

# Procedures and methods of benefit assessments for medicines in Germany

Geertruida E. Bekkering · Jos Kleijnen

© The Author(s) 2008. This article is published with open access at Springerlink.com

## Contents

<b>Abstract</b> . . . . .	S6
<b>1. Background</b> . . . . .	S8
<b>2. Legal framework</b> . . . . .	S9
<b>3. Requirements of the assessment process</b> . . . . .	S10
3.1 Transparency of procedure . . . . .	S10
3.1.1 Topic identification and prioritisation. . . . .	S10
3.1.2 Tendering and awarding the commission . . . . .	S11
3.1.3 Defining the research question . . . . .	S11
3.1.4 Developing the protocol . . . . .	S11
3.1.5 Assessment of the evidence . . . . .	S11
3.1.6 Publication of the preliminary report . . . . .	S12
3.1.7 Review of the preliminary report. . . . .	S12
3.1.8 Recommendations and final report. . . . .	S12
3.1.9 Appeals and planning updates to the report . . . . .	S12
3.1.10 Conclusion . . . . .	S13
3.2 Active participation of the affected parties . . . . .	S13
3.2.1 Topic identification and prioritisation. . . . .	S13
3.2.2 Defining the research question(s). . . . .	S13
3.2.3 Participation in other phases of the assessment . . . . .	S14
3.3 Conclusion . . . . .	S15
<b>4. Requirements of assessment methods</b> . . . . .	S15
4.1 Evidence-based medicine . . . . .	S15
4.1.1 Best available evidence . . . . .	S15
4.1.2 Internal versus external validity. . . . .	S16
4.1.3 Role of systematic reviews in EBM. . . . .	S16
4.1.4 Assessment of the evidence for each outcome . . . . .	S17
4.1.5 Evidence-based methods to prevent bias in the assessment. . . . .	S17
4.2 Choice of comparators . . . . .	S17
4.3 Benefit . . . . .	S18
4.4 Conclusion . . . . .	S18
<b>5. International standards/methods used by NICE and IQWiG</b> . . . . .	S18
5.1 Methods of NICE and IQWiG . . . . .	S19
5.2 Assessment process used by NICE and IQWiG . . . . .	S19
5.3 Conclusion . . . . .	S20

<b>6. How should a benefit assessment be implemented?</b> . . . . .	S20
6.1 Topic identification and prioritisation . . . . .	S20
6.2 Tendering and awarding the commission. . . . .	S22
6.3 Defining the research question . . . . .	S22
6.4 Developing the protocol . . . . .	S25
6.5 Assessment of the evidence . . . . .	S25
6.5.1 Search strategy . . . . .	S25
6.5.2 Submission by manufacturers . . . . .	S25
6.5.3 Selection of studies . . . . .	S26
6.5.4 Data-extraction process . . . . .	S26
6.5.5 Data synthesis. . . . .	S26
6.5.6 Subgroup analyses. . . . .	S26
6.6 Publication of the preliminary report . . . . .	S26
6.7 Review of preliminary report. . . . .	S27
6.8 Recommendations and final report . . . . .	S27
6.9 Planning update report and appeal decision . . . . .	S27
6.10 Conclusion. . . . .	S27
<b>References.</b> . . . . .	S28

**Abstract** The Federal Joint Committee (FJC; *Gemeinsamer Bundesausschuss*, G-BA) defines the health-care elements that are to be reimbursed by sickness funds. To define a directive, the FJC can commission benefit assessments, which provide an overview of the scientific evidence regarding the efficacy and benefits of an intervention. This paper describes the operational implementation of the legal requirements with regard to the benefit assessments of medicines. Such benefit assessments are sometimes referred to as “isolated benefit assessments,” to distinguish them from benefit assessments as part of a full economic evaluation.

The FJC has the freedom to commission these assessments from any agency; however, to date the majority have commissioned the Institute for Quality and Efficiency in Health Care (IQWiG). Nevertheless, the content of this paper applies integrally to any institute commissioned for

such assessments. In this report, ‘the institute’ is used when the text refers to any of these institutes.

The legal framework for benefit assessments is laid out in the German Social Code Book version V (<http://www.sozialgesetzbuch.de>), Sects. 35b (§ 1), 139a (§ 4–6) and Sect. 139b (§ 3). It is specified that:

- The institute must guarantee high transparency.
- The institute must provide appropriate participation of relevant parties for the commission-related development of assessments, and opportunity for comment on all important segments of the assessment procedure.
- The institute has to report on the progress and results of the work at regular intervals.
- The institute is held to giving the commission to external experts.

Based on the legal framework, the institute must guarantee a high procedural transparency. Transparency of the whole process should be achieved, which is evidenced by clear reporting of procedures and criteria in all phases undertaken in the benefit assessment. The most important means of enhancing transparency are:

1. To implement a scoping process to support the development of the research question.
2. To separate the work of the external experts performing the evidence assessment from that of the institute formulating recommendations. Therefore, the preliminary report as produced by external experts needs to be public, and published separately from any subsequent amendments or (draft-)reports made by the institute, which includes the institute’s recommendations.
3. To implement open peer review by publishing both the comments of the reviewers and their names.

---

The expertises were commissioned and funded by the German Association of Research-based Pharmaceutical Companies, Berlin, Germany (<http://www.vfa.de/en/articles/index-en.html>). The authors G. E. Bekkering, Jos Kleijnen had full editorial freedom.

---

G. E. Bekkering (✉)  
Department of General Practice, Catholic University of Leuven,  
Kapucijnenvoer 33, Leuven 3000, Belgium  
e-mail: Trudy.Bekkering@med.kuleuven.be

G. E. Bekkering  
BeSyRe Bekkering Systematic Reviews, Geel, Belgium

J. Kleijnen  
Kleijnen Systematic Reviews Ltd.,  
Westminster Business Center, 10 Great North Way,  
Nether Poppleton, York, UK  
e-mail: jos@systematic-reviews.com

Based on the legal framework, the institute must provide for adequate participation of relevant parties. These include organisations representing the interests of patients; experts of medical, pharmaceutical and health economic science and practice; the professional organisations of pharmacists and pharmaceutical companies; and experts on alternative therapies. Patients and health care professionals bring in new insights with respect to research priorities, treatment and outcomes.

The relevant parties should be identified and contacted whenever the global scope of the assessment has been drafted. Subsequently, the relevant parties should be involved in defining the research question, developing the protocol and commenting on the preliminary report. To implement the involvement of relevant parties in defining the research question a scoping process is suggested. For the other phases, written comments followed by an oral discussion should be used. Finally, the relevant parties should have the right to appeal the final decision on judicial grounds. None of these steps mean that the institute would lose any part of its scientific independence.

From the relevant sections of the legal framework with respect to the assessment methods, it can be concluded that:

1. The institute must ensure that the assessment is made in accordance with internationally recognised standards of evidence-based medicine (EBM).
2. The assessment is conducted in comparison with other medicines and treatment forms under consideration of the additional therapeutic benefit for the patients.
3. The minimum criteria for assessing patient benefit are improvements in the state of health, shortening the duration of illness, extension of the duration of life, reduction of side effects and improvements in quality of life.

EBM refers to the application of the best available evidence to answer a research question, which can inform questions about the care of patients. The optimal design, even for effectiveness questions, is not always the randomised, controlled trial (RCT) but depends on the research question and the outcomes of interest. To increase transparency for each question, the levels of evidence examined should be made explicit. There is no empirical evidence to support the use of cutoff points with respect to the number of studies before making recommendations. To get the best available evidence for the research question(s), all relevant evidence should be considered for each question, and the best available evidence should be used to answer the question. Separate levels of evidence may have to be used for each outcome.

There are many ways in which bias can be introduced in systematic reviews. Some types of bias can be

prevented, other types can only be reported and, for some, the influence of the bias can be investigated. Reviews must show that potential sources of bias have been dealt with adequately.

