L questionnaire). The French value set which prevails at the time of the evaluation should be used.

Failing that, as a transitional measure, the EQ-5D-3L classification system (French EQ-5D-3L value set and EQ-5D-3L questionnaire) should be used.

In the absence of a utility score from the EQ-5D system, a mapping approach should be opted for, in order to arrive at an EQ-5D utility score, provided that a mapping function based on the methodological quality standards is available and has been validated.

Other approaches are not recommended for the base-case analysis of the reference case (see specific questionnaire with the valuation of preferences in the general population, standard gamble and time trade-off methods that directly value the utility associated with the patients' health state, and the ordinal approach). These can be subjected to a sensitivity analysis.

The estimation of utility scores through an approach revealing preferences for a hypothetical health state via vignettes or through a visual analogue scale is not acceptable, even within the framework of a sensitivity analysis.

Estimation of utility scores based on a generic multi-attribute approach

In a cost-utility analysis*, health-related quality of life is measured through a utility score*. The various approaches available to estimate a utility score* are presented in the insert below (Brazier, Ratcliffe, et al. 2017).

Insert 1: Different approaches used to estimate a utility score (Brazier et al., 2017)

1 _ Multi-attribute approaches, based on general population preferences

Health state data collected from the population analysed via a questionnaire, then rated through a multi-attribute utility function defined according to general population preferences.

– Generic questionnaire with valuation based on general population preferences
– Specific questionnaire with valuation through a mapping function, in order to arrive at a generic measure based on general population preferences
– Specific questionnaire with valuation based on general population preferences

2 _ Individual preference revelation approaches

– Patient estimation of the utility score associated with his/her health state

The patient directly assesses his/her health state through a choice-based method (standard gamble, time trade-off, visual analogue scale).

– Estimation by an individual of the utility score associated with a hypothetical health state

The respondent is provided with the description of a health state (via a vignette) and values it using a choice-based method (standard gamble, time trade-off, visual analogue scale).

3 _ Ordinal approaches (discrete-choice method, ranking method)

The respondent is provided with descriptions of several health states that he/she compares and ranks. Statistical models are then used to derive utility scores associated with the health states.

Utility scores are estimated using a generic multi-attribute approach.

Among the different approaches, the use of a multi-attribute approach – which makes it possible to estimate the utility score* associated with a health state described according to several attributes – is recommended.

To promote the comparability of economic evaluations, an approach leading to a generic measurement of health states should be used.

The health state data should be collected from directly concerned individuals and its valuation should be based on general population preferences.

The health state data should be collected from a representative sample of the population analysed. When this information cannot be collected from the individuals directly involved (e.g. young children, mentally ill), the data may be collected from a parent/friend or a healthcare professional but needs to be duly justified (see section 2.3.3, page 38).

The valuation of health states via utility scores reflects the preferences of the general French population regarding these health states. The choice of general population preferences is based on the argument that health interventions are mainly state funded in France.

The EQ-5D classification system is recommended for the reference case analysis

The use of the EQ-5D classification system is recommended for the estimation of utility scores

Among the seven main generic multi-attribute classification systems available[15], only the EQ-5D-5L (Andrade, Ludwig, & Goni, 2020), EQ-5D-3L (Chevalier & de Pouvourville, 2013) and HUI systems (Le Gales, Buron, & Costet, 2002) currently have a value set based on a representative sample of the general French population.

To ensure the consistency of the methodology applied in the studies, the EQ-5D system is recommended for the base-case analysis* of the reference case analysis*. The EQ-5D classification system is composed of a) a generic questionnaire that makes it possible to describe the health state according to 5 dimensions (mobility, self-care, usual activities, pain/discomfort, anxiety/depression) and b) a value set to estimate the multi-attribute utility of the health state described.

- Initially, each of the 5 dimensions were described through 3 response levels (EQ-5D-3L, Annex 4, p.70). Since 2009, a new version has been available, in which each of the 5 dimensions is described through 5 response levels, thus improving the sensitivity of the instrument (EQ-5D-5L, Annex 5, p.72).
- The French value set which prevails at the time of the evaluation should be used. At the time of publication of this document, the value set published by Andrade et al. is the one recommended (Andrade, Ludwig, & Goni, 2020) from the age of 16 years. A sensitivity analysis using the algorithm developed by EuroQol to value the EQ-5D-5L based on French EQ-5D-3L[16] scores should systematically be presented during a transitional period of one year starting on the date of validation of this guidance document by the Board (van Hout & Janssen, 2012) (EuroQol).

In the absence of health state data from an EQ-5D-5L questionnaire, the EQ-5D-3L classification system should be used as a transitional measure.

A mapping approach may be used to obtain an EQ-5D score

A mapping* function makes it possible to derive an EQ-5D score from the health state data collected through questionnaires other than the EQ-5D-5L or EQ-5D-3L questionnaires, but this increases uncertainty and the risk of error (Brazier, Yang et Tsuchiya 2008) (Rowen, Brazier et Roberts J. 2009) (Longworth et Rowen 2011). Moreover, there are few studies which validate such functions in France.

The validity conditions for the mapping* function used should be discussed with regard to the current guidelines (Longworth et Rowen 2011) (Longworth et Rowen 2013) (Wailoo, Hernandez et Manca 2016).

The choice of the mapping* function used should be duly justified, in particular its application to the simulated population, with regard to clinical characteristics (in particular severity) and demographic specificities, also taking into consideration the availability in the trial of variables integrated in the mapping* function. If several valid mapping* functions are available, the functions not used in the base-case analysis* should be tested in a sensitivity analysis.

The Health Economics Research Centre of the University of Oxford publishes a database of studies mapping* to EQ-5D-3L, updated every year (Dakin 2013) (HERC 2016)[17]. At the time this guidance document was written, there was no available function for mapping* to the EQ-5D-5L classification system.

[15] The seven instruments are: QWB (1976), 15D score (1989), EQ-5D-3L(1995), HUI3 (2002), SF-6D (2002), AQoL-8D (2009), and EQ-5D-5L (2009).

[16] The EuroQol Group coordinated a study covering both the EQ-5D-3L and EQ-5D-5L systems, in order to develop a mapping function between the EQ-5D-3L value set and the EQ-5D-5L descriptive questionnaire. The methodology and the value set used to estimate the EQ-5D-3L score based on EQ-5D-5L questionnaires is available on the website https://euroqol.org/eq-5d-instruments/eq-5d-5l-about/valuation-standard-value-sets/crosswalk-index-value-calculator/

[17] The database is available on the website https://www.herc.ox.ac.uk/downloads/herc-database-of-mapping-studies

The uncertainty generated by the parameters of the mapping* function is to be included in the sensitivity analysis (Wailoo, Hernandez et Manca 2016).

The use of any other approach is not recommended.

Classification systems based on a specific questionnaire with the valuation of general population preferences

Numerous classification systems resting on a disease-specific questionnaire and a valuation method based on general population preferences have been developed[18] (Brazier, Harper et Thomas 1998) (Brazier, Roberts et Plattas 2005) (Glied et Smith 2013), in order to overcome the criticised lack of sensitivity of systems based on a generic questionnaire (Brazier, Ratcliffe, et al. 2017) (Garau, Shah et Towse 2009.) (Lorgelly, Doble et Rowen 2014).

To date, these classification systems have different levels of validity and none of them has a French value set. Moreover, the use of specific questionnaires does not allow the comparison of interventions in different therapeutic areas.

The results drawn from these specific classification systems may be included in a sensitivity analysis if the data are available.

Individual preference revelation approaches

The individual preference revelation approaches covered here are the Standard Gamble (SG), the Time Trade-Off (TTO) and the Visual Analogue Scale (VAS).

There are two different scenarios, depending on whether the questions are put to a patient on his/her own health state or to an individual from the general population on a hypothetical health state.

In the first case, the three methods mentioned can be used to estimate the utility score* that the patient associates with his/her own health state. The utility score obtained is based on the patient's own health preferences rather than on those of the general population, as recommended earlier.

Moreover, numerous publications have shown that, for the same health state, utility scores estimated through the standard gamble and time trade-off methods diverge, due to factors that are not associated with health state preferences. To date, there is no decisive argument in favour of either of these two methods (Brazier, Ratcliffe, et al. 2017). For the sake of comparability of the studies, these methods are not recommended in base-case analyses.

The results derived from a standard gamble or time trade-off method, involving the patient's estimation of his/her own health state, may be included in a sensitivity analysis if the data are available.

Visual analogue scales are not recommended, whether in a base-case analysis* or a sensitivity analysis*, given that publications have demonstrated that they do not constitute a valid basis to estimate utility scores (Torrance, Feeny et Furlong 2001) (Brazier, Ratcliffe, et al. 2017).

In the second case, the individuals questioned have not experienced the health state evaluated. It is thus necessary to provide them with a description of the health state through vignettes. These methods can be used to estimate a utility score* associated with this health state, based on the respondent's preferences. The choice-based methods (standard gamble, time trade-off, and visual analogue scale) used to evaluate hypothetical health states described using vignettes are not recommended, whether

[18] For example, questionnaires are available for asthma (Revicki, Leidy, & Brennan-Diemer, 1998) (Kharroubi, 2014) (Yang, 2011), rhinitis (Revicki, Leidy, & Brennan-Diemer, 1998), stroke (Poissant, Mayo, Wood-Dauphinee, & Clarke, 2003), visual functions (Rentz, Kowalski, & Walt, 2014), menopause (Brazier, Roberts, & Plattas, 2005), multiple sclerosis (Goodwin, Green, & Spencer, 2015), over-active bladder (Yang, Brazier, & Tsuchyia, 2009) (Kharroubi, 2014), urinary incontinence (Brazier J., 2008), multiple myeloma (Rowen, Brazier, & Young, 2011) and palliative radiotherapy (Costa, Aaronson, & Fayers, 2014). An instrument applicable to all cancers (QLU-C10D) is under development (King, Costa, & Aaronson, 2016) (Norman, Viney, & Aaronson, 2016).

in a base-case analysis* or sensitivity analysis*, due to the numerous limitations associated with the use of vignettes (low-quality evidence, lack of comparability between studies, including for the same pathology, lack of standardisation, non-representativeness of the distribution of possible health states) (Brazier, Ratcliffe, et al. 2017).

Ordinal approaches (Discrete-Choice Method, Ranking Method)

Ordinal methods (e.g. discrete-choice method, ranking method, best-worst scaling method) provide an order of preference for health states. The respondent compares the health states described according to several dimensions and ranks them (ordinal metric). Statistical models are then used to derive utility scores associated with these health states (cardinal metric).

These methods – initially used to assess how an individual values and arbitrates between the characteristics of healthcare programmes, or to establish the healthcare programmes' willingness to pay – have raised interest over recent years for the estimation of utility scores associated with a health state (Ali et Ronaldson 2012) (Brazier, Rowen et Yang 2012). Yet, in the absence of a standardised methodology, these methods are not recommended for the base-case analysis* of the reference case*.

The results derived from an ordinal method may be included in a sensitivity analysis, if the data are available. In that case, the presentation and justification of all methodological choices are indispensable (Rowen, Brazier et Van Hout 2015) (Brazier, Ratcliffe, et al. 2017) (Soekhai, de Bekker-Grob et Ellis 2019).

Consequences on the quality of life – measured through a specific instrument, without preference-based valuation – should be presented in a supplemental analysis.

Studies using disease-specific quality of life questionnaires[19] with no preference-based valuation method cannot be used in a cost-utility analysis*.

Nevertheless, patient-reported outcome measures* (PROM) – which provide data on how patients feel about their health state – constitute useful information for decision-making.

These studies can also be used to examine the utility scores estimated using the EQ-5D classification system and point out the most relevant sensitivity analyses.

Where relevant, these studies should be appended to the economic evaluation. The method used for the validation of the specific questionnaire should be described. It must have been the subject of at least one scientific publication.

2.3.2. Sources to document utility scores

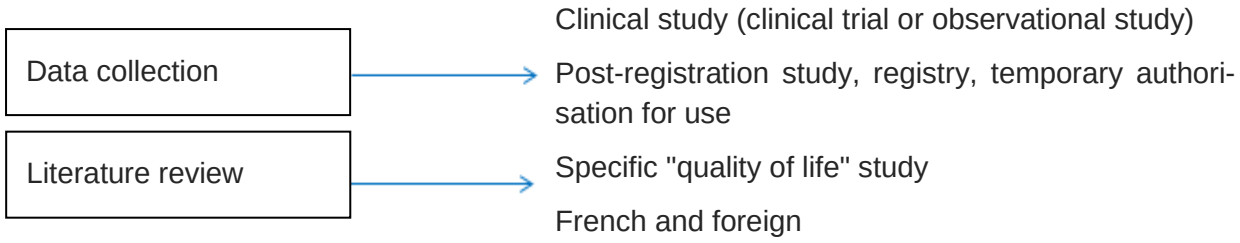
Guideline 16

The utility scores used to adjust life years should be derived from an *ad-hoc* study specifically designed for the collection of the required quality-of-life data or drawn from a systematic literature review. They should not be based on the opinions of experts.

For the collection and processing of quality-of-life data for the estimation of a utility score, the same methodological rigour should be applied as for the collection and processing of effectiveness and tolerability data.

[19] A quality of life questionnaire is said to be specific when it has been produced and validated for a specific disease (e.g. St Georges questionnaire for chronic respiratory diseases, KDQol questionnaire for terminal chronic renal insufficiency, QOLOD questionnaire for obesity and its treatments). The above-mentioned questionnaires have been translated into French and validated.

To document the utility scores used in the QALY calculation, the two main sources to be considered are(1) data from an *ad-hoc* study, or (2) a literature review.



Utility scores produced within the framework of an *ad-hoc* study

The population included in the study should be representative of the population analysed*. In the case of a multicentre study, the heterogeneity of samples between centres should be analysed (demographic variables, clinical risk factors, complication rates, etc.).

For the collection and processing of the data to be used to evaluate quality of life, the same methodological rigour should be applied as for the collection and processing of effectiveness and tolerability data. The study protocol, the collection of data during the study, as well as its statistical processing, should follow the current guidelines (Wolowacz, Briggs et Belozeroff 2016). In particular, the periodicity and duration of the data collection process should be based on the development of the disease and frequency of clinical events.

Special attention should be paid to the treatment of missing data. When the data to be used to evaluate quality of life are reported, the missing data should be described quantitatively and their random or non-random character should be discussed. The correction method used should be described.

When analysis is possible, the estimation of a minimum significant difference is recommended in order to improve the interpretation of the data collected (Walters et Brazier 2005) (Pickard, Neary et Cella 2007).

Extracting a utility scores from a literature review

Where relevant, utility scores are extracted from a systematic literature review conducted in accordance with methodological standards (Petrou, Kwon et Madan 2018). The method used for the identification and selection of the studies should be transparent (systematic literature review research protocol, analysis of the studies' methodological quality, and presentation of the flow chart).

The selected studies should apply the utility score estimation methods set out in this guidance document (see page 32, Utility score estimation method).

The use of a single study to document all utility scores is preferable. Where several publications are necessary, the studies' characteristics, their data collection methods and the characteristics of the populations covered should be clearly described, along with an analysis of their heterogeneity.

French studies should be preferred. In the absence of French publications, foreign studies may be used. In that case, the characteristics of the population in the selected study should be discussed in relation to the characteristics of the French population analysed* under real utilisation conditions (registers, epidemiological studies, etc.).

If several publications with a good methodological quality have been identified during the systematic review to document a utility score*, the utility scores not used in the base-case analysis* should be tested in a sensitivity analysis.

Experts' opinions

The use of experts' opinions to make up for the absence of published utility scores or quality-of-life data is not acceptable in a reference case analysis*, either in a base-case analysis* or a sensitivity analysis*. However, experts' opinions may be used to corroborate scores stemming from recognised methodologies.

2.3.3. Cases involving specific methodological difficulties

Guideline 17

From the age of 16 years, the EQ-5D classification system should be used. Before the age of 16 years, the use of a paediatric system is recommended. In the absence of a value set for the valuation of French preferences, foreign value sets may be used.

In cases of impaired cognitive states, the use of the EQ-5D questionnaire via a version filled in by a proxy of the patient should be the preferred approach.

Paediatrics

Several generic classification systems have been developed and validated for children and adolescents, but not with a French value set.

From the age of 16 years, the EQ-5D system should be used.

Before the age of 16 years, the use of paediatric systems is recommended. In the absence of a French preference value set, a foreign value set is acceptable.

Specific quality-of-life data may be collected and presented for information purposes, as it will make it possible to discuss the utility scores estimated using the selected generic system.

Table 1: Overview of the classification systems suited to a paediatric population

<table>
<tr><th>Instrument</th><th>Questionnaire</th><th>Age</th><th>Preference valuation</th></tr>
<tr><td>HUI2
HUI3</td><td>
<table>
<tr><td>Age of Subject</td><td colspan="2">Self-Administration (15Q)</td><td colspan="2">Interviewer Administration (40Q)</td></tr>
<tr><td></td><td>Self-Assessment</td><td>Proxy Assessment</td><td>Self-Assessment</td><td>Proxy Assessment</td></tr>
<tr><td>< 5 years</td><td>NR</td><td>NR</td><td>NR</td><td>NR</td></tr>
<tr><td>5–8 years</td><td>NR</td><td>Yes[1]</td><td>NR</td><td>Yes[1]</td></tr>
<tr><td>9–12 years</td><td>NR</td><td>Yes[2]</td><td>Yes[1]</td><td>Yes[2]</td></tr>
<tr><td>> 12 years</td><td>Yes[1]</td><td>Yes[2]</td><td>Yes[1]</td><td>Yes[2]</td></tr>
</table>
</td><td></td><td>HUI2 (CAN): SG with parents
HUI2 (UK): general population valuation
HUI3 (France): general population valuation</td></tr>
<tr><td>HSCS-PS</td><td>Adapted from HUI2 and HUI3</td><td>2.5 – 5</td><td>No published algorithm</td></tr>
<tr><td>CHU9D
(Stevens)</td><td>Self-assessment
Proxy</td><td>7-17</td><td>UK valuation: SG with general adult population
Australian valuation: discrete choice method with 11-17 year olds.</td></tr>
</table>

<table>
<tr><th>Instrument</th><th colspan="2">Questionnaire</th><th>Age</th><th>Preference valuation</th></tr>
<tr><td>EQ-5D-Y (Van Reene, 2014)</td><td colspan="2"><table>
<tr><td>0-7 years</td><td>For children aged 4 to 7 years, a proxy version can be used.</td></tr>
<tr><td>8-11 years</td><td>EQ-5D-Y, self-administered</td></tr>
<tr><td>12-15 years</td><td>The EQ-5D-Y questionnaire is recommended but, depending on the requirements of the study, the standard version can also be used.</td></tr>
</table></td><td></td><td>No value set available EuroQol does not recommend the use of the adult value set before the age of 16 years.</td></tr>
<tr><td>EQ-5D</td><td colspan="2"></td><td>From 16 years of age</td><td>French valuation: TTO and discrete choice with a general adult population</td></tr>
<tr><td>AQoL-6D Adolescent version</td><td colspan="2">Self-assessment</td><td>15-17 years</td><td>Valuation: TTO with students
Available for Australia, New Zealand, Fiji and Tonga</td></tr>
<tr><td>15D</td><td colspan="2" rowspan="2">Self-assessment
Interview
Proxy assessment</td><td>From 16 years of age</td><td>Valuation: Rating scale with a general Finnish population (not recommended)</td></tr>
<tr><td>16D</td><td>12-15 years</td><td>Valuation: Rating scale with Finnish children aged 12-15 years (not recommended)</td></tr>
<tr><td>17D</td><td colspan="2">Interview
(proxy for children under 8)</td><td>8-11 years</td><td>Valuation: Rating scale with parents of Finnish children aged 8-11 years (not recommended)</td></tr>
<tr><td>AHUM</td><td colspan="2">Self-assessment</td><td>Older children and adolescents</td><td>UK valuation: TTO with a general adult population</td></tr>
</table>

Impaired cognitive states

The EQ-5D classification system provides acceptable reliability in a certain number of common diseases such as depression, anxiety and personality disorders, but its use is more problematic in psychotic and severe and complex non-psychotic disorders (Brazier, Connell et Papaioannou 2014) (Mulhern 2014).

In such cases, EuroQol[20] recommends the use of a version of the EQ-5D questionnaire to be filled in by a patient proxy (someone close to the patient or a healthcare professional).

Rare events

Trials are generally ill-suited to estimating utility scores associated with a rare event (e.g. an adverse effect or acute event). Care should be taken to ensure that the conditions for the collection of the event data in the trial – in particular the periodicity of the event – are conducive to the production of reliable useful data.

[20] http://www.euroqol.org/eq-5d-products/eq-5d-5l/proxy-version.html

Where the occurrence of rare events has a major impact (different frequency between trial arms, high cost, impact on the quality of life), a pragmatic approach should be anticipated ahead of the trial[21].

In the absence of such a study, a systematic literature review should be conducted in order to identify the utility scores associated with each event, with a discussion of the differences observed between the trial and the literature. Extreme values should be tested in a multivariate sensitivity analysis.

The method consisting in estimating a mean utility difference between individuals having experienced the event, and those who have not, may be used if the data collected in the trial are not solid enough to estimate the disutility associated with the event (very rare event, unsuitable data collection timeline, etc.) or if the disutilities derived from the literature are not reliable (multiple sources, non-comparable populations, etc.). The limitations of the approach should be discussed and the uncertainty explored.

[21] The collection of rare event data requires the adoption of specific protocols (e.g. specific study targeting patients with risks of complications, shorter collection periodicity for quality-of-life data in the trial, collection of quality-of-life data for all clinical contacts not covered by the protocol, data collection periodicity sampling).

3. Methodological choices for the evaluation of costs

3.1. General principles

Guideline 18

The evaluation of the total cost of an intervention is based on the intervention's production costs, which implies the identification, measurement and valuation of the resources consumed.

The scope of the cost items evaluated depends on the perspective adopted.

- Under a collective perspective, all resources consumed in the production of the overall patient care are taken into consideration. They cover the domestic sphere (e.g. informal care), the healthcare sphere (e.g. stays, procedures, healthcare products) and the medico-social sphere (e.g. stays and personal care services).
- Under the healthcare system perspective, the resources considered are restricted to those involved in the production of the patient care (stays, procedures, and healthcare products).

Only direct costs should be considered in the reference case analysis.

A direct cost analysis may be presented as a supplemental analysis.

The evaluation of the total cost of an intervention is based on the analysis of the intervention's production costs, which involves the identification, measurement and valuation of the resources consumed.

HAS distinguishes between the cost of the resources required for the production of the interventions evaluated (direct costs*), and the cost of the resources rendered unavailable because of a poor health state or death (indirect costs*).

In the reference case analysis* – whether in a base-case analysis* or sensitivity analysis* – only direct costs are considered.

The scope of the resources to be considered depends on the perspective of the evaluation.

Under a collective perspective, all resources consumed in the production of the overall patient care should be identified, irrespective of their nature (i.e. domestic, healthcare or medico-social spheres) or their funders (i.e. users, mandatory and supplementary health insurance, social aid, state funding, hospitals).

Under the healthcare system perspective, only healthcare system resources are identified.

Direct costs

To estimate the direct costs* associated with an intervention, the resources to be considered include the following:

- the resources consumed for the intervention (e.g. acquisition costs, administrative costs, as well as the treatment of any adverse events linked with the intervention);
- the resources consumed for the care, and which may change following the intervention (e.g. follow-up care, care linked to comorbidities, care provided by informal caregivers, concomitant treatments, end-of-life care).

Depending on the chosen perspective (see Annex 7, page 75), they may consist in the consumption of:

- hospital care, outpatient care, medical goods, transport, etc.;
- resources consumed due to changes in the organisation of the care[22];
- the expenditures required for the transition from the prevailing situation to that in which the intervention becomes routine[23];
- the other resources consumed in the medico-social and domestic spheres (see collective perspective).

Indirect costs

While indirect costs are excluded from the reference case analysis*, the presentation of a supplemental analysis* including indirect costs* may be relevant.

Indirect costs* concern resources not used in patient care, but which have been rendered unavailable due to the patient's poor health state or premature death. The losses of resources considered are generally losses of productivity, due to a total stoppage (absenteeism) or partial reduction of the productive activity of the population analysed*, whether this activity is paid or unpaid.

3.2. Identification, measurement and valuation of direct costs

Guideline 19

The evaluation of costs involves three stages: the identification, measurement and valuation of the resources associated with the intervention.

All resources associated with the intervention should be identified over the relevant time horizon. The difficulty of taking a resource into account should be duly justified.

The amounts of resources consumed should be measured using high-quality data stemming from appropriate methodology, along with clearly referenced and validated sources.

The valuation of the resources should be based on their unit production cost in France. Where valuation by cost of production is not possible, the French tariffs may be used.

The evaluation of the costs of an intervention is calculated by applying a unit cost (valuation phase) to the amount consumed, expressed in physical units (measurement phase) of the resources identified as directly associated with the intervention or the portions of the care likely to be altered by the intervention (identification phase).

[22] When part of the costs is tied to a particular organisational set-up, any predicable organisational change should be taken into account in the evaluation, over the entire time horizon. This could, for instance, consist of a change in a surgical technique (duration of the intervention and post-operative care), the mode of administration of a treatment (oral/intravenous, frequency, transfers), the organisation of the care over the duration of the hospital stay, the outpatient care, etc.

[23] This may consist of the costs linked to the initial training of staff for the introduction of an intervention, the co-existence of several interventions during the build-up of the new intervention, etc.

3.2.1. Identification of resources consumed

All of the resources coming under the scope of the direct cost evaluation should be identified in an exhaustive way, over the entire time horizon* chosen (see Annex 7, page 75). Only the resources coming under the scope of the chosen perspective will be considered.

Future costs not directly linked to the interventions being evaluated should not be considered (e.g. ageing costs, costs associated with the onset of other pathologies).

The identification of resources should be based on relevant and appropriate sources of information: databases, guidelines for clinical practices, data in published literature, clinical or pilot studies, experts' opinions (see Annex 13, page 100).

The sources of information should be mentioned, and their choice should be clearly explained when several sources are available. In particular, the impossibility of including a relevant resource in the cost calculation should be duly justified.

3.2.2. Measurement of resources consumed

The choice of the data sources used, out of the data sources available to the evaluator, should be justified.

Resources linked to the acquisition or production of the interventions

For resources linked to the acquisition or production of the interventions being evaluated, the data stemming from trials underpinning the evaluation of their effectiveness and tolerability should be preferred, in order to ensure consistency between the evaluation of effectiveness and costs. Data stemming from different sources may be tested in a sensitivity analysis (e.g. real-life data on treatment durations or compliance).

Other resources

For other resources (e.g. linked to disease follow-up or end-of-life care), the measure of the quantities consumed should be based on high-quality sources, giving priority to French sources using real-life conditions in the field or pathology studied: observational prospective study, retrospective study of patient files, registry data, good practice guidelines, or French databases[24]. Where relevant, data collected within the scope of a randomised controlled trial or experts' opinions may be used.

- When the data are collected in a randomised controlled trial, the available data rarely cover the full range of resources consumed. The data collection method should be presented in a transparent way and the validity of the data collected should be discussed. Any measures taken to make up for the data not collected in the trial should be described. If the randomised controlled trial is conducted in several countries, special attention should be paid to the statistical analyses used for the aggregation of the data. Moreover, the conditions under which the data are used in the French context should be discussed.
- The data obtained through experts' opinions should be clearly identified and the methodology adopted should be presented in a transparent way: criteria for the selection of the experts, number of experts approached and number of experts having responded, presentation of the experts (name, specialisation and disclosure of potential interests[25]), method used to record the

[24] Numerous databases exist and can provide useful information to quantify the resources consumed: databases of the National Health Insurance and PMSI (Medical IT system), data relative to the market shares of healthcare products, medical registries, etc. Some allow unrestricted access, free of charge, while others only allow restricted or fee-paying access.

[25] In particular, it should be specified whether the expert has ties with a stakeholder in the evaluation beyond this specific opinion.

opinions, and presentation of the questionnaire. For quantitative parameters, a formal method of elicitation is preferred (see Annex 13, page 100).

Micro-costing approach

The collection of field data using the micro-costing method may prove necessary to identify and measure the resources consumed (Guerre, Hayes, & Bertaux, 2018) (see Annex 9, page 89) when:

- the intervention evaluated is new and the related cost items have not yet been identified and measured (e.g. new treatment, procedure not listed in the CCAM, medical device not included in the list of products and services qualifying for reimbursement or in a Healthcare Resource Group (HRG));
- the intervention evaluated significantly alters the existing care and the associated costs have not been identified and measured (e.g. computer-aided surgery, robotic surgery);
- the intervention evaluated entails a high variability of care among individuals or care institutions;
- the intervention includes non-market goods for which there are no standardised costs (health programmes, therapeutic education, new training course to be set up, etc.).

Furthermore, in the absence of data and as a last resort, the measurement of a resource (e.g. the amount of use of remaining resources) may be based on the formulation of assumptions. In all such cases, the assumptions should be clearly set out, justified and tested in a sensitivity analysis.

3.2.3. Valuation of resources consumed

The valuation of the resources consumed should be based on unit production costs.

In the absence of unit production cost data, or if the data are deemed inappropriate, medical tariffs may be used as a valuation basis as they represent community-recognised prices. Any surcharges payable by patients are to be added to the tariffs.

When data are available, the difference between the tariff, or face price, and the actual price should be documented and undergo a sensitivity analysis.

Resources for which there is no set tariff (non-listed procedures, non-refundable medical devices or medicines, heavy equipment, mobile medical units, etc.) should be valued at the average real price if observable, or via another method to be detailed.

Each resource consumed should be valued for France, in euros, for a year of reference. The preferred method to express costs in that year of reference should be based on:

- the consumer price indices of the healthcare products and services, available on the INSEE website[26], to convert the costs in the year of reference;
- the Purchasing Power Parity method, or any other duly justified method, to transpose costs established in a foreign country to France.

The following sections describe available valuation methods (see Annex 8, page 76).

Hospital costs

Hospital costs are valued in close keeping with hospital stay production costs.

Valuation of Homogeneous Patient Groups (HPGs) via the National Cost Study (NCS)

To date, the National Cost Study (NCS) produces the valuation, which is closest to the hospital production cost, since it is based on the cost-accounting of a sample of public and private institutions.

[26] https://www.insee.fr/fr/statistiques/series/102342213?MENAGES_IPC=2330043&PRIX_CONSO=2409126

First, the hospital stays of interest are identified and classified per Homogeneous Patient Group (HPG), using the Medical IT system (PMSI) in the relevant field (MSO, HAH or FCR) [27].

The heterogeneity of stays and institutions observed in the PMSI should be considered.

- When the stay of interest is likely to be linked to several HPGs, the breakdown of the underlying reason (principal/associated procedure or diagnosis) between the different HPGs should be considered.
- When the stay of interest takes place in both public and private institutions, the breakdown of the activities between the two sectors should be considered.

Second, the identified stays of interest are to be valued using the NCS, something which requires certain corrections.

- The cost of the stay to be considered should exclude structural costs.
- The estimation of the average cost of stay should exclude the cost of products funded on top of HRG-based tariffs. The cost of these products should be added in the arms of the interventions concerned by applying the requested price or the price stated in the Official Gazette. For products whose price is listed in the Official Gazette, a sensitivity analysis should be conducted on the basis of the actual average price recorded in the PMSI.
- The fee surcharges of private clinics should be added[28].

When the NCS data do not reflect the characteristics of a stay associated with an intervention being evaluated, any modification made to the average cost stemming from the NCS must be clearly explained and justified.

Valuation of Healthcare Resource Groups (HRGs) via their medical fees

Since the NCS is a cost study involving a voluntary sample of institutions and, consequently, stays, there are inherent limitations in its construction[29]. NCS-based valuation is not recommended when the uncertainty surrounding the average cost estimation is too great, i.e. when:

- the survey rate[30] is below 20%;
- the sampling error (SE) [31] exceeds 30%;
- the sampling error, and thus the confidence interval, cannot be calculated, in particular due to the fact that the estimation is based on a sample of less than 30 stays or stays produced by a single institution;
- a change in classifications impacting the interventions studied has not been considered in the NCS, due to the NCS time lag.

In such situations, except in duly justified cases, the hospital stays are valued on the basis of applicable tariffs per HRG[32] (Healthcare Resource Group) or specific flat-rate fees (organ removal, home dialysis

[27] MSO: Medicine, Surgery, Obstetrics and Odontology / HAH: Hospitalisation at Home / FCR: Follow-up Care and Rehabilitation. There is no NCS in the field of psychiatry.

[28] Data available on ATIH's ScanSanté website: fee surcharges of private clinics per CCAM-listed procedure or category of procedures.

[29] NCS-based estimations of production costs represent average costs and may conceal sharp variations among institutions or even among patients, as they partly rest on accounting policies that may vary from one institution to another.

[30] The survey rate is the percentage of representation of the NCS sample in relation to national activity.

[31] The sampling error (SE) is defined as the relative error made by the average cost assessor per Homogeneous Patient Group (HPG) due to institution sampling variations; it determines the limits of the confidence interval.

[32] HRG-based tariffs published annually in the Official Gazette.

and self-dialysis[33]). All resources not included in the hospital stay tariff should be valued in addition to that tariff:

- the prices of medicines and medical devices funded separately from HRG-based fees[34];
- the tariffs applicable to daily supplements (resuscitation, intensive care, neonatal care, continuous patient surveillance, peritoneal dialysis, radiotherapy, antepartum);
- the portion of the fees paid by the patients or supplementary health insurance (fee surcharge, inpatient charges[35], etc.);
- medical fees in the case of hospital stays or procedures carried out in for-profit private institutions.

Micro-costing approach

If the micro-costing method is used (Guerre, Hayes, & Bertaux, 2018), each resource consumed must be valued using a real unit cost, which may be determined through a bottom-up or top-down approach (see Annex 9, page 89). The unit cost lists (RTC) available on ATIH's ScanSanté website may be used as a reference for the valuation of certain activities[36].

Outpatient costs

Medical and paramedical consultations [37] (including external consultations) are valued on the basis of total applicable fees (including the consultation fee, travel expenses and surcharges) and the specificity of the independent healthcare professionals' activities[38]. The consultation valuation should take account of the healthcare professionals' specificity and give preference to the data provided for approved healthcare professionals.

Medical laboratory procedures are valued on the basis of the numbers and amounts corresponding to the reimbursement bases available on the National Health Insurance site[39].

Other clinical and technical medical procedures are valued according to the relevant classification (CCAM, NGAP, TNB) and the associated tariffs, to which are added any fee surcharges, supplementary compensation and flat-rate technical fees.

Medical devices and medicines are most often valued on the basis of their tariff, except when this tariff does not represent all of the expenses borne by the various funders:

- medical devices for which a reference price (tarif forfaitaire de responsabilité – TFR) has been set may be marketed at a price to be freely determined by the manufacturer. Such devices are

[33] Set fees for haemodialysis in a medicalised dialysis unit, simple self-dialysis, assisted self-dialysis, home-based haemodialysis, automated peritoneal dialysis, continuous ambulatory peritoneal dialysis, training in automated/continuous ambulatory peritoneal dialysis, training in haemodialysis in a medicalised dialysis unit.

[34] When a device or medicinal product is funded on top of the HRG-based tariffs, it is valued on the basis of the requested price or the price listed in the Official Gazette. For products whose price is listed in the Official Gazette, a sensitivity analysis should be conducted on the basis of the actual average price recorded in the PMSI. Medicinal products under a temporary usage authorisation are to be valued on the basis of the maximum indemnity; drops of 10 to 20% should be tested in a sensitivity analysis.

[35] Inpatient charges represent the patients' share of accommodation and cleaning costs stemming from their hospitalisation. These charges are payable for each day of hospitalisation, including the day of their discharge. They currently stand at €18 per day in a hospital or clinic and €13.50 per day in the psychiatric ward of a health institution. They may be applied over an average length of stay. There are certain situations in which the charges may be waived (https://www.ameli.fr/assure/remboursements/reste-charge/forfait-hospitalier).

[36] The latest data dates back to 2013.

[37] Consultations of midwives and auxiliary medical personnel (nurses, physiotherapists, chiropodists, speech therapists, orthopaedists).

[38] https://www.ameli.fr/l-assurance-maladie/statistiques-et-publications/donnees-statistiques/professionnels-de-sante-liberaux/index.php

[39] https://www.ameli.fr/l-assurance-maladie/statistiques-et-publications/donnees-statistiques/actes-de-biologie-medicale/index.php

valued at the volume-weighted average price (inclusive of VAT) calculated on the real quantities sold;

- non-refundable medicines and devices sold at a price exceeding the applicable tariff are valued at the price actually paid.

Dispensing fees are to be added under the conditions defined by the regulations applicable on the date of submittal of the evaluation[40].

The valuation of the average cost of acquisition of a class of medicinal products[41] takes account of the market share and prices of the medicinal products in the class. The method used to take account of dosage heterogeneity should be detailed.

The valuation of transport costs is based on the average reimbursed amounts published (see Annex 8, page 76). Non-reimbursed transport costs are estimated on the same basis as reimbursed transport costs.

If one of these elements is not considered, its justification is required.

Specific cases

Difficult-to-assess resources

Certain resources, for which there is no set tariff, are difficult to value (non-reimbursed transport, organising of a health programme, caregivers' time, etc.).

Their evaluation is nevertheless sought.

Data sources which allow the valuation of direct costs beyond hospital and outpatient care are often heterogeneous. The options available should be identified and the chosen option should be justified. The options not used should be tested in a sensitivity analysis.

Anticipated changes in costs over time

If the price of a technology or the cost of its implementation are likely to decrease due to the wider distribution of the equipment or technical skills, this anticipated drop 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.

Calculation of cost per event avoided

In the case of an analysis per event avoided, the inclusion of the costs of the event avoided in the total cost estimation is debatable. On the one hand, it potentially generates double accounting by taking into account the benefit of the intervention (the event avoided) in the health outcome, and by giving it a monetary value in the costs (Mullins, 2006). On the other hand, excluding such costs would mean not considering all of the consequences of the intervention.

The impact on the results of the evaluation may be considerable, for instance if the intervention makes it possible to avoid a serious life-long disability.

[40] Several fees may be added to the price of the medicinal product, per box, per quarterly box, per prescription of refundable medicinal product, per complex prescription, per prescription for young or elderly patients, per prescription comprising one or more specific medicinal products. These fees are subject to an agreement between the USPO (French Union of Retail Pharmacists) and the National Health Insurance and are published in the Official Gazette.

[41] For example, if the assessment requires the valuation of the cost of treatment with antihypertensive medicines as one of the components of an intervention.

Il is recommended not to include the costs of the event avoided in the assessment of the total cost. The result of the evaluation is then interpreted as the amount required to avoid the event and may be compared with the cost of treating the event, in a cost minimisation analysis.

3.3. Identification, measurement and valuation of indirect costs in a supplemental analysis

Guideline 20

The impact of an intervention on the time dedicated to a professional or leisure activity should be measured in terms of specific lost time by category of activities affected.

The method used to appraise this lost time is to be chosen by the evaluator and justified.

The resources rendered unavailable due to the patient's poor health state or premature death generally concern the time dedicated to a productive activity, whether this activity is paid or unpaid. The impact of the interventions on the activities of the persons undergoing them, and possibly those of the persons closest to them, should be measured in terms of duration of the different categories of activities affected.

The choice of the valuation method (e.g. in terms of human capital, or friction costs for productivity-based valuation) is left up to the author of the analysis but will need to be justified.

For this supplemental analysis*, the same methodological rigour should be applied as for the calculation of direct costs in the reference case analysis*. The arguments supporting the interest of the information provided by this analysis and the method used should be clearly set out.

4. Methodological choices concerning modelling

4.1. General principle

Guideline 21

The development of a model should comply with three requirements: justification, validation and exploration of uncertainty.

If there is no solid argument to clearly justify a particular methodological choice over several credible options, the one which is the least favourable to the intervention evaluated, in terms of cost or health outcome differential, should be opted for.

Economic evaluation often relies on modelling techniques in order to meet certain methodological requirements such as the inclusion of all available scientific data in the analysis, the comparison of all relevant options, data extrapolation over time, and the exploration of uncertainty (Briggs, Claxton et Sculpher 2008). In this respect, a model is a tool for the analysis of available data. Modelling cannot be a substitute for the collection of high-quality data, irrespective of the level of technical and computational sophistication developed.

The modelling process is based on a succession of methodological choices, such as the type and structure of the model, the simulated population, the comparators, the duration of the simulation*, the assumptions made on the development of the disease over time or the treatment effect, the selection of data sources, etc.

These choices should be consistent with the evaluation's structural choices*.

Table 2: examples of links between the structural choice of the evaluation and the choice of a model

Structural choice of the evaluation	Choice of the model
Perspective	The perspective determines: – the scope of the health effects and resources to be considered in the model; – the method of valuation of health effects and resources.
Time horizon	The time horizon determines the duration of simulation of the model. – Until the extinction of the cohort in the case of a lifetime horizon*. – To a set number of years or maximum age of the cohort in the case of a specific time horizon.
Population analysed Subgroups analysed	The population simulated in the model is representative of the population analysed.
Comparators	The comparators used in the model are selected on the basis of the comparators identified.

All modelling choices should be clearly explained and justified. If there is no solid argument to clearly justify a particular methodological choice over several credible possibilities, the one which is the least favourable to the intervention evaluated in terms of cost or health outcome differential should be opted for.

The conservative character of the choice made should be justified with regard to its impact on the results of the evaluation[42].

All choices have an impact on the model's results and can generate uncertainty. Consequently:

- model choices and simulations should be validated (see page 53, Model verification and validation);
- the uncertainty should be documented and quantified in order to be taken into account in decision-making (see page 54, Exploration of uncertainty).