Methods used by other agencies that perform benefit assessments are useful to interpret the term ‘international standards’ to which the institute must comply. The National Institute for Health and Clinical Excellence (NICE) is a good example in this respect. NICE shows that it is possible to have transparent procedures for benefit assessments but that this requires detailed documentation. NICE has implemented an open procedure with respect to the comments of reviewers, which makes the procedure transparent. Although the Institute for Quality and Efficiency in Health Care (IQWiG) in Germany invites comments on their protocol and preliminary report by posting them on their website, and comments are made public, the individual comments are not evaluated openly, and therefore it remains uncertain whether or not they lead to changes in the reports. The participation of relevant parties in the assessment process as implemented by NICE guarantees a process that is transparent to all relevant parties.

Transparency of the whole process is assured by clear reporting of procedures and criteria in all phases undertaken in the benefit assessment. In a scoping process, a draft scope is commented on first in writing and subsequently in the form of a scoping workshop. In this way, all relevant aspects can be heard and included in the final scope. The protocol is then developed, followed by evidence assessment. The methods used should be completely reported to show readers that the assessment has been performed with scientific rigour and that bias has been prevented where possible. All relevant parties should have the opportunity to comment on the draft protocol and the draft preliminary report. Each comment should be evaluated as to whether or not it will lead to changes, and both the comments and the evaluation should be made public to ensure transparency of this process. The same procedure should be used for the peer-review phase. Based on the final report of the evidence assessment, the institute forms recommendations and the FJC appraises the evidence.

During the writing of the final report, a separation between the evidence assessment and the evidence-appraisal phase should be implemented. Ideally, this separation should be legally enforced to prevent any confusion about conflict of interests.

Such a process guarantees a feasible combination of the legal requirements for transparency and involvement of relevant parties with international standards of EBM to ensure that the benefit assessments of medicines in Germany are performed according to the highest standards.

**Keywords** Benefit assessments · Evidence-based medicine

**JEL Classification** I11

## 1. Background

In Germany, health care is regulated via sickness funds; for the broad public, about 90% of citizens are covered by social sickness funds. The health-care elements to be reimbursed by social sickness funds are defined by the Federal Joint Committee (FJC; *Gemeinsamer Bundesausschuss*, G-BA). The FJC was founded through the Statutory Health Insurance (SHI) Modernisation Act in 2004 and replaces the former four parallel-existing committees for Physicians and Health insurances, Dentists and Health insurances, Coordinating Committee (*Koordinierungsausschuss*) and Committee for Hospitals. The FJC is supervised by the Federal Ministry of Health (FMH).

The tasks of the FJC are defined by the German Social Code Book V (<http://www.sozialgesetzbuch.de>) and are specified by its code of procedure [1]. The main task of the FJC is to formulate which costs of health-care elements, including medicines, are to be reimbursed by the sickness funds. To define a directive, the FJC can commission a scientific institute to carry out benefit assessments or economic evaluations. These benefit assessments are sometimes referred to as “isolated benefit assessments,” to distinguish them from benefit assessments as part of a full economic evaluation. Such assessments provide an overview of the scientific evidence regarding the benefits<sup>1</sup> of a medicine. Based on the recommendations of the commissioned institute, directives are formulated by the FJC regarding whether or not to reimburse for the medicine or technology in question. Assessments can be commissioned for any new licensed medicine (with patented active medicinal ingredients) or for any “medicine of relevance”. Decisions regarding what to commission are made by the FJC itself based on the work of internal working groups. Medicines to be considered are defined by the members of the FJC, but criteria are unclear. Currently, the decision to commission is based on cost implication or any aspect of epidemiologic relevance; however, we were unable to find more concrete criteria.

A benefit assessment of medicines evaluates the clinical benefits and harms of a medicine and follows the design of a systematic review. The steps of a benefit assessment are displayed in Box 1. Since 1 April 2007, the legislature has also provided the FJC with the option of requesting a full economic evaluation.<sup>2</sup> A full economic evaluation should always be preceded by a benefit assessment. However, there is an important principal difference between benefit assessments<sup>3</sup> that will be followed by an economic evaluation and benefit assessments without such an evaluation. Benefit assessments that are intended to be followed by full economic evaluations need to have a broader scope with respect to ‘benefit’. International standards need to be applied for economical evaluations (see below).

### Box 1 Steps in a typical benefit assessment

The benefit assessment is divided into the following ten steps

1. Defining the preliminary research question
2. Tendering and awarding the commission
3. Defining the scope
4. Developing the protocol (in IQWiG terms: report plan)
5. Assessment of the evidence: in this phase literature is searched, critically appraised, and analysed
6. Publication of the preliminary report
7. Review of the preliminary report
8. Publication of the final report
9. Submission to the commissioning entity
10. Appeal and planning update of the report

*IQWiG* Institute for Quality and Efficiency in Health Care

For both benefit assessment as well as full economic evaluation, concrete specifications are stipulated in the legislature regarding the material criteria to be applied and the procedural requirements. The most appropriate way to carry out an economic evaluation following international standards of health economy has been presented in detail by von der Schulenburg et al. [4].

This paper describes the operational implementation of the legal requirements with regard to the benefit assessment of medicines.

According to its code of procedure, the FJC has the freedom to commission any agency to perform a benefit assessment (<http://www.sozialgesetzbuch.de>, Sect. 38, part

<sup>1</sup> Those outcomes of adopting a given course of action that do not involve the use of resources. They can relate to changes in clients’ health and well being, and also to the psychological and physical benefits derived by people, other than the client, who are affected by substance misuse (families/friends of the client, victims of crime, etc) [2].

<sup>2</sup> Studies in which a comparison of two or more treatments or care alternatives is undertaken, and in which both the costs and outcomes of the alternatives are examined [3].

<sup>3</sup> We will use the term ‘benefit assessments’ when such an assessment is performed with the intention to conduct a full economic evaluation and ‘isolated benefit assessment’ when no such intention exists.

F, clauses 1 and 2). However, up to now the majority of commissions have been given to the Institute for Quality and Efficiency in Health Care (IQWiG). The IQWiG was established with the Statutory Health Insurance (SHI) Modernisation Act of 2004. Among other tasks, the Institute is supposed to conduct a benefit assessment of medicines [see Sects. 35b § 1 and 139a § 3 no 5 of the German Social Code Book V (old version, i.e. in use until March 2007)]. The content of this paper applies to all institutes that can be commissioned to perform such assessments. For the remainder of this document, the term ‘the institute’ is used to refer to any such institute.

This paper consists of three parts. Firstly, methodological requirements for the institute’s methods based on the legal framework will be discussed. Secondly, methods used abroad to perform benefit assessments will be described—these methods form the international standards with which the institute must comply based on the legal framework. Finally, the last part outlines the recommended procedures and methods in detail, based on previous sections.

## 2. Legal framework

In this section the legal mandate for the institute’s methods and procedures will be outlined and commented on. Consequences of the legal mandate will be described separately for the assessment process and the assessment methods.

The legal framework for benefit assessments is laid out in the German Social Code Book version V (<http://www.sozialgesetzbuch.de>). The relevant sections of the legal framework are:

Section 35b § 1: “(1) 1. Pursuant to Sect. 139b § 1 and 2, the Institute for Quality and Efficiency in Health Care can be commissioned to assess the benefit or the cost-benefit ratio of pharmaceuticals. 2. Assessments according to clause 1 can be made for each pharmaceutical with patented active ingredients that has become eligible for prescription for the first time, as well as for other pharmaceuticals of significance. 3. The assessment is made based on a comparison with other pharmaceuticals and therapy forms in consideration of the additional therapeutic benefit for the patients in proportion to the costs.<sup>4</sup> 4. With regard to patient benefit, especially the improvement of the state of health, a reduction in the duration of illness, an extension of the duration of life, a reduction of side effects and an improvement in the quality of life should be taken into account

<sup>4</sup> From an economic perspective this wording is incorrect; it should be ‘additional costs’ instead of ‘costs’.

appropriately, as should the suitability and reasonableness of cost absorption by the community of insured people<sup>5</sup> when making an economic assessment. 5. The Institute makes commission-related decisions on the methods and criteria for the development of assessments pursuant to clause 1 based on the international standards of evidence-based medicine and health economics acknowledged by the respective expert circles. 6. During the commission-related development of methods and criteria and the generation of assessments, the Institute ensures high procedural transparency and appropriate participation of the parties mentioned in Sect. 35 § 2 and Sect. 139a § 5. 7. The Institute shall publish the respective methods and criteria on the Internet. 8. Clauses 3 through 7 shall also apply to benefit assessments that have already been started.”