4.2. Model type and structure

Guideline 22

The type of model and its structure should be defined so as to represent the patients' clinical progression and care, without introducing any needless complexity.

The technical characteristics of the model should be suited to the specificities of the evaluation (mode of progression over time, degree of heterogeneity and interaction among individuals, and degree of randomness) and comply with applicable recommendations.

The structure of the model (statuses, events, links) should be defined so as to capture the costs and health outcomes associated with the progression of the disease and comparative care.

The selected model type should be justified.

The choice of the model used should be justified through an analysis of possible options (see Annex 10, page 92). The most appropriate model type should be chosen, mainly on the basis of the following:

- the way in which the model integrates time (duration of the simulation*, possibility of recurrence, discrete or continuous mode of progression over time);
- the existence (or not) of interactions among individuals;
- the most appropriate statistical unit for groups of individuals with homogeneous characteristics (cohort model) and individuals with heterogenous individual characteristics (individual-centred model);
- the model's ability to take parameter randomness into account (deterministic* model versus stochastic* model).

Choices concerning the structure of the model should be clearly explained.

The model structure is defined through statuses and events, and how they are linked. It should be consistent with the clinical progression of patients and their care. Several elements should be described and justified, including:

[42] For example, when choosing the parameter function to be used to adjust survival data, the conservative character of the choice should be discussed in view of the survival differences that determine the evaluation of efficiency, rather than the survival criterion for each of the interventions studied.

- the events or statuses included in the model, in connection with the history of the disease or the consequences of the interventions being evaluated;
- the chronology of statuses or events.

The model does not have to reflect all of the aspects of the pathology or clinical progression, but it should be detailed enough to account for factors likely to result in differences in costs or health outcomes between the interventions compared. For example, in a transition model, special attention should be focused on the way in which the model structure integrates intercurrent events* (e.g. adverse events and treatment stoppages).

Data availability should be considered in the choice of the model. For example, in a partitioned survival model (or area-under-the-curve model), the maturity of the overall survival data is a determining element to be discussed.

When there is uncertainty as to the structure to be adopted, the choice of the structure used should be justified and its consequences on the conclusions of the evaluation should be discussed.

4.3. Values and assumptions associated with model parameters

Guideline 23

Observed parameter values and distributions should be used to document the model.

When there is no observed value for a parameter, it should be estimated through an *ad hoc* calculation or calibration method. The estimation method used should be detailed and justified.

The uncertainty associated with the estimation of the value of the model parameters should be explored and quantified.

When an extrapolation technique is used, all assumptions should be detailed and justified, in particular those relating to the treatment duration and effect size.

The observed values of model parameters should be sought as a priority.

The values associated with health outcome parameters (effectiveness, tolerability, quality of life) and cost parameters should stem from a systematic and exhaustive research process that can cover numerous sources of data: systematic literature review, meta-analyses, use of medical or administrative databases, clinical trials, observational epidemiological studies (cohort studies, case-control studies, cross-sectional studies), surveys, registries, etc.

When there are several sources for a particular parameter, a statistical summary of the data through a pairwise meta-analysis* should be opted for (see page 27. Quantification of comparative effectiveness and safety). If a particular source is selected, its choice should be justified, and the parameter should be the subject of a sensitivity analysis including the other plausible data sources not used.

Each parameter's statistical distribution (central and dispersion characteristics) and data source should be documented. When the parameter's distribution is not precisely known, the elements likely to document it should be presented. This information should be used in the sensitivity analyses to be conducted.

Non-observed values of model parameters should be estimated through an *ad hoc* calculation or based on assumptions.

Where the value of a parameter is not an observed value, the existing knowledge should be examined, explicitly distinguishing between what is controversial, what is poorly known, and what is not documented at all.

In numerous cases, it is possible to estimate the value of a parameter and its distribution through an *ad hoc* calculation, independently from the actual modelling process[43]. The plausibility of the estimations should be discussed. The calculation should be validated through external data (see page 53, Model verification and validation) and the impact of its integration in the model should be examined through an exploration of uncertainty (see page 54, Exploration of uncertainty).

Where the data are insufficient, the value of a parameter may be estimated using a calibration* method, so as to adjust the parameter value in line with the model simulations using external data not used to configure the model.

In the absence of elements making it possible to estimate the value of the parameter and its distribution using a robust methodology, it may be necessary to resort to assumptions or experts' opinions (see Annex 13, page 100). The value and distribution of the parameter should be justified, and the associated uncertainty should be explored and quantified in a sensitivity analysis.

Extrapolation is required when the time horizon of the analysis is longer than the period of observation of available data.

When effectiveness/tolerability/utility/cost data are not available for the entire time horizon* used, an extrapolation method is required. The data extrapolation can have a major impact on the results of the evaluation.

The following should be explained and justified in a rigorous and exhaustive way:

- the point(s) in time from which the model integrates extrapolated data instead of observed data;
- the assumptions used (e.g. growth, decrease or stability over time) to extrapolate effectiveness/tolerability data, utility scores and the variables required for the estimation of costs. Special attention should be paid to the extrapolation assumptions – relating to the treatment duration and effect size – used after the observation period and after the stoppage of treatment;
- the sensitivity analyses conducted to test the impact of these assumptions.

The extrapolation methods should follow the approaches recommended in published literature. For example, methods for the extrapolation of survival data from parametric models have been described by Latimer (2011) (2013), Collett (2015) and Woods, Sideris, & Palmer (2017). The algorithm developed by Latimer in 2011 is a useful tool to select the parametric models to be used in a base-case analysis* and the models to be tested in a sensitivity analysis (see Annex 11, page 94).

The proportion of risks between the intervention evaluated and its comparators should systematically be discussed.

In the case of high uncertainty due to the lack of medium/long-term data, the base-case analysis* should adopt the conservative assumption, out of the plausible assumptions, for the extrapolation of the relative treatment effect. The approach should be discussed in detail and tested in a sensitivity analysis*.

[43] Examples of *ad hoc* calculations: estimation of an initial cardio-vascular risk using a risk equation; adjustment of available data to the characteristics of the population analysed; estimation of the occurrence of an event and comorbidity; projection of survival curves using censured data.

4.4. Model verification and validation

Guideline 24

The model's ability to produce consistent and credible simulations should systematically be explored through the technical verification of the model and its internal validation, and through a simulation validation procedure in order to ensure that:

- the model and the simulated results are intuitively consistent (face validity);
- the simulated results are consistent with external data not used for the configuration of the model (external or predictive validity);
- the model generates results which are comparable to those of other models whose validity has been recognised (cross-validity).

All significant differences and all inconsistencies should be examined, and their origin should be sought.

Each model should undergo a technical verification and a validation procedure aimed at ensuring the model's ability to reproduce consistent and plausible results in keeping with the history of the disease and the expected effects of the interventions studied (Eddy, Hollingworth, & Caro, 2012).

The use of a standardised grid to present the different analyses conducted is recommended (see Annex 12, page 95).

Technical verification and rigorous internal validation are required.

The model should undergo a formal quality check using proven methodologies (e.g. double programming, test of the mathematical logic of connections between the model's parameters and results, reproducibility test, extreme values test) in order to identify any programming errors, data integration errors or inconsistencies in the model's mathematical logic. This technical verification should be carried out by a third party who has not taken part in the development of the model.

The internal validation of the model should end with the comparison of the simulations and data introduced.

Any counter-intuitive results should be explored and discussed.

The consistency and plausibility of the simulations generated should be discussed.

The model's ability to produce consistent, plausible results should be established along three lines.

- Face validity*: Does the model's structure, along with its assumptions, the values and distributions of integrated parameters, and its simulated results and conclusions, make sense and are they consistent with the intuition? By nature, the evaluation of face validity is qualitative; it takes place at a very early stage of the development of the model and in an iterative manner.
- External validity*: Are the model simulations at different points of the simulation (intermediate results and final results) consistent with other data (e.g. national statistics, epidemiological data, registries, cohorts, other observational databases, relevant clinical trials)? External validation applies to all interventions simulated by the model. The comparison of the model's simulations with data stemming from clinical trials or real-life conditions is an essential stage of the model validation process. The data used for external validation is not the same as that used to configure the model and should stem from a population which is sufficiently comparable to the

simulated population*. The choice of using the available data, either to calibrate the model or to validate the model simulations, should be examined on a case-by-case basis.

- Cross-validity*: Does the model generate results in line with those of other validated models – whether French or foreign – developed independently for the same interventions? Consistencies and differences between the models should be clearly identified and explained.

All significant differences and inconsistencies should be examined, and their origin should be sought. Where relevant and feasible, the required adjustments should be made to the model.

4.5. Exploration of uncertainty

Guideline 25

A systematic exploration of the sources of uncertainty associated with the evaluation's structural choices, the modelling choices and the model parameter estimations should be presented according to an appropriate methodology.

- Sensitivity analyses should quantify the impact of a different structural choice in the reference case analysis (e.g. perspective, time horizon, population analysed, comparators, discount rate).
- Sensitivity analyses should quantify the impact of methodological choices and modelling assumptions (e.g. model structure, data sources, calculation methods or assumptions to estimate the value of parameters not directly observed). The impact of the assumptions used for the extrapolation of treatment effects should be systematically explored.
- The uncertainty associated with the parameters of the model should be systematically explored using two complementary approaches: a probabilistic sensitivity analysis, based on a second-order Monte Carlo simulation, and deterministic sensitivity analyses identifying the parameters (or combinations of parameters) which have the greatest influence on the results of the evaluation.

For all of the sensitivity analyses presented, the credibility of the options tested should be justified, along with an interpretation of their results and their contribution to the comprehension of the evaluation.

When a scenario which is fundamentally different to that used in the reference case analysis is put forward, the presentation of its results should include a thorough exploration of the uncertainty through deterministic and probabilistic sensitivity analyses.

A model is always an imperfect representation of reality (Ghabri, Cleemput et Josselin 2017). Its relevance rests on the strength of the arguments presented to justify it and on a stringent validation procedure. In addition, the uncertainty associated with the evaluation's structural choices*, as well as the uncertainty associated with the modelling choices and the uncertainty associated with the estimations of the model parameters, should be systematically explored (Briggs, et al. 2012) (Ghabri, Hamers et Josselin 2016).

The uncertainty exploration principle rests on the analysis of the variations caused by a credible modification of the assumption or parameter studied. The results of the sensitivity analyses, along with their contribution to the comprehension of the evaluation, should be interpreted in a rigorous way.

4.5.1. Exploring uncertainty in the reference case analysis

Uncertainty associated with the evaluation's structural choices

Sensitivity analyses should be conducted to quantify the impact of the structural choices made for the reference case analysis* with respect to the results of the evaluation. The structural choices* tested should be explained on a case-by-case basis.

Uncertainty associated with modelling choices

Sensitivity analyses should be conducted to quantify the impact of the modelling choices on the results of the evaluation and, in particular:

- the model type and structure (e.g. treatment stoppage rules, treatment sequences after stoppage of the treatment under study);
- the sources of data used to populate the model, when several sources exist for the key parameters of the analysis (e.g. effectiveness of interventions, utility scores, resources consumed);
- the calculation methods and assumptions used to estimate the value of parameters not directly observed, when credible alternatives exist (Briggs, et al. 2012);
- the methodology used for the extrapolation of long-term data. These choices should be tested in a systematic, rigorous way. In particular, the assumptions used to extrapolate a relative treatment effect after the observation period and after stoppage of the treatment are crucial in the evaluation of health outcomes over the time horizon. As a minimum, the analyses should make it possible to estimate health outcomes under the following three assumptions: a relative treatment effect which is nil after the observation period, a relative treatment effect which is nil at the end of the treatment period, a relative treatment effect diminishing over time.

Uncertainty associated with the estimation of the model parameters

The uncertainty associated with the estimation of the parameters should be explored through a probabilistic sensitivity analysis.

A probabilistic sensitivity analysis based on a second-order Monte Carlo simulation is systematically required when the parameters' theoretical or empirical distributions are known or can be estimated. Depending on the number of comparators, an acceptability curve* (two options compared) or multi-option acceptability curve* (three or more options) should be presented.

Only parameters whose value is a random variable should be integrated in a probabilistic sensitivity analysis*[44].

The choice of distributions, the values of their parameters, as well as the number of Monte Carlo iterations, should be specified and justified.

Where the modelling methodology greatly hinders the use of a probabilistic analysis based on a second-order Monte Carlo simulation, this point should be clearly explained, and an appropriate methodology should be put forward.

The uncertainty associated with the estimation of the model's parameters should be explored through deterministic analyses making it possible to identify the most influential parameters.

Univariate deterministic sensitivity analyses (or multivariate ones if necessary) should systematically be conducted to identify the parameters that can have the greatest impact on the results of the evaluation.

[44] Discount rates and prices are examples of non-random variables.

The choice of the parameters involved in a deterministic sensitivity analysis* should be justified. They may consist of parameters with a wide variation amplitude, parameters stemming from low-evidence studies, parameters relating to behaviours for which actions can be taken, etc. When a multivariate deterministic analysis is conducted, the author should explain the reasons behind the selection of the parameters.

Since the price (or cost) of the intervention is a parameter of interest, a sensitivity analysis should systematically be presented.

For the cases requiring a CEESP opinion on healthcare products, a deterministic analysis should be conducted on the price of the intervention being evaluated by testing at least three prices below the price used in the base-case analysis. The results presented for each price level should be backed by a probabilistic sensitivity analysis*. The curve linking those price levels and the associated ICERs should be represented on a graph, and the equation linking the price and the ICER should be presented[45].

For the other parameters, the choice of the values tested in a sensitivity analysis* should be presented and justified (e.g. 95% confidence interval, arbitrary variation ±x%).

Where appropriate, the procedure should be supplemented with a threshold analysis – i.e. an analysis to determine the values that alter the conclusions of the evaluation.

4.5.2. Exploring uncertainty in a scenario analysis

In certain cases, it may be pertinent to present the results of a scenario that fundamentally differs from the scenario used in the reference case analysis, by adopting different structural choices, modelling assumptions or parameters.

In that case, the presentation of this scenario's results should be associated with a thorough exploration of uncertainty through deterministic and probabilistic sensitivity analyses.

[45] Using the different ICER values estimated in the price sensitivity analyses, it is possible to estimate a relationship between the ICER and the price of the intervention:
ICER = a x [requested price] + b.
This relationship is to be interpreted with all else being equal and can only be estimated when the price variations do not alter the cost-effectiveness frontier.

5. Presentation and interpretation of results

5.1. Presenting and interpreting results to produce useful conclusions for decision-making

Guideline 26

Quantitative results should be presented and interpreted in a way which is consistent with the objective of the economic evaluation.

The evaluation of cost-effectiveness requires the identification of the interventions on the cost-effectiveness frontier and the presentation of the results based on the incremental cost-effectiveness ratio (ICER) or net benefit (NB).

All of the relevant economic information to aid public decision-making should be extracted from the evaluation.

A clear, justified discussion should make it possible to estimate the robustness of the conclusion of the evaluation and define the conditions under which the conclusion would be altered. This discussion should rest on a critical appraisal of the methods and data used, and on the sensitivity analyses conducted.

The degree of confidence associated with the results should be detailed.

5.1.1. The evaluation of cost-effectiveness involves identifying the interventions on the cost-effectiveness frontier.

For the evaluation of cost-effectiveness, the interventions on the cost-effectiveness frontier* need to be identified. The cost-effectiveness frontier represents all situations for which there are no other interventions that provide a better (or identical) health outcome at a lower cost (non-dominated interventions or interventions maximising the net benefit).

An intervention's position on the cost-effectiveness frontier* makes it possible to conclude that it constitutes an economically rational choice, but does not provide any indication of its acceptability in terms of what the community is willing to pay for an additional health unit (Raimond, Midy et Thébaut 2016).

If it is not possible to build the cost-effectiveness frontier using all of the relevant interventions, then it is not possible to evaluate cost-effectiveness. In such cases, a side-by-side comparison of the evaluated intervention with comparator interventions will provide useful economic information in terms of differences in health outcomes and costs. This information should be presented and discussed.

5.1.2. Analysis of the relationship between additional cost and health gains via one of the two recommended metrics

The cost-outcome analysis* may use two metrics: the incremental cost-effectiveness ratio (ICER) or net benefit (NB) expressed in monetary units (net monetary benefit – NMB) or in health units (net health benefit – NHB).

ICER metric

$$RDCR_{x,y} = \frac{C_y - C_x}{R_y - R_x}$$

Net Benefit metric

$$BMN = \lambda\,(R_y - R_x) - (C_y - C_x)$$

$$BSN = \frac{1}{\lambda}(C_y - C_x) - (R_y - R_x)$$

λ corresponds to society's willingness to pay for a health gain of one unit (in life-years or QALYs).

Until now, ICER has been the most widely used metric. However, several limitations to this metric have been described in published literature since 1998 (Stinnett et Mullahy 1998), targeting in particular the analysis of uncertainty. The Net Benefit metric has been put forward to resolve those issues (O'Brien et Briggs 2002) and to tighten up and simplify the interpretation of the probabilistic sensitivity analysis*, particularly when multiple comparators are used (Briggs, Claxton et Sculpher 2008).

Incremental cost-effectiveness ratio (ICER)

For a result presentation model, please refer to the document "Formatting a technical report. Supporting document for the drafting of technical reports filed with the CEESP". (Haute Autorité de Santé 2016).

- The costs and health outcomes of each intervention should be summarised in a table in order to highlight all positions of strict dominance* or extended dominance*.
- All interventions and the cost-effectiveness frontier* should be represented graphically in terms of cost / health outcome.

The results should be interpreted.

Net Benefit metric

The choice of expressing the result in terms of NMB or NHB is up to the evaluator.

For a result presentation model, please refer to the document "Formatting a technical report. Supporting document for the drafting of technical reports filed with the CEESP". (Haute Autorité de Santé 2016).

- In the absence of any pre-defined reference value(s) for France, the results should be presented graphically in terms of net benefit / willingness to pay.
- The willingness-to-pay interval used to estimate the NB should be wide enough to observe all the interventions that maximise the net benefit.
- A table should show the value of the willingness to pay and net benefit at each point of intersection.

The results should be interpreted.

5.1.3. Analysing the uncertainty associated with the conclusion of the evaluation

Each evaluation involves a degree of uncertainty, imprecision and methodological inconsistency. The results of the analyses making it possible to explore this uncertainty should systematically be interpreted, as the analysis is not limited to summarising quantitative results.

A clear, justified discussion should make it possible to estimate the robustness of the conclusion. For the evaluation of cost-effectiveness, the conditions under which the cost-effectiveness frontier* would be altered should be defined.

5.1.4. Discussing the conclusions of the evaluation

The conclusions of any evaluation should be discussed in order to provide a useful interpretation to aid decision-making. This implies a clear statement on the appropriateness of the information provided to substantiate the final objective.

- The discussion should make it possible to clearly distinguish between demonstrable results and those which stem from an exploratory approach.
- Quantitative results should be interpreted in keeping with the objective of the economic evaluation.
- The degree of confidence associated with the results should be detailed.
- The transposability of the results to current practice should be discussed, based on the credibility of the assumptions, the methodological choices and the data used in the model.

Specificity of an evaluation produced within the framework of Article R. 161-71-3 of the French Social Security Code

For all of the products concerned, the main objective is to evaluate the product's expected cost-effectiveness in patient care at the time of submission of the dossier. The results that meet this objective should be presented in accordance with the procedures set out above.

Concerning products that have already been evaluated, the discussion should inform the decision-maker on the progression of cost-effectiveness since the previous evaluation (re-listing) or between indications (indication extension).

The information discussed will either be quantitative or qualitative, depending on the data available on the prior situation.

5.1.5. Analysing the potential impact of the adoption of an intervention for each funder

An analysis of the interventions' sources of funding should be conducted when an impact is expected concerning possible expense transfers following the choice of an intervention over others. The changes in expenses for each funder[46] should be substantiated.

This implies the separation of the expenses borne by the users, the national health insurance, and the supplementary health insurance schemes. These expenses are valued on the basis of the applicable rules concerning medical tariffs and reimbursement rates. To ensure the effectiveness and comparability of the conclusions of the evaluations, it is recommended, by convention, that the following allocation key be applied[47]:

- the compulsory national health insurance defrays the "Social Security" portion of the tariff, after deduction of the patient deductible or flat-rate fees;
- the supplementary health insurance schemes defray the portion of the tariff which is not non-reimbursed by the compulsory national health insurance (excluding the patient deductible);

[46] The analysis of these transfers between the different funders cannot be carried out in a budget impact analysis, which generally involves adopting a national health insurance perspective.

[47] For example, for an adult consultation with a specialist, invoiced €30 (€7 surcharge), the default allocation will be as follows:

- Compulsory National Health Insurance: €23 (tariff)*0.7 ("Social Security" share) – €1 (patient deductible) = €15.10;
- Supplementary health insurance scheme: €23 (tariff)*0.3 (supplementary health insurance share) = €6.90;
- Patient: €1 (patient deductible) + €7 (fee surcharge) = €8.