Section 35b, § 2: “(2) 1. The assessments according to Sect. 1 are fed to the Federal Joint Committee as a recommendation for decision-making according to Sect. 92, clause 1, § 2, No. 6. 2. They are to be checked at suitable intervals and, if necessary, to be adapted. 3. If new scientific evidence is available, the assessment is to be reviewed at the request of the manufacturers.

Section 139a § 4: “(4) 1. The Institute must ensure that the assessment of the medical benefit is made based on internationally acknowledged standards of evidence-based medicine and that the economic assessment is made based on the relevant internationally recognised standards, especially of health economics. 2. At regular intervals, the Institute must publicly report on the work processes and results including its basis for decision-making. (5) 1. In all important segments of the assessment procedure, the Institute must provide an opportunity for comment to the experts of medical, pharmaceutical and health economic science and practice, to pharmaceutical companies and the relevant organisations representing the interests of patients and the self-help organisations for chronically ill and disabled people, as well as the Federal Government Commissioner for Patients’ Affairs. 2. The comments must be included in the decision. 6. To ensure the professional independence of the Institute, the employees must—prior

<sup>5</sup> This refers to an evaluation from the perspective of social sickness funds. It should be noted, however, that most international guidelines require a societal or national economic perspective for such evaluations [4].

to being hired—disclose all relationships to interest associations and commissioning institutes, especially those of the pharmaceutical and medical products industry, including the type and amount of financial allocations.”

Section 139b § 3: “1. In order to fulfil its tasks according to Sect. 139a § 3 clauses 1–5, the Institute must commission scientific projects to external experts. 2. These experts must disclose all relationships to associations and contract organisations, particularly in the pharmaceutical industry and the medical devices industries, including details on the type and amount of possible remuneration received.”

### 3. Requirements of the assessment process

From the content of the legal framework outlined above it can be concluded that the IQWiG, among other institutes, can be commissioned to assess the benefits of medicines. The legal mandate lays out a number of specific requirements about the process of such assessments:

1. The institute must guarantee a high transparency. Procedures, methods and criteria should be published on the Internet. In the section on transparency of procedure below, we will argue that a high transparency should be applied in all phases of the assessment, starting from topic identification and prioritisation.
2. The institute must provide for appropriate participation of relevant parties for the commission-related development of assessments, and opportunity for comment in all important segments of the assessment procedure. Appropriate participation refers to the possibility to contribute to all important stages of the process. The comments must be included in the documentation. The discussion on active participation of affected parties below explains when relevant parties should participate (process), while “How should a benefit assessment be implemented?” describes how such parties should best participate (methods). Relevant parties should include at least:
  - Relevant organisations representing the interests of patients
  - Experts in medical, pharmaceutical and health economic science and practice
  - The professional organisations of pharmacists and pharmaceutical companies and experts on alternative therapies

These groups are hereafter referred to as ‘relevant parties’. Participation by such relevant parties also creates

some obligations, for example making all relevant information for benefit assessment available.

3. The institute has to report on the progress and results of the work at regular intervals. This requirement is closely related to the requirement for transparency. Below we will argue that, for reasons of transparency, the work of external experts needs to be published separately from any subsequent amendments or (draft-)reports made by the institute.
4. The institute is held to giving the commission to external experts. The external experts, but also employees of the institute, must declare any potential conflicts of interest.
5. The assessment should be updated at regular intervals. If new evidence is available, the assessment is reviewed at the request of the manufacturers.

#### 3.1 Transparency of procedure

Pursuant to the requirements of Sect. 35b § 1 clause 6 of the German Social Code Book V, the institute must guarantee high procedural transparency and participation. Transparency is a basic prerequisite for any research, as this is the only way to show that the process has been performed with scientific rigour and that bias has been prevented as much as possible. Transparency is essential for commissioned institutes to be able to show that the results obtained can be considered valid.

Transparency of the whole process should be achieved by clear reporting of the procedures and criteria used in all phases undertaken in the benefit assessment. The section below describes the crucial points in a benefit assessment where transparency is needed in order to fulfil this requirement.

##### 3.1.1 Topic identification and prioritisation

We argue that, as part of a transparent assessment procedure, the process leading to this assessment, i.e. topic identification and prioritisation, should also be clear. Although the final topic selection may be influenced by political pressures, the actual process should be as transparent as possible. Topic identification should be open to the public and therefore also to all relevant parties. For this to be the case, the FMH and the FJC should establish a procedure for topic identification and prioritisation that involves the public. This could be achieved by publication of the criteria used to select potential topics to be commissioned and the criteria used for prioritisation, should there be multiple suitable topics. A comparable procedure is already being implemented at the German Institute of Medical Documentation and Information (DIMDI), and an

evaluation of these experiences might yield useful information [5].

### 3.1.2 *Tendering and awarding the commission*

To fulfil legislative requirements, the institute must assign scientific research commissions to external experts. However, legislation does not require that external experts are commissioned for all assessments, and in some cases the institute may decide to commission an assessment to an internal expert group, e.g. IQWiG staff. This does not change the operational implementation of the legal requirements with regard to benefit assessment; however, to ensure transparency, it should be clear in which circumstances external experts are required. These criteria should be made public.

To recruit external experts, public and international tenders are preferred, as these ensure that an unlimited number of experts can offer their services. The selection of experts should be made based on reproducible and objective criteria. This phase could be made transparent by making the procedure for the tender, as well as the criteria used to select experts, public, for example by posting them on the institute's website.

### 3.1.3 *Defining the research question*

The first step of an assessment is to define the research question. This is one of the most important phases of a study, as poorly focussed questions lead to unclear decisions about what research to include and how to summarise it [6]. All relevant parties should be involved in this phase, as requested in Sect. 35b § 1 of the German Social Code Book V. Such a procedure ensures that all important aspects are heard and thus that all perspectives are taken into account in formulating the research question. A scoping process is proposed for this, as described in detail in the section on defining the research question below. A scoping process aims to set the boundaries of the assessment with regard to the four elements of the PICO system: patient population, intervention, comparison intervention and outcomes. The scoping process is intended to speed up the total assessment, providing that sufficient input from relevant parties has been incorporated.

In order to make this phase transparent, it should be clear which parties have contributed to the research question and which suggestions were, and which were not, taken into account. Therefore, all suggestions from the parties involved in commenting on the draft research question should be made public, together with documentation as to whether or not a suggestion was considered relevant and has been incorporated into the research question. Also transcripts of meetings that were

held to discuss suggestions and comments should be made public. Beforehand, arrangements must be made to ensure an option for confidentiality concerning patient privacy.

### 3.1.4 *Developing the protocol*

Based on the research question, a protocol (in IQWiG terms: "draft report plan") is subsequently written. To enhance transparency, PICO criteria are used to focus the research question(s):

- Patient population—what patients the assessment refers to
- Intervention—what medicine is evaluated (including dosing instructions and methods of use)
- Comparison intervention—which is/are the current standard treatment(s), including the rationale for choosing this or these treatment(s) as standard
- Outcomes—what outcomes are important and patient-specific

The protocol further contains the following items: background information leading to the research question(s), search strategy, study selection criteria and procedures, study quality assessment, data-extraction strategy and method of synthesis of extracted data.