- individuals defray[48] any fee surcharges, the patient deductible and the cost of non-refundable products and services.

All available data to describe the real breakdown of the funding should be considered, provided that their reliability can be proven.

For example, individuals with recognised chronic conditions are fully reimbursed for the goods and services covered by the programme, but solely within the scope of their exonerating chronic disease, while those without recognised chronic conditions are covered at the usual rate. The funding analysis should reflect as best as possible the breakdown of individuals with and without recognised chronic conditions.

The distinction between the amount to be paid by the patients and the portion covered by supplementary health insurance schemes is difficult to make. The analysis may be improved when data are available on the proportion of users covered by a supplementary health insurance scheme, the level of coverage, and the nature of the goods and services covered.

If a funder cannot be included in the analysis (e.g. due to lack of data), its relative weight in the funding of the interventions should be discussed.

5.2. Presenting the results of the evaluation in a transparent way

Guideline 27

The economic evaluation should be presented is a well-structured, clear and detailed way. The methodology should be transparent. The data and sources used should be clearly presented.

For each intervention, non-discounted values should be presented for each major cost or result component. The total costs and health outcomes obtained on the main criterion should then be calculated and discounted.

Special attention should be paid to the writing and presentation of the evaluation, which entails compliance with two requirements. First, the report should contain enough information to allow a critical appraisal of the validity of the analysis. Second, the report should be written in the clearest possible way (Haute Autorité de Santé 2016).

A short summary (2 pages) should be included in the introduction, presenting, in non-technical language, the subject matter, the method used, the main results and the conclusion of the economic evaluation.

Each major cost and outcome component should be valued without discounting, for all interventions.

The breakdown level should depend on the type of intervention and the measurement methods used. For example, the different cost components may relate to the chronology of the health interventions (acute phase cost, re-intervention cost, chronic phase cost) or the nature of the interventions (hospital cost, cost of healthcare products, transport cost, cost of caregivers, etc.). If the outcomes are estimated in number of QALYs, it is necessary to distinguish between life years and quality-adjusted life years.

The outcomes and total costs of all interventions should then be calculated and discounted.

[48] This practice overestimates the portion paid by patients since supplementary insurance policies may defray fee surcharges, the patient deductible, as well as the cost of certain non-reimbursed health products.

Table of appendices

Annex 1. Sample data extraction sheets

Sample data extraction sheet based on a study aimed at evaluating the effect of a health intervention **Source: (adapted from) Guidelines International Network (Mlika-Cabanne, Harbour, & de Beer, 2011)**	
Reference	Author, title of the article, source, publication date
Sources of funding and competing interests	Quote the source of funding and its type (public research fund, government, academic, industry, etc.)
Setting	Number of centres, countries involved, healthcare setting (urban, rural, etc.)
METHOD	
Study protocol	Describe the study protocol: randomised controlled trial, non-randomised comparative trial, case-control study, retrospective study, cohort study, before and after study, etc.
Eligibility criteria	State the inclusion and exclusion criteria
Interventions	Provide details of the interventions studied (dosage, frequency, treatment duration, length of follow-up, etc.).
Primary outcome measure	
Secondary outcome measure(s)	Where applicable, specify the method used to control the alpha risk
Sample size	Number of patients included in the study. Indicate the power calculation
Randomisation method	Describe the randomisation method
RESULTS	
Number of subjects	Number of patients selected per group; Number of patients analysed per group. Specify whether intention-to-treat analysis.
Study duration	Start and end dates of the study (state if this included follow-up period)
Patient characteristics and group comparability	Summarise the baseline characteristics of the patients and specify any differences
Treatment effect on primary outcome	Summarise the data per treatment group (mean, median, difference, p value, confidence interval)
Treatment effect on secondary outcome	Summarise the data per treatment group (mean, median, difference, p value, confidence interval)
Adverse effects	Summarise the data per treatment group (mean, median, percentage, difference, p value, confidence interval)
Critical appraisal	
Authors' conclusions	Report authors' conclusions
Critical appraisal and reviewer's conclusion	Comment on the study's external validity and internal validity. General comments and conclusion.
General comments	

Sample data extraction sheet based on an cost-effectiveness study Source: HAS Public Health and Economic Evaluation Department (SEESP)	
Reference	Main author, title of the article, source, publication date
Context	State the country(ies) is which the economic evaluation was conducted and the year of validity of the conclusions
Structural choice	
Perspective	State the perspective adopted
Time horizon	State the time horizon (lifetime or a specific duration) and the discount rate
Population analysed	State the indication concerned
Interventions studied	State the interventions included and the interventions excluded.
Analysis method and criteria	State whether the study is a CUA, CEA or cost minimisation study, specifying the outcome criteria (event avoided, life years, or QALY).
Protocol	Specify the protocol of the cost-effectiveness study (trial-based study, modelling).
Modelling	
Type of model	State the model type and structure
Simulated population	Summary of the simulated patients' characteristics and analysis of their representativeness. State whether subgroup analyses have been conducted
Duration of simulation and extrapolation	State the difference between the clinical data observation period and the simulation period. Explain the underlying assumptions in terms of long-term treatment effect.
Approval	State whether a validation study has been conducted
Exploration of uncertainty	Summarise the sources of uncertainty explored and the approaches used
Data	
Description of clinical data	Summary of data sources (clinical trial, meta-analysis, summary review, expert opinion).
Description of useful data	State the evaluation tool (EQ-5D-5L, EQ-5D-3L, TTO, SG, etc.) and the data source (clinical trial, ad-hoc study, meta-analysis, summary review).
Description of cost data	State the cost items evaluated, specifying the year and the reference currency). Summary of cost data sources.
Conclusions	
Results	State the results in terms of ICER or BN for the base-case analysis and subgroups
Uncertainty	Identify the main results making it possible to estimate the uncertainty level.
Critical appraisal and conclusion	

Annex 2. Sample grids to evaluate the quality of an cost-effectiveness study

Sample grid to evaluate the quality of an cost-effectiveness study **Source: (adapted from) Drummond et al. (Drummond & al., 2015)**	**YES**	**NO**
Has a precise, answerable question been formulated?		
The study takes account of both the costs and the outcomes of the interventions.	☐	☐
The study compares all relevant options on the clinical level.	☐	☐
A specific viewpoint was adopted and the study was positioned in a particular decision-making context.	☐	☐
Were competing alternatives described in a comprehensive, detailed manner?		
No important alternative was omitted.	☐	☐
The “do nothing” alternative has been envisaged and studied if relevant.	☐	☐
The alternatives' descriptive elements have been presented (frequency, population analysed, design of the intervention, etc.).	☐	☐
Has the effectiveness of the intervention been established in actual practice?		
Effectiveness has been established by a randomised controlled clinical trial, whose protocol reflects what would normally happen in current practice.	☐	☐
Effectiveness has been established through a summary review of clinical trials of good methodological quality.	☐	☐
Effectiveness has been established through observational data or assumptions, with an analysis of biases in the conclusions.	☐	☐
Have the most important costs and health effects been identified for each alternative?		
Have the different relevant viewpoints been examined, concerning costs as well as health effects.	☐	☐
No important health effect has been omitted. If an important health effect has not been examined, this choice has been justified.	☐	☐
No important cost has been omitted. If an important cost item has not been examined, this choice has been justified.	☐	☐
Have the costs and health effects been measured correctly, in appropriate physical units?		
All identified outcomes and cost items have been measured.	☐	☐
The method used for the quantification of the resources consumed is valid. Unit costs have been detailed (tariffs, market prices, etc.) and are suited to the perspective adopted.	☐	☐
The measurement of health outcomes is suited to the question posed (life years, event avoided, preference score, etc.). The method used to measure the outcomes is valid.	☐	☐
The sources of information are clearly identified and the most relevant source has been given priority.	☐	☐
Have future costs and health outcomes been discounted?		

Sample grid to evaluate the quality of an cost-effectiveness study Source: (adapted from) Drummond et al. (Drummond & al., 2015)	YES	NO
The costs and outcomes have been discounted at the same rate.	☐	☐
The discount rate is known and has been justified.	☐	☐
Has uncertainty been factored into the estimation of costs and health outcomes?		
A sensitivity analysis (deterministic and probabilistic) has been presented, covering all uncertain key parameters.	☐	☐
In the deterministic analysis, the value intervals have been justified.	☐	☐
In the probabilistic analysis, the statistical analyses are suited to the nature of the key parameters and their distribution has been presented and justified.	☐	☐
The uncertainty involved in the conclusions of the economic evaluation is known and has been discussed (confidence intervals, confidence ellipse, acceptability curve).	☐	☐
Is the interpretation of the conclusions of the economic evaluation appropriate?		
An analysis of the differences in the costs and health outcomes of the competing alternatives has been conducted and presented.	☐	☐
If an aggregate indicator has been provided (cost-outcome ratio), it has been correctly interpreted.	☐	☐
The alternatives on the cost-effectiveness frontier have been identified.	☐	☐
The study is transparent on its limitations.		
The conclusions have been compared, from a critical viewpoint, to those of other studies on the same topic.	☐	☐
The study addresses the issue of generalising the conclusions for other contexts or different groups of patients.	☐	☐
The study takes account of other decision-making factors (ethics, funding, organisational/implementation aspects, etc.).	☐	☐

Sample grid for the evaluation of the quality of an cost-effectiveness model **Source: (adapted from) M. Weinstein et al. (Weinstein M. , 2003) (Nuijten, 1998)**	**YES**	**NO**
The structure and assumptions of the model are consistent with current clinical knowledge and economic aspects.		
The cost and outcome indicators are consistent with the perspective of the cost-effectiveness evaluation.	☐	☐
The simulated population is representative of the population analysed Population subgroups have been simulated to take account of differences in event probabilities, quality of life, costs or comparators.	☐ ☐	☐ ☐
All relevant alternatives have been included as comparators.	☐	☐
The model correctly factors in the time dimension. – The duration of the simulation arbitrates between the duration of the impacts in terms of outcome and cost and the uncertainty linked to extrapolation. – The duration of the cycles has been justified (pace of development of the disease, symptoms, decisions concerning treatment or costs). – The costs and outcomes have been discounted at an acceptable rate.	☐	☐
The model's assumptions have been detailed and their empirical validity (or, otherwise, their general acceptability) has been demonstrated.	☐	☐
The structure of the model is consistent with current knowledge concerning the history of the modelled disease and the causal relationships between the different variables.	☐	☐
The simplifications inherent in the structure of the model have been justified and do not alter the conclusions.	☐	☐
The model statuses have been justified in terms of their clinical importance, and their relationship with the final outcome criteria or costs.	☐	☐
If the transition probabilities depend on prior events, the model includes that memory.	☐	☐
The cost and outcome data are relevant and clearly described.		
A systematic review of conclusive data has been conducted and presented on the key variables.	☐	☐
The data sources are transparent. If an identified data source has been excluded, that choice has been justified.	☐	☐

Sample grid for the evaluation of the quality of an cost-effectiveness model **Source: (adapted from) M. Weinstein et al. (Weinstein M. , 2003) (Nuijten, 1998)**	**YES**	**NO**
If the data are based on experts' opinions, the data collection method has been presented in detail and meets the quality standards.	☐	☐
The mathematical model is suited to the subject matter		
The key stages in the development of the model have been detailed.	☐	☐
The mathematical approach used has been justified in relation to alternative approaches.	☐	☐
If the model includes data stemming from other models, the methods used have been described and are in keeping with the biostatic and epidemiological validity criteria. In the case of meta-analysis, the heterogeneity between data sources has been investigated.	☐	☐
The data have been correctly input in the model		
The measurement units, time intervals and population characteristics are mutually consistent throughout the model.	☐	☐
If a Monte Carlo simulation has been used: – the random simulation error is below the size effect; – the sensitivity analysis rests on the generation of pseudorandom numbers with a fixed seed.	☐	☐
If a cohort simulation is used, the sensitivity analysis rests on a simulation using parameter-based probability distribution.	☐	☐
The model includes a comprehensive sensitivity analysis		
A sensitivity analysis is available for all key parameters.	☐	☐
The methods used to conduct the sensitivity analyses are appropriate.	☐	☐
Mathematical models: sensitivity analyses are presented if one of the possible, and credible, alternative approaches can lead to a different conclusion.	☐	☐
The data used in the sensitivity analysis have been justified. – For non-probabilistic sensitivity analyses: lower and upper limits are provided for point estimates. – For probabilistic sensitivity analyses: the specification of probability distributions rests on a clearly defined method.	☐	☐
The uncertainty surrounding the model's conclusions has been studied and described.	☐	☐

Sample grid for the evaluation of the quality of an cost-effectiveness model Source: (adapted from) M. Weinstein et al. (Weinstein M. , 2003) (Nuijten, 1998)	YES	NO
The model is valid.		
The model has undergone technical verification tests	☐	☐
The model has undergone internal validity tests.	☐	☐
The model has undergone external validity tests.	☐	☐
The model has been compared to existing models and the differences have been discussed.	☐	☐

Annex 3. Scientific evidence level and classification

In 2013, HAS published a literature review presenting the various French and foreign systems used for the production of good practice guidelines. This work did not call into question the classifications presented in the guide published by ANAES in 2000.

Table 3: General classification of a study's evidence level (Agence nationale d'accréditation et d'évaluation en santé. 2000) (Haute Autorité de Santé 2013)

Level of evidence	Description of the level of evidence
High	– The protocol is appropriate to address the issue in the best possible way. – The study has been conducted without any major bias. – The statistical analysis is suited to the objectives. – The power is sufficient.
Intermediate	– The protocol is appropriate to address the issue in the best possible way. – The power is clearly insufficient (insufficient number of subjects or insufficient power a posteriori). – There are minor anomalies.
Low	Other types of studies

Scientific evidence is appraised during the aggregation of the results of all selected studies and constitutes the conclusion of the tables in the literature summary. The classification of scientific evidence relies on the existence of data in published literature to address the questions posed, the level of evidence of available studies, and the consistency of their results.

Table 4: Classification of scientific evidence into 3 grades (Agence nationale d'accréditation et d'évaluation en santé. 2000) (Haute Autorité de Santé 2013)

Scientific evidence grade	Level of scientific evidence provided by existing literature
Grade A **Established scientific evidence**	Level 1 – High-powered randomised controlled trials – Meta-analysis of randomised controlled trials – Decision analysis based on properly conducted studies
Grade B **Scientific presumption**	Level 2 – Low-powered randomised controlled trials – Properly conducted non-randomised controlled studies – Cohort studies
Grade C **Low scientific level of evidence**	Level 3 – Case-control studies
	Level 4 – Comparative studies with major biases – Retrospective studies – Case series – Descriptive epidemiological studies (cross-sectional and longitudinal).

Annex 4. EQ-5D-3L classification system

To use this instrument, you must register your study project with the EuroQol Group, which will inform you on user conditions, as well as license fees where applicable.

All required information is available on the website www.euroqol.org

Mobility

I have no problems walking	❑
I have problems walking	❑
I have to stay in bed	❑

Self-care

I have no problem taking care of myself	❑
I have problems washing or dressing myself	❑
I am unable to wash or dress myself	❑

Usual activities *(examples: work, studies, housework, family/leisure activities)*

I have no problems doing my usual activities	❑
I have problems doing my usual activities	❑
I am unable to do my usual activities	❑

Pain/Discomfort

I have no pain or discomfort	❑
I have moderate pain or discomfort	❑
I have extreme pain or discomfort	❑

Anxiety/Depression

I am not anxious or depressed	❑
I am moderately anxious or depressed	❑
I am extremely anxious or depressed	❑

The value set applicable to France was published in 2013 (Chevalier & de Pouvourville, 2013). It makes it possible to calculate the utility score associated with an EQ-5D-3L profile by applying cumulative decrements according to the following function.

U(E) = 1 – u1 - u2 - u3 - u4 - u5 – N3

Table 5: French weighting matrix associated with the EQ-5D-3L classification system

Dimension	Response level	Decrements	
Mobility	1	u1	0
	2		0.15
	3		0.37
Self-care	1	u2	0
	2		0.21
	3		0.32
Usual activities	1	u3	0
	2		0.16
	3		0.19
Pain / Discomfort	1	u4	0
	2		0.11
	3		0.26
Anxiety / Depression	1	u5	0
	2		0.09
	3		0.20
Constant	If at least one dimension is at level 3	N3	0.17

Annex 5. EQ-5D-5L classification system

For each section, tick ONE box – the one which best describes your health TODAY

MOBILITY

I have no problems walking ❑
I have slight problems walking ❑
I have moderate problems walking ❑
I have severe problems walking ❑
I am unable to walk ❑

SELF-CARE

I have no problems washing or dressing myself ❑
I have slight problems washing or dressing myself ❑
I have moderate problems washing or dressing myself ❑
I have severe problems washing or dressing myself ❑
I am unable to wash or dress myself ❑

USUAL ACTIVITIES *(e.g. work, studies, housework, family/leisure activities)*

I have no problems doing my usual activities ❑
I have slight problems doing my usual activities ❑
I have moderate problems doing my usual activities ❑
I have severe problems doing my usual activities ❑
I am unable to do my usual activities ❑

PAIN/DISCOMFORT

I have no pain or discomfort ❑
I have slight pain or discomfort ❑
I have moderate pain or discomfort ❑
I have severe pain or discomfort ❑
I have extreme pain or discomfort ❑

ANXIETY/DEPRESSION

I am not anxious or depressed ❑
I am slightly anxious or depressed ❑
I am moderately anxious or depressed ❑
I am severely anxious or depressed ❑
I am extremely anxious or depressed ❑

The value set applicable to France was published in February 2020 (Andrade, Ludwig, & Goni, 2020). It makes it possible to calculate the utility score associated with an EQ-5D-5L profile by applying cumulative decrements according to the following function.

$U(E) = 1 – u1 - u2 - u3 - u4 - u5$

The values stemming from the age/gender-adjusted model are given priority in the reference case analysis.

Table 6: French weighting matrix associated with the EQ-5D-5L classification system

	Mobility	Self-care	Usual activities	Pain / Discom-fort	Anxiety / De-pression
Non-adjusted model					
Level 2	0.0338811	0.0374599	0.0310561	0.0231809	0.0192951
Level 3	0.0433628	0.0494621	0.0367977	0.0487449	0.0484484
Level 4	0.1778412	0.1715431	0.1558802	0.2660137	0.2027724
Level 5	0.318512	0.2580992	0.2380036	0.4455017	0.261614
Adjusted model (reference case analysis)					
Level 2	0.03759	0.03656	0.03313	0.02198	0.02046
Level 3	0.04774	0.050781	0.03979	0.04704	0.04683
Level 4	0.17949	0.172251	0.15689	0.26374	0.20005
Level 5	0.32509	0.258331	0.24005	0.44399	0.25803

Annex 6. HUI3 classification system

This instrument is not freely available. It is managed by Health Utility Inc. and the fee to use the French version of the questionnaire and associated user manual is CAN$4,000.

All required information is available on the website www.healthutilities.com.

The value set applicable to France was published in February 2002 (Le Gales, Buron, & Costet, 2002)

Table 7: HUI3 weighting matrix

Dimensions		1	2	3	4	5	6
Vision	u1	01:00	0.96	0.86	0.80	0.70	00:49
Hearing	u2	01:00	0.94	0.92	0.87	0.83	0.66
Speech	u3	01:00	0.93	0.89	0.84	0.64	
Ambulation	u4	01:00	0.93	0.87	0.79	0.73	00:57
Dexterity	u5	01:00	0.93	0.87	0.82	0.77	0.61
Emotion	u6	01:00	0.93	0.77	0.65	00:45	
Cognition	u7	01:00	0.94	0.90	0.84	0.76	00:56
Pain	u8	01:00	0.94	0.87	0.73	00:51	

Annex 7. Examples of resources to be identified according to their economic sphere (non-exhaustive list)

Table 8: Type of resources according to their initial economic sphere

Economic sphere	Examples of resources to be identified
Health sphere	– Stays in healthcare institutions (MSO, FCR, HAH, Psy); – Use of emergency services; – Outpatient care: medical consultations (private practice, clinics, healthcare centres), external consultations (in public hospitals or private for-profit or non-profit hospitals), clinical and technical medical procedures, medical imaging, laboratory analyses, procedures by auxiliary medical personnel, non-refundable procedures by auxiliary medical personnel, spa therapy treatments; – Medical goods: medicinal products, medical devices and other medical goods (optical devices, prosthetic devices, consumables and dressings); – Transport: mobile emergency and resuscitation service (SMUR), ambulances, light healthcare vehicles (VSL), taxis, personal vehicles and other means of transport (including public transport) – whether reimbursed or not; – Costs linked to the impact of the intervention studied on the organising of the care (e.g. training, therapeutic education programme, public health programme, changes in practices due to the introduction of an innovative healthcare product, etc.).
Medico-social and social sphere	– Medico-social institutions in the public and private sectors providing care to people suffering from deficiencies or incapacities due to age, disabilities, long-term or chronic diseases, or dependency (including care homes for dependent elderly people); – Social institutions in the public and private sectors exclusively performing social activities: social services involved in child protection, institutions for adults and families in difficulty; – Other institutions involved in social and medico-social activities (e.g. local information and coordination centres (CLICs), resource centres for autistic persons (CRAs); – Social aid services: assistance provided to elderly persons and disabled adults and children (e.g. home help, healthcare assistants, carers, etc.).
Domestic sphere	– Family and non-family caregivers' time.

Annex 8. Examples of databases that can be used for the measurement and valuation of the resources consumed

A. Hospital care

Table 9: Hospital databases

Examples of databases that can be used to measure the resources consumed		
Database	**Type of information**	**Access***
PMSI (MSO, FCR, HAH, PSY)	Exhaustive hospital stay data with anonymised individual data – Patients' sociodemographic characteristics (age, gender) – Type of institution – Characteristics of the stay: type of stay (HPG code), average duration of stay, main/associated/related diagnoses, classified procedures – Specific activities, chemotherapy and dialysis – Mode of admission (including through emergency services), mode of discharge	Aggregate data accessible on ATIH's ScanSanté website, http://www.scansante.fr/ Restricted access to individual data
Files supplementing PMSI data (Fichcomp, Fichsup)	Exhaustive data on hospital stays involving: – products invoiced on top of HRG-based tariffs (medicines: code UCD, medical devices: code LPP) – Medicinal products under a temporary authorisation for use (ATU) – Anti-thrombotic medicines – Organ removal – Flat-rate fee for peritoneal dialysis Exhaustive per-institution data on: – specific external consultations – lactarium activity – activity dedicated to initial registration for oral chemotherapy (active patient queue, number of consultations, etc.)	Aggregate data accessible on ATIH's ScanSanté website, http://www.scansante.fr/ Restricted access to individual data
SAE (annual statistics of healthcare institutions), DREES	Aggregate data available per institution status, concerning heavy equipment, specific activities (emergency medical services), the number of stays/days/sessions, etc.	Unrestricted access on the website: https://www.sae-diffusion.sante.gouv.fr/sae-diffusion/accueil.htm

* This information is provided on an indicative basis according to the information available at the time this document was written.

Table 10: hospital databases

Examples of databases that can be used for the valuation of hospital resources.		
Database	**Type of information**	**Access***

NCS MSO, HAH, FCR,	Production costs for hospital stays in MSO, HAH and FCR	Unrestricted access on the ATIH website http://www.atih.sante.fr/etudes-nationales-de-couts-sanitaires-enc/presentation
ENC EHPAD (national cost study on care homes for dependent elderly people)	Average daily cost of care (total costs and costs per activity/expenditure item) for care home residents	Unrestricted access on the ATIH website http://www.atih.sante.fr/etudes-nationales-de-couts-sanitaires-enc/presentation
Files supplementing PMSI data (Fichcomp, Fichsup)	Information on hospital stays – Purchase price of products invoiced on top of HRG-based tariffs – Purchase price of medicines under a temporary authorisation for use (ATU) – Purchase price of anti-thrombotic medicines – Valuation of stays in institutions formerly under a collective allowance or a national quantified target according to the reimbursement rate or amount reimbursed by the National Health Insurance: HRG-based tariff and supplements taken into account	Restricted access
Units costs list	Per-day cost of clinical services – HAH, MSO, PSY, FCR	Unrestricted access on ATIH's Scan Santé website, http://www.scansante.fr/

* This information is provided on an indicative basis according to the information available at the time this document was written.

B. Use of emergency services

➔ Invoicing rules

Emergency care is remunerated in two ways: a flat-rate fee per admission in the service and an annual allocation.

The aim of the emergency admission and treatment (ATU) fee is to cover the expenses resulting from the admission and treatment of patients in the emergency department of the healthcare institutions authorised to that effect. The ATU fee applies to MSO institutions.

The ATU fee is payable for each admission to the emergency department which is:

- not scheduled,
- not followed by hospitalisation in an MSO service or short-stay unit in the same institution.

This fee is added to the tariffs payable for the consultation, medical procedures and any applicable surcharge. The ATU fee and the cost of any medical procedures or consultations performed in the emergency department cannot be invoiced if the emergency admission is followed by hospitalisation in the same institution.

Moreover, the ATU fee cannot be cumulated with the environmental safety fee, nor the small equipment fee, or an HRG-based tariff if the patient is given a subsequent appointment.

At the time of publication of this document, it amounted to €25.28 (see the Order of 27 February 2008 setting the 2008 National Health Insurance resources for MSO healthcare institutions and the Order of 27 February 2009 for the 2009 resources).

The annual emergency service allocation (French: FAU) aims to cover fixed costs (personnel, equipment, etc.) and is determined according to the number of emergency department admissions for which an ATU fee is invoiced.

C. Non-hospital care: medical consultations, medical auxiliaries, etc.

The National Health Insurance databases can be used to estimate the costs of non-hospital care, according to the type of consultation or procedure refunded to beneficiaries. Certain National Health Insurance databases are freely accessible. However, access to the DCIR's exhaustive data – making it possible to analyse care pathways in a comprehensive way – is restricted (see factsheet on general databases).

Table 11: Database for non-hospital care

Database	Type of information	Access*
Freely accessible National Health Insurance databases on healthcare expenditures		
National Health Insurance healthcare expenditures excluding hospital care, National Data (aggregate data from SNIIRAM)	Total monthly reimbursements per type of service, type of provider, and type of prescriber Information on the amounts reimbursed, the reimbursement bases, the reimbursement rates, and surcharges. Possibility of estimating an average cost, with surcharge (e.g. consultation per type of specialist, transport, medical/paramedical procedures) Period covered: 2010 to 2017	Unrestricted access on the National Health Insurance's Open Data website http://open-data-assurance-maladie.ameli.fr/index.php
Open Damir (aggregate data from SNIIRAM)	Reimbursements by the National Health Insurance (all schemes). This database supplements the previously mentioned national database with information on hospital services directly invoiced to the National Health Insurance, information on the beneficiary (gender, age bracket, region of residence, supplementary universal healthcare coverage – CMUC), information on the care provider and prescriber, etc. Possibility of accurately estimating an average cost with surcharge, for example by targeting the beneficiaries. Period covered: 2009 to 2016	Unrestricted access on the National Health Insurance's Open Data website http://open-data-assurance-maladie.ameli.fr/index.php
Freely accessible National Health Insurance databases on independent healthcare professionals		

Database	Type of information	Access*
SNIR	Information on demographics, the volume of activity, the prescriptions and fees of doctors and other independent healthcare professionals. Possibility of estimating – per specialisation, and per department or region – an average fee (fee with surcharge and travel expenses) per procedure (including consultations, visits and technical procedures), per specialisation Most recent period covered: 2013	Unrestricted access on the National Health Insurance website https://www.ameli.fr/l-assurance-maladie/statistiques-et-publications/donnees-statistiques/professionnels-de-sante-liberaux/donnees-completes/2013-tableaux-personnalisables.php
Freely accessible EcoSanté (IRDES) databases		
Eco-Santé (IRDES)	Data on the population's health state, healthcare consumption (non-hospital care, independent healthcare professionals, hospital activity), national healthcare accounts, the national objective for healthcare expenditures (ONDAM), social protection accounts, social aid and global demographic and economic indicators. For each specialisation: total fees (with or without surcharges), travel expenses, overall activity (consultations, visits, technical procedures), and prescriptions. Source: CNAM data, SNIR file. Most recent data: 2013	Data accessible on the IRDES website http://www.ecosante.fr/index2.php?base=DEPA&langh=FRA&langs=FRA

* This information is provided on an indicative basis according to the information available at the time this document was written.

Medical consultations of general practitioners and specialists

➔ Invoicing rules

Medical consultations are subject to conventional tariffs, to which may be added surcharges and extra fees (nights, Sundays, national holidays, post-hospitalisation follow-up of a high-comorbidity patient, etc.), as well as any additional remuneration (for the follow-up of elderly people, patients with chronic conditions, GP targets, etc.) and flat-rate allocations (for doctors providing primary care, paediatric care, medical supervision of main condition, etc.).

Procedures by medical auxiliaries

➔ Invoicing rules

The invoicing of nursing procedures is based on the General Nomenclature of Medical Procedures (NGAP)[49]. Nursing procedures break down into the following 3 categories: medical nursing procedures (AMI); nursing care procedures (AIS); nursing care programmes (DI).

Various indemnities may be added to the invoiced procedure: complex surcharges, supplementary flat-rate allocations (care programme for dependent patients, etc.), services invoiced according to time spent (daily clinical surveillance and prevention sessions), and travel expenses.

For physiotherapists, the listed fees vary depending on where the services are performed (private practice (AMK) or healthcare institution (AMC)).

[49] General Nomenclature of Medical Procedures (NGAP) in force since the UNCAM decision of 11 March 2005: http://www.ameli.fr/fileadmin/user_upload/documents/NGAP.pdf

➔ Valuation rules

Medical consultations and procedures by medical auxiliaries are valued in close keeping with their production cost for each specialisation by estimating an average cost considering applicable fees and any add-on. External private-sector consultations are valued in the same way as consultations with independent specialists.

In specific cases which involve considerable nursing care and long visits, the average costs estimated at the national level for any type of care may no longer be appropriate. In such cases, it is necessary to take the duration of the visits into account and apply the following rule[50]:

- any visit lasting up to 120 minutes will be valued as 3 AISs (nursing care procedures) + 1 IFD (travel allowance) per 30-minute period
- and any visit exceeding 120 minutes will be valued as 13 AISs per 6-hour period.

For non-reimbursed medical auxiliary procedures (dieticians, occupational therapists, psychologists, etc.) resources for which there is no listed tariff are valued at the average current price if observable, or through another method to be specified.

D. Costs of medico-technical procedures

Clinical and technical medical procedures

➔ Invoicing rules

Medical procedures are invoiced based on CCAM tariffs. A fixed contribution of €18 is payable by the patient for any procedure (performed during a hospital stay or an external consultation) whose tariff is €120 or more in the CCAM or whose coefficient is 60 or more in the NGAP.

➔ Valuation rules

Medical procedures are valued based on applicable CCAM tariffs, to which is added an average national cost taking surcharges into account.

Laboratory procedures

Table 12: Freely accessible National Health Insurance database on medical laboratory expenses

Database	Type of information	Access*
BioIAM		
Open Bio **(aggregate data from SNIIRAM)**	Set of annual databases on reimbursements and the number of beneficiaries of medical laboratory procedures from 2014 to 2016. Information on reimbursed and reimbursable amounts, the total annual number of procedures per pathophysiological group according to the nomenclature of medical laboratory procedures (TNB), based on elements of information on the beneficiaries (age bracket, gender, region of residence) and the prescriber's specialisation. Period covered: 2014 to 2016	Unrestricted access on the National Health Insurance's Open Data website http://open-data-assurance-maladie.ameli.fr/index.php

[50] http://www.sfes.info/Atelier-de-standardisation-des,294.html

* This information is provided on an indicative basis according to the information available at the time this document was written.

➔ Invoicing rules

Laboratory procedures are invoiced based on the tariffs listed in the national laboratory table (TNB) (coefficient*value of key-letter B). Sampling costs or additional flat-rate fees may be added to those tariffs.

➔ Valuation rules

Laboratory procedures are valued in close keeping with their production cost by estimating an average cost considering the cost of the procedures and the various flat-rate fees and add-ons.

Medical imaging procedures

➔ Invoicing rules

Medical imaging procedures are invoiced based on CCAM tariffs.

Certain specificities are to be considered for CAT and MRI procedures: a flat-rate technical fee is invoiced on top of the tariff applicable to the procedure. The amount of these flat-rate fees depends on a scale considering the geographical location, age of the equipment, its power, the number of procedures performed per year, etc. The purpose of these technical fees is to offset the cost of purchase and operation of the equipment.

➔ Valuation rules

Medical imaging procedures are valued based on applicable CCAM tariffs, to which are added a mean national cost taking surcharges into account.

CAT and MRI procedures are valued based on applicable CCAM tariffs, to which are added a mean amount in respect of the flat-rate technical fee. This mean amount is estimated on the basis of an average reference activity (for all geographic areas, CAT class and MRI power combined) and an average tariff for the full-rate technical fee, except in cases where the details on the equipment's distribution per geographic area, class, power and status (amortised / non-amortised) are known. A method for estimating average technical fees for CAT and MRI procedures has been put forward by Robert Launois (2014).

The amounts of the full-rate or reduced-rate fees may be used in a sensitivity analysis.

Concerning bone scan and PET procedures, the tariffs include the supply of radiopharmaceutical products. Where the price of the product is very high, it is added to the tariff applicable to the procedure.

The contrast agents used in radiology in institutions formerly under a collective allocation are valued based on the retail price (inclusive of VAT), to which a discount is applied in a sensitivity analysis.

E. Cost of medicinal products

The breakdown of sales of medicinal products within the same therapeutic class is measured using the databases mentioned below whenever the number of units sold/provided is known.

The following databases may be used to measure and value the amounts reimbursed for medicinal products. Certain National Health Insurance databases are freely accessible. However, access to the DCIR's exhaustive data – making it possible to analyse care pathways in a comprehensive way – is restricted (see factsheet on general databases).

Table 13: Databases on medicinal products

Database	Type of information	Access*
Freely accessible National Health Insurance databases on medicinal product expenses		
Open MEDIC (aggregate data from SNIIRAM)	Set of annual databases on the use of medicinal products provided in retail pharmacies from 2014 to 2016. Information on the amounts reimbursed, the number of boxes provided for each ATC class, according to elements of information on the beneficiaries (age bracket, gender, region of residence) and the prescriber's specialisation. Period covered: 2 years + the year in progress	Unrestricted access on the National Health Insurance's Open Data website http://open-data-assurance-maladie.ameli.fr/index.php
Open PHMEV	Set of annual inter-scheme databases on the medicinal products prescribed by public healthcare institutions and private healthcare institutions of public interest, and provided in retail pharmacies Data on the number of boxes provided, the amount reimbursed per ATC class, according to information on the beneficiary (age bracket and gender) and the institution, Period covered: 2 years + the year in progress	Unrestricted access on the National Health Insurance's Open Data website http://open-data-assurance-maladie.ameli.fr/medicaments/index.php#Open_PHMEV
Retroced_AM	Data on medicinal products within the framework of hospital retrocession by the general health insurance system (including local mutualist sections) Information for each medicinal product by UCD code on reimbursement bases (including a retrocession margin), the amounts reimbursed, and the number of units reimbursed. Period covered: 2 years + the year in progress	Unrestricted access on the National Health Insurance website https://www.ameli.fr/l-assurance-maladie/statistiques-et-publications/donnees-statistiques/medicament/retroced-am.php
BdM_IT	Information on prices (face prices) for – approved medicinal products reimbursed to individuals covered by the French national health system (identified by a CIP code) – medicinal products transferred and invoiceable on top of HRG-based tariffs by healthcare institutions (identified by a UCD code).	Unrestricted access on the National Health Insurance website http://www.codage.ext.cnamts.fr/codif/bdm_it/index.php?p_site=AMELI
Databases on non-hospital sales/prescriptions/dispensing of medicinal products		
ANSM database on taxes applicable to medicinal products, (exhaustive list of filers, annual)	Medicinal products sold by laboratories to hospitals and retail outlets (actual turnover, number of units sold to hospitals and retail outlets, CIP code).	Exhaustive data with restricted access Aggregate data in ANSM's annual report
SDM/SDM Spé/DMSO IQVia (panel of pharmacies, excluding	Dispensing of prescription/non-prescription medicinal products or medical devices in retail pharmacies (CIP/ATC code, units, turnover, manufacturer price	Fee-charging databases

Database	Type of information	Access*
French overseas departments), monthly, history> 5 years)	exclusive of VAT, retail price inclusive of VAT, reimbursement rate).	
GERS databases (exhaustive list of subscribing laboratories, monthly, 10-year history)	Subscribing laboratories' sales to retail pharmacies (Sell-in). Data concerning the number of units sold, total sales exclusive of VAT (only in retail pharmacies), CIP/EphMRA/ATC code.	Fee-charging databases Access restricted to subscribers
Xpr-So-Open-health (panel of pharmacies excluding Corsica, daily, history > 3 years)	Real-time dispensing of healthcare products sold in retail pharmacies (medicinal products, medical devices, homoeopathic products) Information on the number of prescription/non-prescription units sold, origin of the prescriptions (e.g. GP, hospital, midwife), reimbursement rate, number of pharmacies, average price of non-reimbursed medicinal products, CIP/ATC code, etc.	Fee-charging database
Databases on hospital sales/prescriptions/dispensing of medicinal products		
ANSM database on taxes applicable to medicinal products, (exhaustive list of filers, annual)	Medicinal products sold by laboratories to hospitals and retail outlets (actual turnover, number of units sold to hospitals and retail outlets, CIP code).	Exhaustive data with restricted access Aggregate data in ANSM's annual report
Files supplementing PMSI data (Fichcomp)	Exhaustive data on hospital stays involving: – Products invoiced on top of HRG-based tariffs (medicinal products: UCD code, purchase price) – Medicinal products under temporary authorisation for use (purchase price)	Aggregate data accessible on ATIH's ScanSanté website, http://www.scansante.fr/ Restricted access to individual data
HOSPI PHARMA-IMS IQVia (panel of in-house pharmacies, monthly, history >5 years)	Intra-hospital dispensing and consumption of medicinal products by in-house pharmacies (Sell-out). Data available on the number of units dispensed per type of healthcare institution, and the valuation of the specialisation at the catalogue price, by UCD/ATC/EphMRA code or by specialisation.	Fee-charging database
HOSPIWARD-IMS IQVia (panel of in-house pharmacies, monthly, history >5 years)	This database supplements the HOSPI PHARMA database with the following: Analysis of the number of units dispensed (UCD) in the different hospital services, by type of hospital (CHU, CLNCL, etc.)	Fee-charging database
GERS databases (exhaustive list of subscribing laboratories, monthly, 10-year history)	Sales by laboratories partnering healthcare institutions (Sell-in) Data concerning the number of units sold, UCD/EphMRA/ATC code	Fee-charging databases Access restricted to subscribers

Database	Type of information	Access*
AFSSAPS Tax database (exhaustive, annual, since 1987)	Medicinal products sold to hospitals and retail outlets by laboratories (total sales figure, number of units sold to healthcare institutions, CIP code)	Aggregate data in the annual report
Medication dosages		
RCP, ANSM-EMA	Marketing authorisation (MA): multidisciplinary approach, prescription conditions (list, restricted/non-restricted prescription, hospital-only, etc.), recommended dosage	Data accessible on the ANSM website, EMA website or http://base-donnees-publique.medicaments.gouv.fr/

* This information is provided on an indicative basis according to the information available at the time this document was written.

F. Costs of medical devices

The following databases may be used for the measurement and valuation of the amounts reimbursed for medical devices. Certain National Health Insurance databases are freely accessible. However, access to the DCIR's exhaustive data – making it possible to analyse care pathways in a comprehensive way – is restricted (see factsheet on general databases).

Table 14: Databases for medical devices

Database	Type of information	Access*
National Health Insurance databases on medical device expenditures, Open Data access		
LPP'AM	Data on the medical devices included on the list of products and services (LPP) reimbursed for years 2006 to 2015 (general health insurance system, metropolitan France). Information on reimbursed/reimbursable amounts and the quantities reimbursed.	Unrestricted access on the National Health Insurance website https://www.ameli.fr/l-assurance-maladie/statistiques-et-publications/donnees-statistiques/liste-des-produits-et-prestations-lpp.php
LPP	Pricing information (face price) on products and services by code or name	
Databases on non-hospital sales/prescriptions/dispensing of medicinal products		
Database on the tax applicable to medical devices, ANSM (exhaustive list of laboratories, annual, since 1987)	Provider's total sales	Restricted access
SDM/SDM Spé/DMSO IMS-IQVia (panel of pharmacies excluding French overseas departments, monthly, history> 5 years)	Dispensing of prescription/non-prescription medicinal products or medical devices in retail pharmacies (CIP/ATC code, units, total sales, manufacturer price exclusive of VAT, retail price inclusive of VAT, reimbursement rate),	Fee-charging database

Database	Type of information	Access*
Xpr-So-Openhealth (panel of pharmacies excluding Corsica, daily, history > 3 years)	Real-time dispensing of healthcare products sold in retail pharmacies (medicinal products, medical devices, homoeopathic products) Data available on the number of prescription/non-prescription units sold, origin of the prescriptions (e.g. GP, hospital, midwife), reimbursement rate, number of pharmacies, average price of non-reimbursed medicinal products, CIP/ATC code, etc.	Fee-charging database
Databases on hospital sales/prescriptions/dispensing of medicinal products		
Files supplementing PMSI data (Fichcomp)	Exhaustive data on hospital stays involving products invoiced on top of HRG-based tariffs (implantable medical devices: LPP code, purchase price).	Aggregate data accessible on ATIH's ScanSanté website, http://www.scansante.fr/ Restricted access to individual data

* This information is provided on an indicative basis according to the information available at the time this document was written.