The draft protocol should be open for comment. It should be commented on by all relevant parties and an oral hearing should subsequently be held. To enhance transparency, the draft protocol, all comments from relevant parties and the evaluation of these comments should be posted on the Internet. Transcripts of meetings held to discuss all relevant comments should also be made public.

### 3.1.5 *Assessment of the evidence*

The assessment of the evidence encompasses all steps of a systematic review of the evidence, from the literature search to evidence synthesis. Although all steps of the process are described in detail in the protocol, during the assessment it is likely that new decisions will have to be made or that previous decisions will seem unfeasible. Furthermore, subjective judgments in this process are inevitable. A transparent approach will ensure readers that a rigorous approach has been taken and that bias has been prevented as much as possible. Transparency in this phase is ensured if the reader is able to follow all the steps taken in the process. Any changes that have been made to the protocol should be reported and made public as an amendment and should be subject to comment and discussion in an oral hearing. More details on how the process of evidence assessment should take place are given below.

### 3.1.6 Publication of the preliminary report

The preliminary report describes the process and the results of the assessment. To enhance transparency, the report should be constructed according to the recommendations of the guidelines for the reporting of systematic reviews as much as possible (QUOROM guidelines; <http://www.equator-network.org>) [7].

To present the evidence clearly, evidence tables should be used. A standard way of constructing evidence tables has not been identified, mainly because this depends on the research question [8]. However, all results and characteristics of the included studies that may have influenced the results or which are relevant for the generalisability of results should be presented in a way that enables easy comparison between studies.

To ensure transparency, the report as written by external experts should be made public. This also applies when the assessment is performed internally by staff of the commissioned institute.

### 3.1.7 Review of the preliminary report

The entire assessment process provides multiple opportunities for comments by the relevant parties. All comments and the evaluation of each individual comment with regard to whether or not it is relevant and has led to any changes should be made public. Such a procedure, as already implemented by the National Institute for Health and Clinical Excellence (NICE) [9], will enhance transparency of the review procedure.

In addition, an external review of the preliminary report takes place. A similar peer review process has become established for scientific publications. This phase is important to verify the work of the review group. The peer review process works best if done transparently. To ensure transparency, it should first be clear how the institute invites peer reviewers. The process of inviting reviewers, and the criteria for selecting them and for performing the review should be made public. Second, peer reviewers' comments should be documented and evaluated, similar to the procedure for comments of relevant parties. The result of the evaluation should be documented and published as part of the final report. Whenever a comment is considered not relevant this should be justified. This is the only way to ensure that the comments of the reviewers have been taken into account or have been omitted for the right reasons.

To improve the quality of the system of peer review, two large medical journals examined the effects of an open peer review system, meaning that the names of the reviewers were revealed to the authors [10, 11]. Both studies showed that open reviewing is not detrimental to the quality of such reviews; thus there is no reason to omit the reviewer's

name. Arguments that favour open peer review include increased accountability, fairness and transparency.

### 3.1.8 Recommendations and final report

Following publication of the preliminary report of the evidence assessment, recommendations are made by the institute. These are presented in the final report, which therefore includes more than an evidence assessment.

Legally, the appraisal based on this report is the task of the FJC. However, the recommendations of the institute also include a form of appraisal. Thus, the institute performs an assessment as well as an appraisal and, because of this, a potential conflict of interest arises. This should be prevented by separating the two steps, similar to the way this is implemented by NICE [9]. Box 2 below presents NICE's definitions of assessment and appraisal.

**Box 2** Difference between evidence assessment and evidence appraisal according to the National Institute for Health and Clinical Excellence (NICE) [12]

---

“The *assessment* process consists of an objective analysis of the quality, findings and implications of the (mainly research) evidence available as it relates to the appraisal question and context. The *appraisal* process, in contrast, is a consideration of the outputs of the assessment process within the context of additional information supplied by relevant parties such as clinical specialists and patient experts. The appraisal decision is a judgment on the importance of a range of factors that differ from appraisal to appraisal.”

---

Means of increasing transparency in this phase are the implementation of an open procedure that shows the results of the evidence assessment (as stated in the preliminary report) separate from the recommendations of the final report and separate from the results of the appraisal, as well as a clear link between each recommendation and the evidence on which it is based. However, ideally, a separation between evidence assessment and appraisal should be implemented, similar to the procedure used by NICE.

### 3.1.9 Appeals and planning updates to the report

Pursuant to the Social Code Book V Sect. 35b, clauses 2 and 3, the assessments are to be checked at suitable intervals and, if necessary, adapted. Furthermore, if new scientific evidence becomes available, the assessment is to be reviewed at the request of the manufacturers, who should make all information about benefit research available; with this in mind, a registry of ongoing trials would be useful. Transparency in this step is ensured if the processes for both appeal and update are made public using objective criteria.



### 3.1.10 Conclusion

Based on the legal framework, the institute must guarantee a high procedural transparency. Transparency of the whole process should be achieved by clear reporting of procedures and criteria in all phases undertaken in the benefit assessment. The most important means of enhancing transparency are:

1. Implementation of a scoping process to support development of the research question.
2. Publication of the comments of all parties involved, together with a justification as to whether or not these comments led to changes in the final documents.
3. Separation of the product of the evidence assessment from that of the evidence appraisal by publishing the work of external experts.
4. Implementation of open peer review by publishing both the comments of the reviewers and their names.

### 3.2 Active participation of the affected parties

The institute is legally required to provide for appropriate participation of relevant parties for the commission-related development of assessments. Appropriate participation refers to the possibility to contribute to all important stages of the process. This includes specification of the problem up to and including the appraisal of evidence at the FJC. This section focusses on the process (*when* relevant parties should participate), whereas the methods (*how* relevant parties should participate) are described in detail later (“[How should benefit assessments be implemented?](#)”).

#### 3.2.1 Topic identification and prioritisation

Topic identification should be open to the public and therefore also to the relevant parties. A public procedure should be implemented. An example of collaboration between clinicians and patients is implemented by the Database of Uncertainties about the Effects of Treatments (DUETs) [13], where research questions about the effects of treatments are included in the database only if they have been requested by both patients and clinicians.

#### 3.2.2 Defining the research question(s)

To assist in defining the research question(s) with the active participation of relevant parties, a scoping process is recommended. The purpose of the scoping process is to provide a framework for the assessment. Issues of interest, for example population, intervention and outcomes, should be defined as clearly as possible. Such a scoping process consists of two steps. First, a draft scope is prepared and

sent to all relevant parties, who are requested to comment in writing. Subsequently, a scoping workshop is held in which all opinions are discussed. Such a method ensures that all relevant aspects are heard and included in the final scope. This should lead to research questions that are relevant to all parties that have participated, whilst still guaranteeing the scientific independence of the institute.

The institute should organise the scoping workshop, in which the FJC, the institute, the external experts and relevant parties should take part. The participation of the FJC should ensure that the research question(s) will be directly transferable to the clinical question of the FJC, and thus that the results of the assessment meet the needs of the FJC. Alternatively, a one-staged scoping process organised by the FJC could be implemented.

The research institute in question is commissioned by the FJC. A potential commission is based on a policy question for which a decision is needed, for example, what are the best treatment options for patients with brain tumours? Subsequently, the research question is refined using the PICO criteria, for example, with respect to the participants, do we refer to primary tumours or secondary tumours, do we refer to children or adults? How should the intervention be defined: radiotherapy, chemotherapy and/or surgery? Do we include proton radiation? How many proton therapy facilities are available in Germany? Which useful and ethical comparison intervention would be available: e.g. photon radiation, watchful waiting, sham treatment? Concerning outcomes, do we focus on mortality or are we interested in morbidity, quality of life and adverse events? The minimum set of outcomes has been defined in Social Code Book V.