The LPPR (list of products and services qualifying for reimbursement) lays down the reimbursement basis. The price is freely determined for most devices, even though certain medical devices have a maximum sale price set by agreement with the CEPS. Certain products are reimbursed as part of a fixed package price. The data sources for the real prices charged are essentially private sources.

G. Cost of medical/non-medical transportation

The following databases and reports can be used for the measurement and valuation of the amounts reimbursed for transport. Certain National Health Insurance databases are freely accessible. However, access to the DCIR's exhaustive data – making it possible to analyse care pathways in a comprehensive way – is restricted (see factsheet on general databases).

Table 15: Transport sources

Database	Type of information	Access*
Freely accessible National Health Insurance databases on healthcare expenditures		
National Health Insurance healthcare expenditures excluding hospital care, National Data (extracted from SNIIRAM)	Total monthly reimbursements per type of service, type of provider, and type of prescriber Information on the amounts reimbursed, the reimbursement bases, the reimbursement rates, and surcharges. – Possibility of estimating an average cost, including surcharge, according to the type of vehicle used (e.g. ambulance, light medical vehicle, taxi). – Possibility of estimating the breakdown between the different modes of transport reimbursed. Period covered: 2010 to 2017	Unrestricted access on the National Health Insurance's Open Data website http://open-data-assurance-maladie.ameli.fr/index.php
Open Damir (extracted from SNIIRAM)	Total reimbursements by the National Health Insurance for all schemes combined.	Unrestricted access on the National Health Insurance's Open Data website

Database	Type of information	Access*
	This database supplements the previously mentioned national database with information on hospital services directly invoiced to the National Health Insurance, information on the beneficiary (gender, age bracket, region of residence, supplementary universal healthcare coverage – CMUC), information on the care provider and prescriber, etc. A more accurate estimation of an average cost with surcharge is possible, for example by targeting the beneficiaries. Period covered: 2009 to 2016	http://open-data-assurance-maladie.ameli.fr/index.php
Commission des Comptes de la Sécurité Sociale, June 2016	Information on the number of journeys reimbursed and the total amounts reimbursed, by type of vehicle used. Possibility of estimating an average amount reimbursed per journey, by type of vehicle.	Unrestricted access: CCSS (June 2016). Social Security accounts, 2015 results, 2016 forecasts. *Éclairage 3.2 "Les dépenses de transport et leurs disparités régionales"* (Transport expenditures and their regional variations, p. 108. http://www.securite-sociale.fr/IMG/pdf/rapport-ccss-juin2016.pdf

* This information is provided on an indicative basis according to the information available at the time this document was written.

H. General databases

Table 16: General databases

Database	Type of information	Access*
Data from the SNDS (*Système National des Données de Santé* – National Health Data System): Exhaustive data presented in aggregate and anonymised individual formats or EGB (individual data sample, created in 2005, data starting in 2003)	Anonymised individual data** Patients' socio-demographic data: gender, age, national health insurance branch office, department/municipality of residence, universal healthcare coverage (CMU), supplementary universal healthcare coverage (CMU-c) Data concerning the healthcare professional or institution (geographic details) In-kind services provided: medico-technical procedures (CCAM code), paramedical procedures, medicinal products (CIP code), medical devices (LPP code), technical procedures (NGAP code), laboratory procedures (NABM code), medical transport, hospital stays Monetary compensation provided: sick pay, disability benefits Medical information: chronic conditions (ALD code), Homogeneous Patient Groups (GHM code)	Restricted access (access according to authorisation profiles as defined by the law relative to exhaustive data)

Database	Type of information	Access*
Portail Epidémiologie France (ITMO Santé Publique d'Aviesan)	Online catalogue of the main health-related databases (excluding clinical trials) by topic and from French sources, which may be useful for the development of public health research and expertise: administrative databases, databases stemming from surveys, records of morbidity, cohorts, longitudinal studies, case-control studies, and cross-sectional studies	Data accessible on the INSERM website https://epidemiologie-france.aviesan.fr/epidemiologie-france/accueil
Databases on the longitudinal follow-up of patients, diagnoses and prescriptions (independent practitioners)		
LPD–IQVIA (formerly Thalès) (representative panel of doctors, monthly, history >10 years)	Panel of independent general practitioners and specialists Data available on the patient profile (age, gender), clinical profile (diagnosis, comorbidity), clinical tests, treatments (initial, repeat), patient follow-up and prescriber profiles.	Fee-charging databases
Disease Analyzer–IQVIA (representative panel of general practitioners, monthly, history >10 years)	Panel of independent general practitioners Data available on the patient profile (age, gender), clinical profile (diagnosis, comorbidity), clinical tests, treatments (initial, repeat), patient follow-up and prescriber profiles	
GERS database - GERS SAS (representative panel of general practitioners, monthly, history >10 years)	Panel of independent general practitioners and specialists Data available on the patient profile (age, gender), clinical profile (diagnosis, comorbidity), reimbursed tests, prescribed treatments (initial, repeat), patient follow-up, prescriber profiles and hospital stays	Fee-charging databases
Diagnosis and prescription database (independent practitioners)		
LPD IQVIA database (formerly Thalès) representative panel of doctors, monthly, history >10 years	Panel of independent general practitioners and specialists, Data available on the patient profile (age, gender), clinical profile (diagnosis, comorbidity), clinical tests, treatments (initial, repeat), patient follow-up and prescriber profiles	Fee-charging databases
Disease Analyzer IQVIA (panel of general practitioners, monthly, since 2000)	Panel of independent general practitioners Data available on the patient profile (age, gender), clinical profile (diagnosis, comorbidity), clinical tests, treatments (initial, repeat), patient follow-up and prescriber profiles	
EPPM – IQVIA (representative panel of independent doctors, quarterly, history > 10 years)	Panel of independent general practitioners and specialists, Data on specialists' prescriptions (dosages, treatment duration, co-treatment, etc.), diagnoses (CIM-10), characteristics of the patient and doctor, procedures, etc.	

* This information is provided on an indicative basis according to the information available at the time this document was written.

** The data from the PMSI and SNIIR-AM data warehouse are anonymised through a common identifier, which makes it possible to chain individual data between the two systems. All the individual data can be processed in an aggregate manner.

Annex 9. Hospital cost evaluation methods

A. Micro-costing method: precise identification of the resources consumed per intervention

This approach consists in direct and detailed observation of the resources consumed at each stage of patient care (e.g. treatment, surgical intervention, hospital stay) or consumed in the use of a health technology (e.g. in the case of innovative medical devices accompanying new medical practices) (Gold, et al. 1996). It makes it possible to appraise a health intervention's entire production cost.

The principal advantage of this method is the precision of the ensuing cost estimation. However, it is time-consuming and may potentially limit the generalisation of the results obtained.

The micro-costing approach is particularly suited for the following:

- services involving a significant proportion of labour costs or structural costs (Wordsworth, Ludbrook et Caskey 2005);
- cost components involving high variability between patients (Swindle, Lukas et Meyer 1999);
- innovation, provided it is precisely documented.

Identification and measurement of the resources consumed

The data is generally collected in real time, based on the direct observation of the resources required for the production of a health intervention (Guerre, Hayes et Bertaux 2018).

This observation makes it possible to take stock of the amounts of work (e.g. implementation and production time through the timing of the different stages) and capital expenditures (e.g. equipment, consumables, materials of various kinds, and structures whether they are fully dedicated to the activity or not) involved in the care provided to the patient.

This rigorous process requires the accurate prior identification of the different stages involved in the health intervention.

Valuation of resources consumed

The valuation of the resources consumed is based on two main approaches.

1/ The bottom-up approach: all cost components are identified and valued individually for each patient (the resources in terms of personnel, equipment and consumables are identified per unit and added up to obtain the total cost).

This is the approach most frequently associated with micro-costing since it provides a patient-centred viewpoint: calculation of the health intervention's specific unit costs and precise analysis of the costs induced by the most resource-intensive patient subgroups.

The bottom-up micro-costing approach (Wordsworth, Ludbrook et Caskey 2005) is considered as the "gold standard" to evaluate the costs of a health intervention. The cost data must be presented and described in an appropriate way (Morelle, Plantier et Dervaux 2018).

2/ The top-down approach: all cost components are identified but valued for a typical patient.

With the top-down approach, resources are valued using various sources of aggregate data (such as the healthcare institution's accounting data, the use of allocation keys or, more generally, models stemming from epidemiological and economic data) and appraised in relation to the patients covered by the evaluation. For example, this may include the running cost of an operating room for a given period, spread over the number of patients operated during that period. This approach thus assesses the cost

of an average patient and does not provide any insight into the differences in the care provided to each patient on an individual basis (this approach gives average costs per patient, unlike the bottom-up approach which makes it possible to calculate the unit costs of the care provided to each patient).

B. Gross-costing method

The gross-costing method consists in allocating the total amount of the costs borne by a healthcare institution to a particular service, and then to a patient, using predetermined allocation keys.

Identification and measurement of the resources consumed

The costs are thus calculated using the hospitals' accounting data at an aggregate level. The identification of the cost components thus corresponds to a high level of aggregation, often by assigning mean values stemming from the national administrative databases (Mogyorosy et Smith 2005).

Valuation of resources consumed

There are also two different approaches for the valuation of the resources consumed: bottom-up and top-down gross costing.

Gross costing relies on strong assumptions which may have an important impact on cost estimation precision. Indeed, the assumptions are that:

- there is no high variability between the individuals in the group for which the costs are being measured;
- the variations in practices are negligible;
- when we calculate an average cost per day of hospitalisation (since hospital days are often used as the main component of the cost of a hospital stay), we assume that the cost of the hospital stay is proportional to the duration of the stay.

To sum up, gross costing may be used when:

- a high level of precision is not necessary (e.g. modification of the duration of the stay as part of an innovative organisational strategy);
- the perspective is broad enough (society).

C. Choice between the micro-costing and gross-costing methods

The choice between the micro-costing and gross-costing methods is first an trade-off between the required precision, feasibility and cost.

Gross costing has several advantages which partly offset the drawbacks of micro costing:

- in terms of feasibility: since hospital cost data consists of aggregate data, its estimation can be done fairly quickly;
- in terms of cost: this method is inexpensive as it largely relies on administrative databases;
- the results of the study are easier to generalise.

It also has several drawbacks, including a major one given the lack of precision since a cost cannot be associated with a specific component of the hospital stay because of the aggregate data, which is examined at an overall level. With this method, differences in terms of consumption of resources (different patient profiles) are thus not known and individual variations are not considered.

Consequently, the gross-costing method cannot be used to measure modifications in the resource consumption structure (such as the impact of the introduction of a new medicine on the cost of a service), nor to conduct analyses on specific subgroups.

Annex 10. Choice of the model

Table 17: Type of models (Briggs, Wolstenholme, & Blakely, 2016, according to Brennan, Chick, & Davies, 2006).

		Cohort/aggregate-level/counts		Individual-level	
		Expected value, continuous state, deterministic	Markovian, discrete state, stochastic	Markovian, discrete state	Non-Markovian, discrete state
No interaction	Untimed	Decision tree rollback or comparative risk assessment	Simulation decision tree or comparative risk assessment	Individual sampling model: Simulated patient-level decision tree or comparative risk assessment	
	Timed	Markov model (deterministic)	Simulation Markov model	Individual sampling model: Simulated patient-level Markov model	
Interaction between entity and environment	Discrete time	System dynamics (finite difference equation)	Discrete time Markov chain model	Discrete time individual event history model	Discrete time Discrete event simulation
	Continuous time	System dynamics (ordinary differential equation)	Continuous time Markov chain model	Continuous time individual event history model	Continuous time Discrete event simulation
Interaction between heterogeneous entities/spatial aspects important		NA	NA	NA	Agent-based simulation

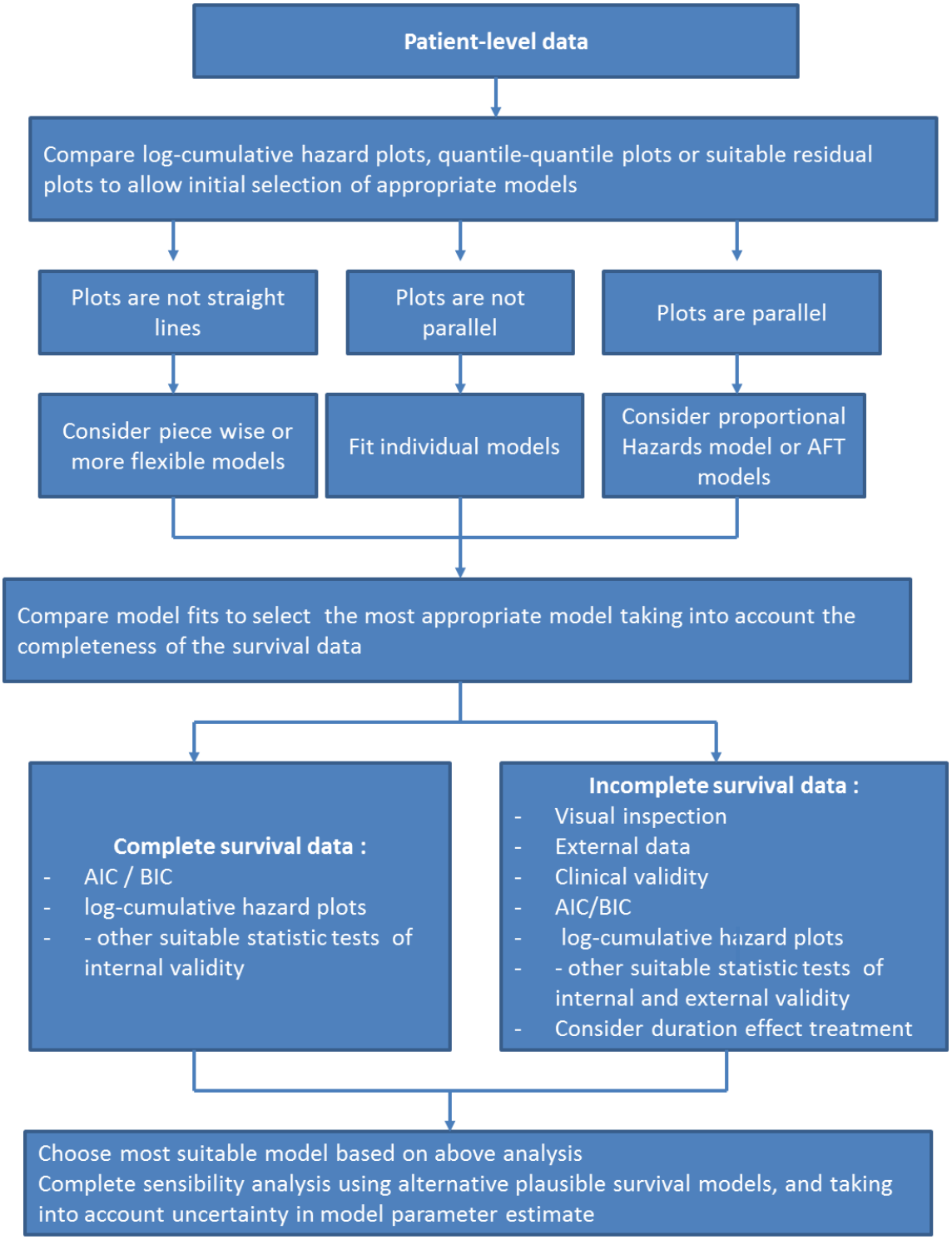
Table 18: Questions to guide the choice of a model

Source: (adapted from) Brennan, 2006 (Brennan, Chick, & Davies, 2006) and Caro 2012 (Caro & Briggs, 2012)		
Question	Example	Choice of the model
1. Is parameter variability (e.g. treatment effect variability) an important element to consider?	The effects of the intervention are small and variable over time	Need for a model that provides stochastic results
2. Is the heterogeneity of the population analysed likely to have an impact on cost-effectiveness?	Certain specific characteristics may lead to the identification of subgroups with a different cost-effectiveness	Models allowing simulations at the individual level are more flexible and better able to incorporate additional variables or modifications of assumptions
3. Do individual risk factors affect outcomes in a non-linear fashion?	Effects of age, the natural history of the disease, and comorbidity	Need to subdivide states in an aggregate model. Need to consider the use of a model allowing individual simulations if there is a high number of risk factors.
Question	Example	Choice of the model

Source: (adapted from) Brennan, 2006 (Brennan, Chick, & Davies, 2006) and Caro 2012 (Caro & Briggs, 2012)		
4. Can some of the parameters have multiple effects that may result in interactions between these parameters?	Comorbidities in the diabetes may affect the progress of the renal failure or retinopathy	Individual simulation model to be preferred
5. Is there any time spent in non-Markovian states?	Poor survival rate after an operation, moving from one age group to another, length of stay in hospital	Need to use 'fixes' in Markovian models or use non-Markovian models
6. Are there too many dimensions to be able to use a cohort approach?	Large number of risk factors and /or subdivision of states to overcome non-Markovian effects	Individual simulation probably necessary
7. Is there any "recycling" of states?	Recurrence of the same illness (e.g. heart attack, no longer any response to a medicine)	A decision tree is probably not appropriate
8. Is the phasing or timing of the events requiring decisions important?	In smokers, if lung cancer occurs before bronchitis, then the patient may die before bronchitis occurs	It is possible to have different branches in the decision tree, but a Markov model or macro-simulation model may be necessary
9. Are there direct interactions between individuals?	Infectious disease models	Models allowing interactions
10. Are there interactions due to constrained resources?	Models with resource constraints	Models allowing interactions
11. Are numerous events likely to occur in one-time unit?	Natural disaster, epidemic, risk of comorbidities (diabetes)	Need to use a model with short time intervals or a continuous-time model
12. Are interactions between individuals occurring in small populations?	Use in hospital catchment area (rather than on the national scale)	Need to consider micro-simulations because of the inaccuracies in using fractions of individuals
13. Are there intervention delays due to resource constraints which may affect costs or health outcomes?	Rapid treatment with angioplasty and stents after a myocardial infarction	Need for stochastic output and interactions
15. Can a marginal change in parameters produce a non-linear change in the performance of the intervention?	Intensive care units are suddenly full and newly arriving patients must be transferred elsewhere	Discrete event simulation is useful

Annex 11. Algorithm for the selection of survival models

Source: Nicholas Latimer. NICE DSU Technical Support Document 14: Survival analysis for economic evaluations alongside clinical trials - extrapolation with patient-level data. Report by the decision support unit June 2011 (last updated March 2013).



Annex 12. Verification and validation of a model – AdViSHE tool

Source: Vemer P., Corro Ramos I., van Voorn GA, et al. AdViSHE: A Validation-Assessment Tool of Health-Economic Models for Decision Makers and Model Users. Pharmacoeconomics. 2016 Apr; 34(4): 349-61.

Part A: validation of the conceptual model

A conceptual model describes the underlying system (e.g., progression of disease) using a mathematical, logical, verbal, or graphical representation.

Please indicate where the conceptual model and its underlying assumptions are described and justified.

A1/ Face validity testing (conceptual model): Have experts been asked to judge the appropriateness of the conceptual model?

If yes, please provide information on the following aspects:

– Who are these experts?
– What is your justification for considering them experts?
– To what extent do they agree that the conceptual model is appropriate?

If no, please indicate why not.

Aspects to judge include appropriateness to represent the underlying clinical process/disease (disease stages, physiological processes, etc.); and appropriateness for economic evaluation (comparators, perspective, costs covered, etc.).

A2/ Cross validity testing (conceptual model): Has this model been compared to other conceptual models found in the literature or clinical textbooks?

If yes, please indicate where this comparison is reported.

If no, please indicate why not.

Part B: techniques used to validate the data serving as input in the model

Please indicate where the description and justification of the following aspects are given:

– search strategy;
– data sources, including descriptive statistics;
– reasons for inclusion of these data sources;
– reasons for exclusion of other available data sources;
– assumptions that have been made to assign values to parameters for which no data was available;
– distributions and parameters to represent uncertainty;
– data adjustments: mathematical transformations (e.g., logarithms, squares); treatment of outliers; treatment of missing data; data synthesis (indirect treatment comparison, network meta-analysis); calibration; etc.

B1/ Face validity testing (input data): Have experts been asked to judge the

appropriateness of the input data?

If yes, please provide information on the following aspects:

– Who are these experts?
– What is your justification for considering them experts?
– To what extent do they agree that appropriate data has been used?

If no, please indicate why not.

Aspects to judge may include but are not limited to potential for bias; generalizability to the target population; availability of alternative data sources; any adjustments made to the data.

B2/ Model fit testing: When input parameters are based on regression models, have

statistical tests been performed?

– If yes, please indicate where the description, the justification and the outcomes of these tests are reported.
– If no, please indicate why not.

Examples of regression models include but are not limited to: disease progression based on survival curves; risk profiles using regression analysis on a cohort; local cost estimates based on multi-level models; metaregression; quality-of-life weights estimated using discrete choice analysis; mapping of disease-specific quality-of-life weights to utility values.

Examples of tests include but are not limited to: comparing model fit parameters (R2, Akaike information criterion (AIC), Bayesian information criterion (BIC)); comparing alternative model specifications (covariates, distributional assumptions); comparing alternative distributions for survival curves (Weibull, lognormal, logit); testing the numerical stability of the outcomes (sufficient number of iterations); testing the convergence of the regression model; visually testing model fit and/or regression residuals.

Part C: techniques used to validate the computerised model

If there are any differences between the conceptual model (Part A) and the final computerised model, please indicate where these differences are reported and justified.

C1/ External review: Has the computerised model been examined by modelling

experts?

If yes, please provide information on the following aspects:

– Who are these experts?
– What is your justification for considering them experts?
– Can these experts be qualified as independent?

– Please indicate where the results of this review are reported, including a discussion of any unresolved issues.

If no, please indicate why not.

C1/ External review: Has the computerised model been examined by modelling

experts?

If yes, please provide information on the following aspects:

– Who are these experts?
– What is your justification for considering them experts?
– Can these experts be qualified as independent?
– Please indicate where the results of this review are reported, including a discussion of any unresolved issues.

If no, please indicate why not.

Aspects to judge may include but are not limited to absence of apparent bugs; logical code structure optimised for speed and accuracy; appropriate translation of the conceptual model.

C2/ Extreme value testing: Has the model been run for specific, extreme sets of

parameter values in order to detect any coding errors?

If yes, please indicate where these tests and their outcomes are reported.

If no, please indicate why not.

Examples include but are not limited to zero and extremely high (background) mortality; extremely beneficial, extremely detrimental, or no treatment effect; zero or extremely high treatment or healthcare costs.

C3/ Testing of traces: Have patients been tracked through the model to determine

whether its logic is correct?

If yes, please indicate where these tests and their outcomes are reported.

If no, please indicate why not.

In cohort models, this would involve listing the number of patients in each disease stage at one, several, or all time points (e.g., Markov traces). In individual patient simulation models, this would involve following several patients throughout their natural disease progression.

C4/ Unit testing: Have individual sub-modules of the computerised model been tested?

If yes, please provide information on the following aspects:

– Was a protocol that describes the tests, criteria, and acceptance norms defined beforehand?
– Please indicate where these tests and their outcomes are reported.

If no, please indicate why not.

Examples include but are not limited to turning sub-modules of the program on and off; altering global parameters; testing messages (e.g., warning against illegal or illogical inputs), drop-down menus, named areas, switches, labelling, formulas and macros; removing redundant elements.

Part D: techniques used to validate the model outcomes.

D1/ Face validity testing (model outcomes): Have experts been asked to judge the

appropriateness of the model outcomes?

If yes, please provide information on the following aspects:

- Who are these experts?
- What is your justification for considering them experts?
- To what extent did they conclude that the model outcomes are reasonable?

If no, please indicate why not.

Outcomes may include but are not limited to: (quality-adjusted) life years; deaths; hospitalizations; total costs.

D2/ Cross validation testing (model outcomes): Have the model outcomes been compared to the outcomes of other models that address similar problems?

If yes, please provide information on the following aspects:

- Are these comparisons based on published outcomes only, or did you have access to the alternative model?
- Can the differences in outcomes between your model and other models be explained?
- Please indicate where this comparison is reported, including a discussion of the comparability with your model.

If no, please indicate why not.

Other models may include models that describe the same disease, the same intervention, and/or the same population.

D3/ Validation against outcomes using alternative input data: Have the model outcomes been compared to the outcomes obtained when using alternative input data?

If yes, please indicate where these tests and their outcomes are reported.

If no, please indicate why not.

Alternative input data can be obtained by using different literature sources or datasets but can also be constructed by splitting the original data set in two parts, and using one part to calculate the model outcomes and the other part to validate against.

D4/ Validation against empirical data: Have the model outcomes been compared to empirical data?

If yes, please provide information on the following aspects:

– Are these comparisons based on summary statistics, or patient-level datasets?
– Have you been able to explain any difference between the model outcomes and empirical data?
– Please indicate where this comparison is reported.

If no, please indicate why not.

D4.A/ Comparison against the data sources on which the model is based (dependent validation)

D4.B/ Comparison against a data source that was not used to build the model (independent validation)

Part E: Other validation techniques

E1/ Other validation techniques: Have any other validation techniques been performed?

If yes, indicate where the application and outcomes are reported, or else provide a short summary here.

Examples of other validation techniques: structured "walk-throughs" (guiding others through the conceptual model or computerised program step-by-step); naïve benchmarking ("back-of-the-envelope" calculations); heterogeneity tests; double programming (two model developers program components independently and/or the model is programmed in two different software packages to determine if the same results are obtained).

Annex 13. Collecting and summarising experts' opinions

Source: Australia. Guidelines for preparing submissions to the pharmaceutical benefits advisory committee. Pharmaceutical Benefits Advisory Committee. Version 5.0 September 2016.

1) Justifying the use of experts' opinions.

Experts' opinions should be used to make up for the lack of data or to support existing data.

2) Describing the method used to gather and analyse experts' opinions.

The methods used to gather experts' opinions range from questionnaire-based surveys involving a statistical analysis, to the qualitative or quantitative summarising of interviews across a selected panel of experts.

A copy of the questionnaire or scenarios submitted to the experts should be provided.

The results and their variability should be presented and interpreted, with a discussion of the limitations and biases of the method used. The analysis and presentation of qualitative studies and interviews should comply with applicable guidelines (O'Hagan, Buck, & Daneshkhah, 2006) (Grigore, Peters, & Hyde, 2013) (Soares, Bojke, & Dumville, 2011).

The weight of experts' opinions in the conclusions should be clearly set out.

All values based on experts' opinions should be tested in a sensitivity analysis.

Information to be provided	Remarks
Criteria for the selection of the experts	Give preference to a group of experts likely to prescribe the treatment. Opt for an entire group or one that is selected in a random way. The generalisability of the opinions of the members of a scientific committee is difficult to estimate.
Number of experts approached	The number of experts approached should be consistent with the number of prescribers concerned. This information is to be provided.
Number of participating experts	Appraise whether the number and characteristics of non-respondents are likely to reduce the representativeness of the group of experts. This information is to be provided.
Participating experts	The experts cannot remain anonymous.
Disclosure of conflicts of interest	Provide a conflicts of interest declaration for each participant, specifying the nature of any conflict, including financial. Specify, where applicable, the remuneration associated with the participation.
Information given	Provide the technical document or any information given to respondents.
Collection method	Provide full details of the method used to collect the experts' opinions (e.g. individually or as a group) Describe the collection process (interview, telephone, self-complete questionnaire). State whether an iteration technique was used (e.g. Delphi method).

Information to be provided	Remarks
Questions asked	Describe the tool and its development. Provide the questionnaire or interview guide. If a pilot study was conducted, provide the results and describe the modifications it brought about. Discuss potential biases for each of the questions.
Number of responses per question	For each question, state the number of missing data items. Discuss the impact or non-response on the representativeness of the opinions given. This information is to be detailed.
Summarising method	Describe the approach used to summarise the responses: qualitative approach (e.g. majority opinion or Delphi technique) or quantitative approach (e.g. mean, median, mode). Describe the approach used to estimate the variability of opinions: qualitative approach (e.g. range of opinions expressed, extreme opinions, common opinion) or quantitative approach (e.g. min/max, confidence interval, percentile).

Glossary

Discounting: technique used for determining the present value of future costs or benefits.

Therapeutic adherence: degree of patient acceptance of their treatment.

Supplemental analysis: analysis not required by HAS, but which may be conducted to provide additional information (e.g. exploratory analysis of cost-effectiveness in subgroups, quality-of-life study, evaluation of the health of caregivers, PROM, PREM).

Cost-benefit analysis (CBA): type of economic evaluation making it possible to measure the costs and benefits of a health intervention in monetary units in order to determine its net social value.

Cost-consequence analysis (CCA): type of economic evaluation in which costs and consequences are presented separately in a detailed way, without their subsequent aggregation.

Cost-effectiveness analysis (CEA): type of economic evaluation making it possible to measure the incremental cost of a supplementary health unit, expressed in physical units (e.g. life-year gained, clinical event avoided).

Cost-outcome analysis (COA): type of economic evaluation making it possible to measure the incremental cost of a supplementary health unit, expressed in physical units (cost-effectiveness analysis) or utility (cost-utility analysis).

Cost-utility analysis (CUA): type of economic evaluation making it possible to measure the incremental cost of a supplementary health unit, expressed in utility (e.g. quality-adjusted life year).

Reference case analysis: HAS-required analysis in which the base scenario complies with the recommended structural choices. It comprises a base-case analysis, non-exploratory subgroup analyses, as well as the thorough exploration of uncertainty through sensitivity analyses.

Base-case analysis: analysis conducted on the population analysed, based on the middle values of the model parameters, as well as the adopted assumptions and modelling choices.

Subgroup analysis: analysis conducted on a subgroup of the population analysed. Depending on the robustness of the analysis, it may be integrated in the reference case analysis or presented as a supplemental analysis when the results are exploratory.

Sensitivity analysis: analysis aimed at exploring the uncertainty surrounding the result of the base-case analysis. It may be conducted using a deterministic or probabilistic approach.

Scenario analysis: analysis presenting an alternative scenario to the one used for the reference case analysis, based on structural choices which are different to those of the base-case scenario. It comprises a base-case analysis and a thorough exploration of uncertainty through sensitivity analyses.

Analysis of extremes: sensitivity analysis whose objective is to estimate results by combining the most pessimistic parameters, assumptions and model choices (worst-case scenario) or the most optimistic ones (best-case scenario).

Threshold analysis: sensitivity analysis involving the variation of model parameter values in order to determine the values that correspond to a specific result (e.g. modification of the cost-effectiveness frontier, predefined ICER).

Deterministic sensitivity analysis: approach aimed at characterising the influence of one or several parameter(s), an assumption or a model choice on the results of the evaluation by testing plausible variations or modifications determined by the evaluator.

Probabilistic sensitivity analysis: approach aimed at exploring the overall statistical uncertainty generated by the statistical variability of point estimates of model variables (mainly effectiveness, tolerability, utility and cost). The values of the data inputs to the model vary simultaneously according to a stochastic process (e.g. second-order Monte Carlo simulation).

Frame of reference: set of methodological choices made by HAS to conduct a reference case analysis.

Calibration: method used to optimise the parameters of the model, aimed at improving the adjustment of simulations to empirical data.

Conservative choice: when there is a choice between two credible methodological positions, the conservative choice is the one which is favourable to the comparator in terms of cost-effectiveness.

Structural choice of the evaluation: set of methodological choices made by the authors of the evaluation upon its design. They determine the methods used to evaluate health outcomes and costs and to choose the model

Direct comparison: approach comparing point estimates of two or more interventions during the same study.

Indirect comparison: approach comparing point estimates of two or more interventions stemming from different studies, using a common comparator.

Naïve comparison: approach comparing point estimates of two or more interventions, without the use of a bias reduction method.

Acceptability curve (evaluation with a single comparator): representation, on the same graph, of the probability of cost-effectiveness of the product evaluated (i.e. its ICER<λ) as a function of the reference values plotted on the ordinate axis.

Multi-option acceptability curve (evaluation using more than one comparator): representation, on the same graph, of the probability of cost-effectiveness of each comparator (i.e. maximising the net benefit) as a function of the reference values plotted on the ordinate axis.

Markov curves: graphic representation of the change in the distribution of the cohort simulated in each model status according to the duration of the simulation.

Opportunity cost: value of the best non-implemented alternative, estimated through the measurement of the benefits foregone due to the allocation of available resources to a given use.

Direct costs: valuation of the resources consumed by the health intervention (e.g. acquisition costs, administrative costs, adverse event handling costs) and by the care provided (e.g. follow-up care, comorbidity-related care, care provided by an informal caregiver, concomitant treatments, end-of-life care).

Indirect costs: valuation of resources not directly consumed by the care provided (see direct costs) but rendered unavailable due to the patient's poor health state or premature death.

Surrogate endpoint: observed variable correlated with the variable of interest, which cannot be observed.

Strict dominance: situation in which a health intervention is less costly than its comparator for an identical or higher effectiveness level, or situation in which an intervention is more effective than its comparator for an identical or lower cost. The comparator is said to be strictly dominated.

Extended dominance: situation in which a combination of two health interventions is less costly than the comparator for an identical or higher effectiveness level, or more effective than the comparator for an identical or lower cost. The comparator is said to be under extended dominance.

Simulation duration: duration over which the model simulates results in terms of costs and health outcomes.

Comparative effectiveness: estimation of the effectiveness gained in comparison with the conventional treatment.

Technical cost-effectiveness: a health intervention is technically efficient if it is not dominated (strict or extended dominance). Technically efficient interventions make up the cost-effectiveness frontier.

Allocative cost-effectiveness: allocative cost-effectiveness characterises interventions that support the optimal allocation of collective resources, by maximising individual health benefits in a constrained budget.

Economic evaluation: type of analysis that examines the difference in health outcomes provided by an intervention in relation to the difference in cost it generates.

Cost-effectiveness evaluation: type of economic evaluation that makes it possible to identify the cost-effectiveness frontier and estimate the ICER or INB of the interventions studied.

Intercurrent event: event that occurs during a cycle, but which does not generate a transition to another model status. The consequence of the occurrence of an intercurrent event is generally limited to an impact on costs and utility.

Transitional event: event that occurs during a cycle and which generates a transition from one model status to another.

Cost-effectiveness frontier: The interventions on the cost-effectiveness frontier are identified as all non-dominated interventions (strict or extended dominance).

Time horizon: period of time during which the costs and health effects associated with the intervention will be taken into account in the evaluation.

Lifetime horizon: period of time running until death. A lifetime model produces costs and health outcomes until all of the individuals in the simulated cohort are in the "deceased" status.

Specific time horizon: period of time determined by the evaluator upon the design of the economic evaluation, defining a shorter period for the stoppage of the evaluation than the patients' death.

Mapping: technique aimed at developing and using algorithms to transpose health measurements (clinical indicator, symptoms score, PROM) onto a utility scale.

Pairwise meta-analyses: Statistical analysis that combines data stemming from several studies comparing the same interventions in parallel, and which generates an overall estimate of their relative effect.

Network meta-analysis: Statistical analysis that combines data stemming from several studies to estimate the relative effectiveness of different health interventions, whether or not they have been compared in a pairwise analysis.

Deterministic model: Model whose parameters are constants.

Stochastic model: Model whose parameters are random variables characterised by a probability distribution.

Observance: patients' effective conduct with respect to their treatment. This generally reflects the patient's ability to comply with all prescription specifications (dosage, intake frequency, intake schedule, etc.).

Therapeutic option: any curative, preventive or palliative measure that may be envisaged in a care pathway for a specific population. The available options may consist of clinical practice guidelines, current practice, marketing authorisations, etc.

Persistence: patient's ability to follow the treatment over a defined period.

Piggy-back study: economic evaluation attached to a clinical trial.

Overall care: actions implemented to meet all of the patient's requirements, for their physical, mental and social wellbeing.

Population analysed: population over which cost-effectiveness is claimed. Group of individuals whose health is affected by the interventions being compared, either directly (sick persons, individuals screened, etc.) or indirectly (caregivers, non- vaccinated people, etc.).

Simulated population: the individuals simulated in the model, whose characteristics correspond to the patients included in the clinical trials or observational studies from which the data is derived.

Patient Reported Outcome Measure (PROM): Measure of the patient's health status reported directly by the patient, without any interpretation by the doctor or other third party (general health status, quality of life, functional status, symptoms, etc.).

Patient Reported Experience Measure (PREM): Patient-reported measure of the care experience (overall satisfaction, information received, attention paid to pain, waiting times, relations with care providers, etc.).

Utility scores: measure of the health state preference on a scale where 1 indicates perfect health and 0 indicates death.

Supportive care: all care and support required by sick persons, in parallel with any specific treatments, throughout serious illnesses.

Healthcare strategy: set of complementary healthcare and support actions implemented for the prevention or treatment of a given pathology in a patient.

Face validity: quality criterion of a model based on current available knowledge, evaluating the credibility of the structure, input data, assumptions made, and simulations produced.

Internal validity: quality criterion of a model based on the technical verification of the model and the consistency of the simulations with observations used for the development of the model.

Cross-validity: quality criterion of a model relating to the consistency of the simulations made with simulations stemming from other models, implemented in conditions that are similar enough to be compared.

External validity: quality criterion of a model based on the consistency of the simulations with empirical observations not used in the development of the model.

Technical verification: quality criterion of a model based on tests aimed at ensuring the accuracy of the coding and mathematical calculations used for the simulations.

Bibliographic references

Ades, A., Caldwell, D., & Reken, S. (2012, January). NICE DSU TECHNICAL SUPPORT DOCUMENT 7: EVIDENCE SYNTHESIS OF TREATMENT EFFICACY INDECISION MAKING: A REVIEWER'S CHECKLIST. Récupéré sur nicedsu.org.uk: http://nicedsu.org.uk/wp-content/uploads/2016/03/TSD7-reviewer-checklist.final_.08.05.12.pdf

Agence nationale d'accréditation et d'évaluation en santé. (2000). Guide d'analyse de la littérature et gradation des recommandations. Paris: ANAES.

Ali, S., & Ronaldson, S. (2012). Ordinal preference elicitation methods in halth economics and health services research: using discrete choice experiments and ranking methods. Br Med Bull, 103(1), 21-44.

Andrade, L., Ludwig, K., & Goni, J. (2020). A French Value Set for the EQ-5D-5L. PharmacoEconomics, https://doi.org/10.1007/s40273-019-00876-4.

Angelis, A., & Kanavos, P. (2016). Value-based Assessment of new medical technologies: towards a robust methodological framework for the application of multiple criteria decesion analysis in the context of health technology assessment. PharmacoEconomics, 34, 435-446.

Baker, R., Bateman, I., & Donaldson, C. (2010). Weighting and valuing quality-adjusted life-years using stated preference methods: preliminary results from the social value of a QALY project. Health Technology Assessment, 14(27), 1-162.

Brazier, J., Connell, J., & Papaioannou, D. (2014). A systematic review, psychometric analysis and qualitative assessment of Generic Preference-Based Measures of Health in Mental Health Populations and the estimation of mapping functions from widely used specific measures. NIHR Journals Library.

Brazier, J., Harper, R., & Thomas, K. (1998). Deriving a preference based single index measure from the SF-36. J. Clinical Epidemiology., 51(11), 1115-1129.

Brazier, J., Ratcliffe, J., Salomon, J., & Tsuchiya, A. (2017). Measuring and valuing health benefits for economic evaluation. Oxford: Oxford University Press.

Brazier, J., Roberts, J., & Plattas, M. (2005). Estimating a preference-based index for a menopause specific health quality of life questionnaire. Health and Quality of Life Outcomes, 13.

Brazier, J., Rowen, D., & Yang, Y. (2012). Comparison of health state utility values derived using time trade-off, rank and discret choice data anchored on the full health-dead scale. Eur. J. Health Econ., 13(5), 575-87.

Brazier, J., Yang, Y., & Tsuchiya, A. (2008). A review of studies mapping (or cross walking) from non-prefrence based measures of health to generic preference-based measures. Récupéré sur eprint.whiterose.ac.uk: http://eprint.whiterose.ac.uk/10912/

Brennan, A., Chick, S., & Davies, R. (2006). A taxonomy of models structures for economic evaluation of health technologies. Health Economics, 15, 1295-1310.

Briggs, A., Claxton, K., & Sculpher, M. (2008). Decision modeling for health economic evaluation. Oxford University Press.

Briggs, A., Weinstein, M., Fenwick, E., & et al. (2012). Model parameter estimation and uncertainty: a report od the ISPOR-SMDM Modeling good practices Task Force 6. Value Health, 15, 835-842.

Caro, J., & Briggs, A. (2012). Modelling good research practices task force. Value in Health, 15, 796-803.

Chevalier, J., & de Pouvourville, G. (2013). Valuing EQ-5D using time trade-off in France. Eur J Health Econ, 14(1), 57-66.

Collett, D. (2015). Modelling Survival Data in Medical Research. Third edition. CRC Press.

Commissariat général à la stratégie et à la prospective. (2013). Evaluation socioéconomique des investissements publics. Rapport de la mission présidée par E. Quinet.

Cooper, N., Sutton, A., & Achana, F. (2015). RFP Topic: use of network meta-analysis to inform clinical parameters in economic evaluations. Report for the CADTH.

Dakin, H. (2013). Review of studies mapping from quality of life or clinical measures to EQ-5D: an online database. Health and Quality of Life Outcomes, 11, 151.

Dias, S., Sutton, A., & Welton NJ. (2013). Evidence synthesis for decision making 6: embedding evidence synthesis in probabilistic cost-effectiveness analysis. Med Decis Making, 33(5), 671-8.