In addition, the types of studies that are relevant to answering the question should be specified. However, the study designs chosen should not be used as a means to exclude other designs from the assessment, as for each question the best evidence available at the time of the assessment should be sought. This issue will be discussed below in the section on [Evidence-based medicine](#). The resulting question should be relevant for the German health care system and should also cover the problem for which a decision is needed. This is an important stage of the project and should be given appropriate attention. All potentially relevant parties should be identified and contacted at this stage.

Subsequently, a preliminary subject draft is formulated, which then can be better specified with input from a scoping group. Such a group includes all relevant parties specific for the commission. Consultation of relevant parties ensures that all relevant aspects have been considered, and such a group therefore needs careful consideration. The group should include representatives of patient organisations and care givers. Experts from medical, pharmaceutical and

health economic science and practice should also be involved. For example, in an assessment with respect to diabetes, one would need not only clinicians with a strong clinical background but also clinicians with a scientific background. In addition, there are many health-care professionals who deal with patients with diabetes, such as nurses, physiotherapists, home-care assistants, dieticians, and wound care and preventive services professionals.

At least those participants who are defined by law, e.g. professional organisations of pharmacists and pharmaceutical companies, and patients, have to be involved. Furthermore, depending on the topic, the scoping group may be expanded to include additional circles, such as patients' relatives or experts on alternative therapies. This leads to a greater acceptance of the assessment results and to an indispensable collection of expertise in the assessment process.

The involvement of stakeholders in research is generally considered to be important [14, 15]. Some empirical evidence for this comes from studies that focussed on the input of patients or consumers, which include past, current and future users, care givers and people representing any of these groups. Broad involvement is believed to lead to research that is more relevant to people's needs and concerns, more reliable and more likely to be put into practice [16]. Based on a survey, Hanley et al. [17] conclude that consumer involvement in the design and conduct of controlled trials is growing, and seems to be welcomed by most researchers. Another argument for the involvement of relevant parties refers to the subjective judgments. Subjective judgment takes place in each assessment. A scoping workshop may balance the subjective judgments (value judgments) from one group to another.

Empirical evidence shows that individuals' biases may be better balanced in multidisciplinary groups. For example, when presented with the same evidence, a single-specialty group will reach different conclusions than a multidisciplinary group; a multidisciplinary panel may provide a more divergent viewpoint than panels composed entirely of practitioners who apply the interventions [18]. Coulter et al. [19] showed that the composition of a panel influences the ratings, and those who use a given procedure are more likely to rate it as appropriate than those who do not use the procedure.

NICE considers the contribution of health professionals to be unique. These outline the professional view of the place of a technology in current clinical practice [12]. Clinicians, therefore, would be able to provide evidence on issues such as:

- Patient group variations, in particular differential baseline risk of the condition and capacity for different subgroups of patients to benefit;

- The particular circumstances in which treatment is delivered, including the need for concomitant treatments, the setting in which treatment is delivered and the requirements for additional professional input;
- The treatments that are currently used as standard practice and whether these may differ from what is considered best practice.

There is empirical evidence showing that the preferences of patients and health professionals differ with respect to research priorities, treatment and outcomes. For example, a review revealed a number of mismatches in priorities for health research between professionals and the public [20]. Devereaux et al. [21] showed considerable variability between physicians and patients in weighing up the potential outcomes associated with atrial fibrillation and its treatment. In addition, there was considerable variability within the group of patients and within the group of physicians [21]. Differences in the preferences of patients and health professionals are difficult to predict; they vary in direction and magnitude and are often specific for a given condition [22]. Chard et al. [23] showed a mismatch between the views of professionals and patients on the management of osteoarthritis. And, based on input from patients, fatigue has been added as a core outcome in the evaluation of interventions in rheumatoid arthritis [24]. Other studies also show a lack of consensus among key stakeholders, among them patients, family members and health care professionals on desired outcome priorities for adolescents' mental health services [25], schizophrenia [26] and rheumatology [27]. Such results indicate that the input of patients and health professionals is especially relevant in the phase of formulating the research question.

### 3.2.3 Participation in other phases of the assessment

Participation should occur at all important steps of the assessment procedure. Apart from the above-mentioned scoping process to develop the research question, this includes in particular:

- Written comments and oral discussion for the protocol
- Written comments and oral discussion for the preliminary report
- Appeal for the final decision of the FJC

The form of participation will be outlined below (see section heading "[How should a benefit assessment be implemented?](#)").

Empirical evidence supporting the importance of such participation in research planning and design is beginning to emerge. A systematic review showed that the main rationale for involving patients affected by cancer in research, policy and planning, and practice was the unique

perspective that patients can bring to research. However, the impact of patient involvement has been only sparsely investigated [28]. In a short-action research pilot study, consumers were involved in all stages of the health technology assessment (HTA). Consumers made unique contributions to the HTA Programme; when seeking research topics, face-to-face discussion with a consumer group was more productive than scanning consumer research reports or contacting health information services. Consumers were willing and able to play active roles as panel members in refining and prioritising topics and in commenting on research plans and reports [29]. Several case series describe how input from patients led to changes in the methods, procedures and measures used in the design of a randomised controlled trial (RCT) on, for example, breast cancer [30] or stroke [31].

The experiences of NICE indicate that it is possible to involve stakeholders in health-care decisions, although it demands commitment from the entire organisation and specific managerial arrangements; depending on the circumstances, it can also be costly [32].

Although the participation of stakeholders, i.e. both patients and individuals from all relevant health professional groups, is one of the domains of the AGREE instrument that is used internationally to assess the quality and reporting of clinical guidelines [33], empirical evidence on the importance of the involvement of health-care professionals in clinical research is sparse. The Consumers' Advisory Group for Clinical Trials (CAG-CT) is a good example of the value of such input, however. This group consists of both patients with breast cancer and breast cancer health professionals. Marsden et al. [34] described how this group adequately contributed to the design of a national randomised breast cancer trial.

### 3.3 Conclusion

Based on the legal framework, adequate participation of relevant parties at all steps of the processes must be provided for. There is empirical evidence showing that patients and health-care professionals have their own preferences with respect to research priorities, treatment and outcomes.

The relevant parties should be identified and contacted whenever the global scope of the assessment is available. Subsequently, the relevant parties should be involved in defining the research question, developing the protocol and commenting on the preliminary report. In all phases, written comments followed by an oral discussion should be used. However, for the research question a more explorative scoping workshop should be implemented, in which the FJC, the institute and the relevant parties, as well as the external experts participate. Finally, the relevant parties should have the right to appeal the final decision.

## 4. Requirements of assessment methods

From the relevant sections of the legal framework with respect to the assessment methods, it can be concluded that

- The institute must ensure that the assessment of the medical benefit is made in accordance with internationally recognised standards of evidence-based medicine (EBM).
- The benefit assessment is conducted in comparison with other medicines and/or treatment forms under consideration of the additional therapeutic benefit for the patients. This requires the definition of current treatment standard(s) with which a (new) intervention is compared. Co-interventions which are widely used should be allowed.
- The minimum catalogue of criteria for assessing patient benefit, as given by law, are
  - Improvement of the state of health
  - Shortened duration of illness
  - Extension of the duration of life
  - Reduction of side effects
  - Improvement in the quality of life

In the sections below, we will argue that, in principle, each separate outcome refers to a separate research question, and that to find the best available evidence a separate consideration of appropriate study types for each question is warranted.

### 4.1 Evidence-based medicine

The institute must ensure that the benefit assessment is made in accordance with recognised standards of EBM. EBM (or better, evidence-based health care) represents the integration of best research evidence with clinical expertise and patient values in making decisions about the care of patients [35]. The content of EBM is timely, as this refers to the best evidence currently available, also called 'best available evidence'. Therefore, there is a difference between optimal evidence (for example an RCT could be conducted) and best available evidence (for example, when no RCTs have been performed for a certain outcome and thus best evidence refers to cohort studies). Best available evidence implies that a hierarchy of evidence levels exists.

#### 4.1.1 Best available evidence

From a scientific point of view, the strongest design for evaluating the efficacy of therapeutic interventions is an RCT [36]. The basic principle of RCTs is that the patient is allocated randomly to one of the chosen number or types of interventions to compare their effects. Allocating the

patients at random is the best guarantee that the patient groups are comparable because it avoids selection bias. A limitation of RCTs is that they often use strict inclusion criteria and consequently exclude large proportions of the target population. Moreover, patients who choose not to participate in the trial may differ from those who do. This implies that the effects documented in studies may not be representative of the effects that would be seen if the interventions were used on the entire target population. In addition, RCTs will see only what they look for, and such designs often have a limited follow-up period. For example, it is unlikely for RCTs to report on rare side effects or on long-lasting effects. Therefore, an RCT is not always the only and optimal design for questions about the effects of health care. This depends, for example, on the outcomes that are to be assessed.

#### 4.1.2 Internal versus external validity

The extent to which bias is minimised in a clinical trial is referred to as internal validity. Internal validity is defined as the extent to which the results of a study are correct for the circumstances being studied [37]. External validity, in contrast, refers to the extent to which the results of a study provide a correct basis for generalisations to other circumstances [37], for example, other patient populations and interventions including different comparator technologies. The design of RCTs is typically characterised by high internal validity, sometimes at the expense of applicability. This is labelled ‘efficacy’, referring to the effects of an intervention under ideal circumstances. The design of observational studies, in contrast, may have higher external validity at the expense of internal validity. Observational studies may evaluate practice more pragmatically in the clinical setting, which is labelled ‘effectiveness’. Both study designs may contribute to the ‘best available evidence’, and the limitations of each design should be taken into account when formulating recommendations.

Based on the principle of ‘best available evidence’, it is not possible to conclude that there is too little evidence to perform a benefit assessment. The commissioning of a benefit assessment is driven by a clinical question for which the FJC needs to make a decision. Such a decision can be taken, in principle, on the lowest level of evidence (expert opinion).

If no RCTs have been carried out, or if the RCTs do not report (valid) information for the outcome in question, results of other studies should be assessed. For this reason the strength of the analysed evidence should always be presented together with the recommendation.

There is no widely agreed evidence to support methods that propose the use of cut-off points with regard to a minimum number of studies before making

recommendations. In contrast, empirical evidence suggests that sometimes one trial may be enough while at other times many studies may still give too little evidence. Three studies compared results of meta-analyses with those of trials [38–40]. A systematic comparison of these three empirical assessments concluded that the disagreements may be less prominent for primary outcomes and that overall the frequency of significant disagreements beyond chance is 10–25% [41, 42]. Disagreement may also exist among trials [43] and among meta-analyses [44]. These discrepancies suggest that a dogmatic approach with respect to a minimum number of studies needed is difficult to support. Instead, in accordance with the definition of EBM, the best available evidence should be used to answer the clinical question. All evidence should be scrutinised to determine how well it matches the clinical question with regard to characteristics of patients, interventions and outcome events. In order to determine what evidence is best, factors such as trial size and trial quality need to be examined to evaluate the validity of the study or meta-analysis.

The importance of using evidence from designs other than only RCTs is illustrated by the updated methods of NICE: “Non-RCT evidence will be required, not just for those situations where RCTs are unavailable, but also to supplement information from RCTs when they are available” [45]. This statement is stronger than the previous NICE methods, which allowed that “in some circumstances non-RCT evidence may be needed to supplement what is available from RCTs ....” [12].

In order to account for uncertainty in the available evidence, Claxton et al. [46] propose the use of a new framework for decision-making by NICE. The analysis combines all available data, accounting for the uncertainty explicitly, and establishes which intervention, for a particular group of patients, has the highest expected cost-effectiveness. This approach needs further work.

#### 4.1.3 Role of systematic reviews in EBM

The hierarchy of evidence as introduced by the definition of EBM is typically labelled ‘levels of evidence’, based on internal validity for the purpose of the study. The levels of evidence aim to make the evidence underlying a conclusion transparent. There are several approaches to grading the strength of evidence (e.g. [47, 48]). For effectiveness questions, NICE uses a four-level system to rank the different types of study design according to their relative internal validity for estimating relative treatment effect:

- Level 1: RCTs
- Level 2: controlled observational studies, e.g. cohort studies, case-control studies

- Level 3: observational studies without a control group, e.g. case series
- Level 4: expert opinion [12]

From the perspective of EBM, a systematic review of RCTs is the most powerful and useful evidence available [35]. Therefore, this design should be the highest level of evidence. This is illustrated by the ranking system of the Oxford Centre for Evidence-based Medicine [49], which produced a five-level hierarchy in which systematic reviews of RCTs are the highest level for effectiveness questions of treatment interventions. This scheme also presents levels of evidence for other study types such as prognostic and diagnostic studies. Whenever a systematic review has been found that addresses the research question, it should be examined closely to determine whether a new systematic review is still necessary. The decision to exclude such a review, for example if the quality is very poor, should be justified in the preliminary report.

#### 4.1.4 Assessment of the evidence for each outcome

Since the benefit assessment is based on different patient-relevant outcomes, the research for each of these outcomes must also be conducted based on the principle of “best available evidence”. The commission to assess the benefit of one medicine in one indication typically leads to a number of research questions. As a result, for each research question (or outcome) a separate search may be required. Then, for each question (outcome) the number and designs of studies should be assessed. If there are multiple high-quality RCTs that report on the relevant question, evidence of a lower level such as non-randomised and observational studies may be excluded. However, if there are few, small or poor-quality RCTs only, lower level evidence should be considered as well.

Subsequently, the evidence for each outcome separately is evaluated in a similar manner. This calls for a flexible rather than a dogmatic approach, depending on whether the highest available level of evidence provides answers to the questions. Often, that will become clear only in the course of doing the assessment.

#### 4.1.5 Evidence-based methods to prevent bias in the assessment

Two leading international organisations have developed handbooks on how to perform systematic reviews of RCTs: the Cochrane Collaboration [6] and the Centre for Reviews and Dissemination (CRD) [50]. Both handbooks describe explicit methods for limiting bias in the review and providing more reliable results, both of which are needed to draw valid conclusions in order to make decisions. Below

follows an overview of important limitations in systematic reviews that may lead to bias and how they should be dealt with. For more details on these limitations, as well as other methods for preventing bias in the assessment, we refer to the Cochrane and CRD handbooks [6, 50].

*Publication bias* Studies that show beneficial results are more likely to be published and therefore more likely to be included in systematic reviews; this may introduce bias [51]. Bias in the retrieval of studies can be countered by using an extensive search strategy: searching multiple electronic databases and using multiple sources of study reports (electronic databases, manual search, trial registries, reference lists, etc.). Where possible, statistical methods should be used to examine the evidence for publication bias [6, 50]. Evidence for publication bias weakens the strength of the conclusion from a systematic review.

*Reporting bias* Within studies, beneficial results are more likely to be reported and therefore more likely to be included in systematic reviews [51]. Reporting bias can be prevented to a certain extent by not selecting certain outcomes. Instead, all results on all relevant outcomes should be presented in a systematic review. However, this still depends on the outcomes that are reported in the primary studies. Registration of ongoing trials is the only way to reduce the problem of reporting bias.

*Methodological quality of primary studies* The quality of the primary studies to be included in the review is of the utmost importance. Empirical evidence shows that inadequate quality of studies may distort results from systematic reviews [52]. Therefore, the methodological quality of all studies to be included needs to be assessed. Also, the influence of the quality of included studies on their results should be examined [52]. The use of summary scores from quality scales is problematic [6]. Based on empirical evidence, concealment of allocation, blinding of outcome assessment and handling of patient attrition in the analysis should generally be assessed [52]. It is not always possible to incorporate all methodological quality items in a study; for example, it is not possible to blind patients for psychological interventions.