Dias, S., Welton, N., & Sutton, A. (2013). Evidence synthesis for decision making. 5 : the baseline naturl history model. Medical Decision Making., 33(5), 657-670.

Drummond, M., & al. (2015). Methods for the Economic Evaluation of Health Care Programmes. Oxford: Oxford University Press.

Drummond, M., Sculpher, M., & Torrance, G. (2005). Methods for Economic Evaluation of Health Care prorammes. Oxford: Oxford University Press.

Eddy, D., Hollingworth, W., & Caro, J. (2012). Model transparency and validation: A report of the ISPOR-

SMDM Modeling Good Research Practices Task Force-7. Value in Health, 15, 843-850.

EuroQol. (s.d.). Crosswalk index value calculator. Récupéré sur www.euroqol.org: https://euroqol.org/eq-5d-instruments/eq-5d-5l-about/valuation-standard-value-sets/crosswalk-index-value-calculator/

Faria, R., Hernandez Alava, M., & Manca, A. (2015, May). NICE DSU Technical Support Document 17. The use of observational data to inform estimates of treatment effectiveness in technology appraisal : methods for comparative individual patient data. Récupéré sur scharr.dept.shef.ec.uk: http://scharr.dept.shef.ac.uk/nicedsu/wp-content/uploads/sites/7/2016/03/TSD17-DSU-Observational-data-FINAL.pdf

Garau, M., Shah, K., & Towse, A. (2009.). Assessment and appraisal of oncology medicines: does NICE's approach include all relevant element? What can be learnt from international HTA experience? Report for the Pharmaceutical Oncology Initiative (POI). .

Gedda, M. (2015). Traduction française des lignes directrices PRISMA pour l'écriture et la lecture des revues systématiques et des méta-analyses. Kinesither Rev, 15(157), 39-44.

Ghabri, S., Cleemput, I., & Josselin, J. (2017). Towards a New Framework for Addressing Structural Uncertainty in Health Technology Assessment Guidelines. PharmacoEconomics.

Ghabri, S., Hamers, F., & Josselin, J. (2016). Exploring Uncertainty in Economic Evaluations of Drugs and Medical Devices: Lessons from the First Review of Manufacturers' Submissions to the French National Authority for Health. PharmacoEconomics, 34(6).

Glied, S., & Smith, P. (2013). The Oxford Handbook of Health Economics. Oxford.

Gold, M., Siegel, J., Russel, L., & Weinstein, M. (1996). Cost-effectiveness in Health and Medicine. New-York Oxford: Oxford University Press.

Guerre, P., Hayes, N., & Bertaux, A. (2018). Estimation du coût hospitalier : aproches par "micro-costing" et "gross-costing". Revue d'épidémiologie et de Santé Publique, 66, 565-572.

Haute Autorité de Santé. (2013, avril). Etat des lieux: niveau de preuve et gradation des recommandations de bonne pratique. Consulté le juin 13, 2017, sur www.has-sante.fr: https://www.has-sante.fr/portail/jcms/c_1600564/fr/niveau-de-preuve-et-gradation-des-recommandations-de-bonne-pratique-etat-des-lieux

Haute Autorité de Santé. (2013). Niveau de preuve et gradation des recommandations de bonne pratique. Paris: HAS.

Haute Autorité de Santé. (2016). Format du rapport technique. Document support pour la rédaction des rapports techniques déposés auprès de la CEESP. Récupéré sur www.has-sante.fr.

Haute Autorité de Santé. (2016). Format du rapport technique. Document support pour la rédaction des rapports techniques déposés auprès de la CEESP. Récupéré sur www.has-sante.fr: https://www.has-sante.fr/portail/jcms/c_2639155/fr/format-du-rapport-technique

Haute Autorité de Santé. (2018, Juillet). Rapport d'analyse prospective 2018 : De nouveuax choix pour soigner mieux. Récupéré sur www.has-sante.fr: https://www.has-sante.fr/portail/upload/docs/application/pdf/2018-07/rapport_analyse_prospective_2018.pdf

HERC. (2016, 05). HERC database of mapping studies. Consulté le 10 2016, sur https://www.herc.ox.ac.uk/downloads/herc-database-of-mapping-studies

Higgins, J., Thomas J., Chandler, J., & et al. (2019). Cochrane Handbook for Systematic Reviews of Interventions version 6.0 (Updated July 2019). Récupéré sur www.training.cochrane.org: https://training.cochrane.org/handbook#about-the-handbook

Hutton, B., Salanti, G., Caldwell, D., & et al. (2015). The PRISMA extension statement for reporting of systematic reviews incorporing network meta-analyses of health care interventions: checklist and explanations. Ann Intern Med, 162(11), 777-84.

ICH. (2019, mars). Guide d'introduction à MedDRA Version 22.0. Récupéré sur www.meddra.org: https://www.meddra.org/sites/default/files/guidance/file/intguide_22_0_french.pdf

Jansen, J., Trikalinos, T., Cappelleri, J., & et al. (2014). 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 Health, 17(2), 157-73.

Johannesson, M., & al. (1996). Theory and Method of Economic Evaluation of Health Care. Dordrecht, Bostin and London: Kluwer.

Josselin, J., & Le Maux, B. (2017). Multi-criteria Decision Analysis. Dans Statistical tools for program evauation (pp. 385-416). Springer.

Latimer, N. (2011). NICE DSU Technical Support Document 14: Undertaking survival analysis foreconomic evaluations alongside clinical trials - extrapolation with patient-level data. Récupéré sur http://www.nicedsu.org.uk

Latimer, N. (2013). Survival analysis for economic evaluations alongside clinical trials extrapolation with patient-

level data : inconsistencies, limitations and a practical guide. Med Decis Mak, 33(6), 743-54.

Le Gales, C., Buron, C., & Costet, N. (2002). development of a preference-weighted health status classification system in France: the Health Utilities Index. Health Care Management Science, 5(1), 41-52.

Lebègue, D. (2005). Le prix du temps et la décision publique. Commissariat général au plan. Paris: La documentation française.

Liberati, A., Altman, D., Tetzlaff, J., & et al. (2009). The PRISMA Statement for Reporting Systematic Reviews and Meta-Analyses of Studies That Evaluate Health Care interventions: Explanation and Elaboration. PLoS Med, 6(7), e1000100.

Longworth, L., & Rowen, D. (2011). The use of mapping methods to estimate health state utility values. NICE DSU Technical Support Document 10.

Longworth, L., & Rowen, D. (2013). Mapping to obtain EQ-5D utility values for use in NICE Health Technology Appraisals. Value in Health, 16(1), 202-210.

Lorgelly, P., Doble, B., & Rowen, D. (2014). Condition specific or Generic Preference-based Measures in Oncoloy: the EORTC-8D or the EQ-5D. University of Sheffield.

Marsh, K., Ijzerman, M., & Thokala, P. (2016). Multiple criteria decision analysis for health care decision making - Emerging good practices: report2 of the ISPOR MCDA Emerging Good Practices Task Force. Value in Health, 19, 125-137.

Mlika-Cabanne, N., Harbour, R., & de Beer, H. (2011). Sharing hard labour: developing a standard template for data summaries in guideline development. BMJ Qual saf, 20, 141-145.

Mogyorosy, Z., & Smith, P. (2005). The main methodological issues in costing health care services. A literature review. (C. f. Economics, Éd.) Récupéré sur www.york.ac.uk: http://www.york.ac.uk/media/che/documents/papers/researchpapers/rp7_Methodological_issues_in_costing_health_care_services.pdf

Moher, D., Shamseer, L., & Clarke, M. (2015). Prefered Reporting Items for Systematic Review and Meta-Analysis Protocols (PRISMA-P) 2015 Statement. Syst Rev, 4(1), 1.

Morelle, M., Plantier, M., & Dervaux, B. (2018). Méthodes d'analyse et de traitement des données de coûts. RESP, 66, S101-118.

Muhlbacher, A., & Kaczynski, A. (2016). Making good decisions in healthcare with multi-criteria decision analysis: the use, current research and future developement of MCDA. Appl Health Econ Health Policy, 14, 29-40.

Mulhern, B. (2014). Using generic preference-based measures in mental health: psychometric validity of the EQ-5D and SF-6D. The British Journal of Psychiatry, 205, 236–243.

Mullins, C. (2006). Double counting and the reporting of cost per event avoided. Clin Ther, 28(4), 602-3.

Nocera, S., Telser, H., & Bonato, D. (2003).), The Contingent-Valuation Method in Health Care – An Economic Evaluation of Alzheimer's Disease. Dordrecht, Boston, and London: Kluwer.

Nuijten, M. (1998). Reporting format for economic evaluation. PartII: focus on modelling studies. Pharmacoeconomics, 14(3), 259-268.

O'Brien, B., & Briggs, A. (2002). Analysis of uncertainty in health care cost-effectiveness studies: an introduction to statistical issues and methods. Statistical methods in medical research, 11, 455-468.

Petrou, S., Kwon, J., & Madan, J. (2018). A Practical Guide to Conducting a Systematic Review and Meta-analysis of Health State Utility Values. Récupéré sur link.springer.com: https://link.springer.com/content/pdf/10.1007%2Fs40273-018-0670-1.pdf

Phillippo, D., Ades, A., & Dias, S. (2016, december). NICE DSU Technical Support Document 18: Methods for populaiton-adjusted indirect comparisons in submissions to NICE. Récupéré sur scharr.dept.shef.ac.uk: http://scharr.dept.shef.ac.uk/nicedsu/wp-content/uploads/sites/7/2017/05/Population-adjustment-TSD-FINAL.pdf

Pickard, A., Neary, M., & Cella, D. (2007). Estimation of minimally important differnces in EQ-5D utility and VAS scores in cancer. Health Qual Life Outcomes, 5, 70.

Puhan, M., Schünemann, H., & Murad, M. (2014). A GRADE Working Group approach for rating the qualitéy of treatment effect estimates from network meta-analysis. BMJ, 349.

Raimond, V., Midy, F., & Thébaut, C. (2016). L'évaluation économique des produits de santé innovants : quelle interprétation pour quel usage ? Revue Française des Affaires Sociales, 263-281.

Rowen, D., Brazier, J., & Roberts J. (2009). Mapping SF-36 onto the EQ-5D index: how reliable is the relationship ? health Qual Life Outcomes, 7, 27-38.

Rowen, D., Brazier, J., & Van Hout, B. (2015). A comparison of methods for converting DCE values onto the full health-dead QALY scale. med Decis Making, 35(3), 328-40.

Soekhai, V., de Bekker-Grob, E., & Ellis, A. (2019). Discrete choice experiments in health economics: past, present and future. pharmacoEconomics, 37(2), 201-226.

Song, F., Loke, Y., & Walsh, T. (2009). methodological problems in the use of indirect comparisons for evaluation healthcare interventions: survey of published systematic reviews. British Medical Journal, 338, b1147.

Stinnett, A., & Mullahy, J. (1998). Net health benefits: a new framework for the analysis of uncertainty in cost-effectiveness analysis. Med. Decis. Making, S68-S80.

Strobe - Strengthening the reporting of observational studies in epidemiology. (2007). Checklist of items that should be included in reports of cohort studies. Récupéré sur strobe-statement.org: https://www.strobe-statement.org/index.php?id=available-checklists

Swindle, R., Lukas, C., & Meyer, D. (1999). Cost analysis in the Department of Veterans Affairs: consensus and future directions. Med Care, 4(suppl Va), AS3-8.

Thokala, P., Devlin, N., & Marsh, K. (2016). Multiple criteria decision analysis for health care decision making - An introduction: report 1 of the ISPOR MCDA Emerging Good Practices Task Force. Value in Health, 19, 1-13.

Torrance, G., Feeny, D., & Furlong, W. (2001). Visual analog scales. Do they have a role in the measurement of preferences for health states? Med Decis making, 21, 329-34.

van Hout, B., & Janssen, M. (2012). Interim scoring for the EQ-5D-5L: Mapping the EQ-5D-5L to EQ-5D-3L value sets. Value in Health, 15(5), 708-15.

Wailoo, A., Hernandez, M., & Manca, A. (2016, 09). Use of Mapping to Estimate Utility Values from Non-Preference Based Outcome Measures for Cost per QALY Economic Analysis: Good Research Practices Task Force. Consulté le 10 2016, sur www.ispor.org

Walters, S., & Brazier, J. (2005). Comparison of the minimally important difference for two health state utility measures: EQ-5D and SF-6D. Qual Life res, 14(6), pp. 1523-1532.

Weinstein, M. (2003). principles of good practice fordecision analytic modeling in health care evaluation: report of the ISPOR tast force on good research practices. Modeling studies. Value in health, 6(1), 9-17.

Weinstein, M., Torrance, G., & McGuire, A. (2009). QALYs: the basics. Value in health, 12 Suppl 1, S5-S9.

Wolowacz, S., Briggs, A., & Belozeroff, V. (2016). Estimating Health-State Utility for economic models in clinical studies: an ISPOR Good Research Practices Task Force report. Value in Health, 19(5), 704-719.

Woods, B., Sideris, E., & Palmer, S. (2017). NICE DSU Technical SupportDocument 19. Partitioned Survival Analysis for Decision Modelling in Health Care: A Critical Review. Récupéré sur http://www.nicedsu.org.uk

Wordsworth, S., Ludbrook, A., & Caskey, F. (2005). Collecting unit cost data in multicentre studies. Creating comparable mehtods. Eur J Health Econ, 6(1), 38-44.

Zawodnik, A., & Niewada, M. (2018). Multiple criteria decesion analysis (MCDA) for health care decision making - overview og guidelines. Récupéré sur www.jhpor.com: https://www.jhpor.com/article/2207-multi-ple-criteria-decision-analysis-mcda-for-health-care-decision-mak-ing---overview-of-guidelines

Zorzela, L., Loke, Y., & Loannidis, J. (2016). PRISMA harms checklist: improving harms reporting in systematic reviews. BMJ, 1(352), i157.

Methodological guidance documents referred to

Germany. IQWIG. General Methods for the assessment of the relation of benefits to costs. Version 1.0, November 2009.

Germany. German recommendations on health economic evaluation: third and updated version of the Hanover Consensus. Value in Health 2008; 11(4): 539-544

England and Wales. NICE. Guide to the methods of technology appraisal 2013. http://publications.nice.org.uk/pmg9.

Australia. Guidelines for preparing submissions to the pharmaceutical Benefits Advisory Committee. Pharmaceutical Benefits Advisory Committee. Version 5.0 September 2016.

Austria. Guidelines on Health Economic Evaluation. Institute for Pharma-economic Research. April 2006.

Belgium. Cleemput I, Neyt M., Van de Sande S., Thiry N. Belgian guidelines for economic evaluations and budget impact analyses: second edition. 2012. KCE reports 183B. www.kce.fgov.be.

Canada. Guidelines for the Economic Evaluation of Health Technologies: Canada. Canadian Agency for Drugs and Technologies in Health. 4th edition, 2017. https://www.cadth.ca/about-cadth/how-we-do-it/methods-and-guidelines/guidelines-for-the-economic-evaluation-of-health-technologies-canada.

Canada. PAAB. Guidance Document for the Use of Pharmacoeconomic Studies in Advertising. January 2013

Scotland. NHS Scotland. Guidance to manufacturers for completion of New Product Assessment Form (NPAF). Revised March 2016

Finland. Preparing a Health Economic Evaluation to be Attached to the Application for Reimbursement Status and Wholesale Price for a Medicinal Product. 2015.

France. Guide méthodologique pour l'évaluation économique des stratégies de santé (Methodological guidance for the economic evaluation of health strategies). Collège des économistes de la santé (Health Economists Board). 2003.

Ireland. Health information and quality authority. Guidelines for the Economic Evaluation of Health Technologies in Ireland. Draft 2014

Norway. Norwegian Medicines Agency. Guidelines on how to conduct pharmacoeconomic analyses. 2012.

New Zealand. Prescription for pharmacoeconomic analysis. Methods for cost-utility analysis. Pharmaceutical Management Agency. Version 2.2 2015.

Netherlands. Zorginstituut Nederland. Guidelines for economic evaluations in health care. June 2016.

Sweden. General guidelines for economic evaluation. Pharmaceutical Benefit Board, April 2003.

USA. Sanders G., Neumann P.J., Basu A. et al. (2016) Recommendations for conduct, methodological practices, and reporting of cost-effectiveness analyses. Second panel on cost-effectiveness in health and medicine. JAMA 2016; 316(10): 1093-1103.

USA. ICER (2018) A Guide to ICER's Methods for Health Technology Assessment. https://icer-review.org/wp-content/uploads/2018/08/ICER-HTA-Guide_082018.pdf

Participants

This document was produced by the project managers of the Public Health and Economic Evaluation Department, under the direction of Catherine Rumeau-Pichon, Michèle Morin-Surroca and Carine Ferreti, in collaboration with the Public Health and Economic Evaluation Commission chaired by Christian Saout.

The project was coordinated by Fabienne Midy.

The original French text for the part entitled "Structural Choices in Evaluation Methods" was written by Fabienne Midy.

The original French texts for the part entitled "Choice of Methods for the Evaluation of Outcomes" were written by: Fabienne Midy (General Principles section); Agathe Doutriaux and Sandy Leproust (Effectiveness and Tolerability section); Erwan Autin, Agathe Doutriaux, Fabienne Midy, Cléa Sambuc (Utility section).

The original French text for the part entitled "Methodological Choices for the Evaluation of Costs" was written by Isabelle Bongiovanni and Anne-Line Couillerot.

The original French text for the part entitled "Methodological Choices Concerning Modelling" was written by Sylvain Druais and Fabienne Midy.

The original French text for the part entitled "Presentation and Interpretation of Results" was written by Anne-Line Couillerot, Fabienne Midy, Cléa Sambuc.

We wish to thank the CEESP members who acted as resource persons for certain parts of the document: Pauline Chauvin (Modelling), Valérie Clément (Utility Evaluation), Pierre-Yves Boelle (Clinical Data), Lionel Perrier (Cost Evaluation).

We also wish to thank all of those who took part in the public consultation and auditions; their questions and comments greatly contributed to the enhancement of the methodological guidance document.

Participating SEESP Project Managers, HAS

Erwan Autin

Isabelle Bongiovanni

Anne-Line Couillerot

Agnès Dessaigne

Agathe Doutriaux

Sylvain Druais

Stéphanie Leclerc

Sandy Leproust

Fabienne Midy

Célia Pessel

Anne-Isabelle Poullié

Véronique Raimond

Cléa Sambuc

Laura Zanetti

CEESP Members

ADAM Christophe

ALLA François

BIDEAU Daniel

BUNGENER Martine

BUTHION Valérie
CAMBON Linda
CHAUVIN Pauline
CLEMENT Valérie
COLIN Cyrille
COTE-MESNIER Sophie
DUGUET Christophe
DUPONT Jean-Claude K
FOURNEAU Christelle
FOURNEYRON Emmanuelle
FOURNIER Cécile
IZAMBERT Caroline
LACOSTE Olivier
LAZZAROTTO Sébastien
LE GALES Catherine
MAMZER Marie-France
ORVAIN Jacques
PERRIER Lionel
PICHETTI Sylvain
RUSCH Emmanuel
SERRIER Hassan
SAOUT Christian
VINAY Nicolas
WITTWER Jérôme

Acknowledgements

HAS would like to thank all the participants listed above.

Abbreviations and acronyms

CBA	Cost-benefit analysis
CCA	Cost-consequence analysis
CEA	Cost-effectiveness analysis
COA	Cost-outcome analysis
CUA	Cost-utility analysis
BIA	Budget impact analysis
LTI	Long-term illness (chronic condition)
MA	Marketing authorisation
APE	*Actif à Part Entière* – Conventional healthcare professional
NMB	Net monetary benefit
NB	Net benefit
INB	Incremental net benefit
NHB	Net health benefit
CCAM	Joint classification of medical procedures
CEESP	*Commission de l'évaluation économique et de santé publique* (Economic Evaluation and Public Health Commission)
NCS	National cost study
HPG	Homogeneous Patient Group
HAH	Hospitalisation at home
HAS	Haute Autorité de Santé – the French Health Authority
MAIC	Matching-adjusted indirect comparison
MSO	Medicine-Surgery-Obstetrics
NGAP	*Nomenclature générale des actes professionnels* – General Nomenclature of Medical Procedures
PMSI	Medical IT Programme
PROM	Patient Reported Outcome Measure
QALY	Quality-Adjusted Life Year
ICER	Incremental Cost-Effectiveness Ratio
SG	Standard Gamble
FCR	Follow-on care and rehabilitation
STC	Simulated treatment comparison
TNB	*Table nationale de biologie* (national medical laboratory table)
TTC	*Toutes taxes comprises* (inclusive of VAT)
TTO	Time Trade-Off

VAS Visual Analogue Scale

Find all our publications at

www.has-sante.fr

HAS
HAUTE AUTORITÉ DE SANTÉ

Developing quality in the health,
social, and medico-social sector