## 4.2 Choice of comparators

The comparator is either the best possible treatment or the currently routine treatment. Although the best treatment would be the comparator of choice, treatments representing routine German care should also be included in the evaluation. There may be several comparator treatments, depending on regional differences. The comparator needs to be defined as precisely as possible, especially if the

circumstances of its use differ from the circumstances of use for the intervention being assessed. The choice of one (or more) comparator(s) needs to be discussed in the scoping process and justified in the protocol.

As part of the licensing procedure, medicines are typically compared with placebo. Such trials answer the question whether the medicine is more effective than placebo. For benefit assessments, from the perspective of the health-care system, head-to-head trials comparing one medicine with another are to be preferred if the comparison therapy is the current standard therapy. Head-to-head trials should be evaluated in the same way as placebo-controlled trials. If the assignment to both treatments has been done randomly, such trials are level-1 evidence. If only placebo-controlled trials are available, the additional benefit of medicines can be estimated using adjusted indirect comparisons [53, 54].

### 4.3 Benefit

Benefit assessments investigate the benefits (and harms) of a medicine. The legal requirements specify inclusion of the following benefit parameters in the assessment:

- Improvement of the state of health
- Reduction in the duration of illness
- Extension of the duration of life
- Reduction of side effects
- Improvement in the quality of life

It should be noted that this is a minimal list only, and other benefit parameters should be included as necessary. The scoping process should identify relevant other outcomes to be incorporated in the assessment, for example patient satisfaction.

Benefit is a subjective concept, and patients, clinicians and researchers may have differing opinions on how to define benefit. Even within a group of patients and within a group of professionals ‘benefit’ may be differently interpreted. In addition, the topic of assessment will have an impact on what is considered benefit. For each assessment, benefit should therefore be defined in the scoping process. The input from all relevant parties is crucial for this matter. The different types of benefit may require different study designs in the review: For example, benefit defined as lower mortality or lower morbidity may be assessed using RCTs, while observational studies may be better suited to answer the question if benefit is defined as fewer adverse events.

### 4.4 Conclusion

The institute must ensure that the assessment is made in accordance with internationally recognised standards of

EBM. To increase transparency for each question, the levels of evidence that will be used in the assessment should be made explicit. To get the best available evidence for the research question, all evidence should be considered in order to determine which studies form the best evidence to answer that question. The optimal design, even for efficacy questions, is not always the RCT but depends on the research question and the outcomes. Separate strategies may necessary for each outcome.

There are many ways in which bias can be introduced in systematic reviews. Some types of bias can be prevented, other types can only be reported and for others, the influence of the bias can be investigated. Reviews must show that potential sources of bias have been dealt with adequately.

The comparator is either the best treatment or the currently routine treatment. There may be several comparator treatments. This should be discussed during the scoping process.

For each assessment, benefit should be defined in the scoping process. The input from all relevant parties is crucial for this matter.

## 5. International standards/methods used by NICE and IQWiG

The German legislature requires that international methodological standards be applied (Section 35b of the German Social Code Book V). The absence of a supranational organisation defining binding standards does not mean that there are no internationally accepted methodological standards. A number of organisations and collaborations have published guidance documents for their assessments. These guidance documents reflect worldwide accepted methods which, in our view, should be seen as international standards.

This section provides an overview of a selection of these procedures and methods. For this purpose, the methods of institutes that evaluate technologies for government will be compared with the methods used by IQWiG. We will focus on the National Institute for Health and Clinical Excellence (NICE) as this is a leading institute, with most of its methods available in the public domain. Methods of the EUR-ASSESS group, reporting best practice in undertaking and reporting health technology assessments [8], the Canadian Agency for Drugs and Technology in Health (CADTH, formerly CCOHTA) [55], and several European agencies such as the French National Authority for Health (Haute Autorité de santé, HAS) [56], the Dutch Health Care Insurance Board (College voor Zorgverzekeringen, CvZ) [57], and the Danish Institute for Health Technology Assessment [36] were assessed, but it was not possible to

compare them meaningfully because their methods are available only in limited detail.

### 5.1 Methods of NICE and IQWiG

As the methods of an institute are related to its organisation, an outline of these aspects will be given below. NICE is an independent organisation, situated within the National Health Service (NHS) and is supported by central governmental funding [58]. The institute provides guidance to the NHS in England and Wales on the use of selected new and established technologies. The institute undertakes appraisals of health technologies at the request of the Department of Health and the Welsh Assembly Government [12].

IQWiG was founded as a non-profit, non-governmental private law foundation that has legal capacity and is, in this sense, an independent organisation. IQWiG is responsible for the scientific evaluation of the benefits and harms as well as the quality and efficiency of health-care services. Its responsibilities are to support the FJC in fulfilling its legislative duties by submitting recommendations, and to contribute to continuous improvement in the quality of health care for the public [59]. IQWiG receives funding out of the budgets of the sickness funds (partly from a quota related to hospital admissions and partly from a percentage calculated on cases in ambulatory care) from the FJC. Regular formal jurisdiction is explicitly not possible for benefit or cost-benefit assessments (Sect. 35b, Sect. 4 of the German Social Code Book).

The relationship between the FJC and the IQWiG is defined by the institute's charter within Sect. 7, § 5 clause 2 as follows: "The procedural regulations decided by the Federal Joint Committee on the basis of §91 (3) SGB V must be observed as far as the involvement of the Institute is concerned. The methodological requirements regarding the scientific, cross-sector evaluation of measures to be regulated in the procedural regulations and the demands on the professional independence of external experts must be defined in close consultation with the Institute Director." Thus the IQWiG is obliged to ensure that they perform their work in a way which enables the FJC to work with them in accordance with its own procedural code.

### 5.2 Assessment process used by NICE and IQWiG

The process of NICE has been described in detail [9, 12]. The involvement of relevant parties is one of the key principles in the process of NICE. Organisations that may wish to participate in the appraisal are identified as early as possible, when Ministers have provisionally decided on the list of technologies for appraisal [9]. These

organisations are invited to participate in the scoping process as well as in the assessment process, as consultants or as commentators [9]. Consultants are organisations that participate in the assessment and appraisal, such as the manufacturer(s) or sponsor(s) of the technology or national patient organisations. They may comment on various documents or products, write a submission, and can appeal against the Final Appraisal Determination. Commentators include organisations such as manufacturers of comparator technologies; they can comment only on the various documents. All comments made by consultants or commentators, including the response, will be made public by posting them on the website.

The scoping process consists of the submission of written material and a meeting called a scoping workshop, where issues raised are discussed. The scoping workshop was added to the written commenting phase when the methods were updated in 2004 [9, 60], which suggests that the use of solely written comments is not optimal for the scoping aim. The evidence assessment is performed by an independent group, while NICE performs the evidence appraisal and formulates recommendations.

To improve the work of NICE, it was evaluated in 2001/2002 and again in 2007/2008. A subsequent report describes what NICE does and how it works, changes made since its establishment, and the new challenges it faces. The authors conclude that the institute does a vital job in difficult circumstances and formulated recommendations to improve its functioning [61].

The main difference between the situation in Germany and England with respect to the scoping process is to be seen on the formal level of constitution of the institutes.

At NICE, the scoping process is the starting point of each commission and will lead to a specified protocol, which will be used directly by the commissioned experts. The scoping process is an open and transparent procedure; external experts take part, and their names and all products are made public.

At the IQWiG, the situation is less transparent. The IQWiG may modify the draft scope developed by the FJC, which could lead to an assessment that does not quite match the FJC's need. In its methods, the IQWiG describes the need to keep the external experts confidential, which directly hinders both participation in the scoping process and a transparent review process. In our view, the names of the external experts should be made public shortly after commissioning. Another problem exists because of the split of legal competence between the FJC and the IQWiG: The FJC is the only authority that can make the appraisal; the IQWiG and its external experts perform only the evidence assessment. However, the IQWiG is allowed to make recommendations, which

may lead to an overlap of the assessment and appraisal phases.

The IQWiG does involve relevant parties in parts of the assessment process. However, the processes are not clear. With respect to the scoping of a project: IQWiG drafts the (rough) definition of the research questions. If necessary, this definition will be refined by the project group, with inclusion of external experts (if required). Individual affected persons, patient representatives and/or consumer organisations will “regularly be involved with regard to the topic-related definition of patient-relevant outcomes” [62].

The IQWiG puts their report plans and preliminary reports on the web with the aim of inviting written comments. However, it remains unclear whether all comments are being made public and it also is not clear how the comments are handled. Only amendments to the report plan are described in the reports, meaning that comments of reviewers will be known only if the IQWiG agrees with them. In addition, comments that are not written in German may be omitted. Experts should be literate in German, which limits the pool of experts. The time period given for commenting on the products is perceived to be too short [63].

Table 1 presents an overview of the methods of the IQWiG [64] and NICE [45]. Both institutes published a draft update of their methods in November 2007. The information in the table below includes also information from the draft update of the methods, as this would have incorporated any recent amendments.

### 5.3 Conclusion

Methods used by other agencies that perform benefit assessments are useful in interpreting the term ‘international standard’ to which the institute must comply. NICE shows that it is possible to have transparent procedures for benefit assessments but that this requires detailed documentation. Their documents should be an example to other agencies in this respect. NICE has implemented an open transparent procedure with respect to the publication of the assessments produced by the external experts and comments of reviewers. Furthermore, their separation of evidence assessment and evidence appraisal prevents conflicts of interest in this last phase of the assessment.

Although the IQWiG invites comments on their protocol and preliminary report and posts them on their website, the comments are not clearly individually evaluated; therefore, it is not clear which comments are perceived to be relevant and which are not and whether they have been incorporated.

The participation of relevant parties in the assessment process is implemented properly by NICE, which guarantees a process that is acceptable to all relevant parties.

## 6. How should a benefit assessment be implemented?

This section describes how a benefit assessment should be implemented. It focusses on assessments to be performed in Germany and is based on the legal framework that requires transparent procedures, methods and criteria, the involvement of relevant parties and performance of assessments according to recognised standards of EBM. Experiences of agencies abroad that perform such assessments are taken into account where applicable. For general methodological guidance on benefit assessments, we refer to handbooks on systematic reviews published by the CRD [50] or the Cochrane Collaboration [6], and for general guidance on how to report such assessments to the QUOROM statement [7]. The process of the benefit assessment is presented in Figs. 1 and 2.

The FJC can commission a benefit assessment or a full economic evaluation. If a full economic evaluation is requested, a benefit assessment should be conducted first. However, as the methods of assessing benefit for an economic evaluation require a broader view compared with methods of assessing benefit for a benefit assessment, the methods for the benefit assessment should be broadened if a full economic evaluation of the medicine is requested. Also, it should be known that a benefit assessment based on RCTs has important disadvantages if this assessment is to serve as the basis for an economic evaluation, for example, due to the selective patient population, which differs from the general population. For more details on methods of economic evaluations that comply with international standards, we refer the reader to von der Schulenburg et al. [4] and Antes et al. [63]. This section focusses on benefit assessment without subsequent economic evaluation.

### 6.1 Topic identification and prioritisation

A public procedure should be developed for topic identification and prioritisation for potential future benefit assessments. Anyone should be able to suggest topics, by using a website, for example. A potential list of topics should be produced regularly by the FJC based on a set of criteria such as: “the intervention is likely to result in a significant health benefit across Germany if given to all patients for whom it is indicated” [9]. The list of criteria should be published on the FJC’s website. The topics should then be reviewed by the FMH, who decides which topic should be selected for an assessment. Reasons for (not) selecting a topic from the potential list of topics should be made public to guarantee a transparent procedure.

To prepare the benefit assessment, a draft scope is written. This document contains a first description of the patient population, the intervention and its setting, the



**Table 1** Overview of key issues of the process from the Institute for Quality and Efficiency in Health Care (IQWiG) and the National Institute for Health and Clinical Excellence (NICE)

	IQWiG	NICE
Topic identification and prioritisation		
The criteria for selecting topics are public <sup>aa</sup>	± <sup>a</sup>	+ <sup>o</sup>
The criteria for prioritizing topics are public	– <sup>b</sup>	NA <sup>p</sup>
Tender and commission		
The criteria experts should fulfil are public	+ <sup>c</sup>	± <sup>q</sup>
The criteria for selecting experts are public	– <sup>d</sup>	– <sup>r</sup>
The procedure for the commission is public	– <sup>e</sup>	–
Research question		
Relevant parties are involved in this phase	– <sup>f</sup>	+ <sup>s</sup>
An oral hearing is implemented in this phase	± <sup>g</sup>	+
All comments from relevant parties are public and evaluated	NA <sup>h</sup>	+ <sup>t</sup>
Protocol		
The protocol is published	+ <sup>i</sup>	+ <sup>u</sup>
Relevant parties are involved in this phase	± <sup>j</sup>	– <sup>v</sup>
An oral hearing is implemented in this phase	± <sup>g</sup>	–
All comments of relevant parties are public and evaluated	± <sup>k</sup>	NA
Preliminary report		
The report as written by experts is published	–	+ <sup>w</sup>
Relevant parties are involved in this phase	+	+
An oral hearing is implemented in this phase	± <sup>g/l</sup>	–
All comments of relevant parties are public and evaluated	± <sup>m</sup>	+
Review procedure		
The criteria for inviting reviewers are public	–	– <sup>x</sup>
All comments of the reviewers are public	–	–
Final report		
Results appraisal is published separately from results assessment	–	+ <sup>y</sup>
The underlying evidence for each recommendation is made public	+	+
Appeal		
Appeal possible?	NA <sup>n</sup>	+ <sup>z</sup>

Criterion: + Fulfilled; ± partly fulfilled; – not fulfilled; NA not applicable

<sup>a</sup> Although the methods refer to broad criteria such as burden of disease and burden of cost, more specific criteria were not found

<sup>b</sup> IQWiG receives commissions on several topics at the same time and decides for itself which to do first. It is not clear how these topics are prioritised

<sup>c</sup> Criteria to describe the experience of experts are published (methods version 2, page 98)

<sup>d</sup> It is not clear how the experts are selected from the pool of experts who fulfil the criteria

<sup>e</sup> The bidding is open, but the decision-making is not clear

<sup>f</sup> In the new methods, it is stated that the research question will be defined by the project group (methods version 3, page 17, first and third paragraph)

<sup>g</sup> In the new methods, oral hearings are an option at each step of the process; however, IQWiG decides if a comment is worth being discussed, and this lacks transparency

<sup>h</sup> IQWiG does not involve relevant parties in this phase; hence, there are no comments

<sup>i</sup> Protocol is available on the Internet, for example: <http://www.iqwig.de/index.651.en.html>

<sup>j</sup> Relevant parties are invited to make written comments only (methods version 2, page 102). However, it is stated that they can be involved, meaning that involvement is not always implemented. It is not clear under which circumstances they will not be involved

<sup>k</sup> Comments are published on the Internet but not individually evaluated, and therefore it is not clear what was done with the comments

<sup>l</sup> An oral hearing is optional. It is not clear when such a hearing is implemented (methods version 3, page 16)

<sup>m</sup> Comments are published but not individually evaluated, and therefore it is not clear what was done with the comments, for example: [http://www.iqwig.de/download/N06-01A\\_Dokumentation\\_und\\_Wuerdigung\\_der\\_Stellungnahmen\\_zum\\_Vorbericht.pdf](http://www.iqwig.de/download/N06-01A_Dokumentation_und_Wuerdigung_der_Stellungnahmen_zum_Vorbericht.pdf)

**Table 1** continued

- <sup>n</sup> This is outside the scope of IQWiG's methods. Appeal is possible on the level of the FJC
- <sup>o</sup> Criteria are published in the process document [9], p. 3
- <sup>p</sup> There is no real prioritisation at this stage: suitable topics are referred for a scope and, based on the scope, Ministers decide whether or not to commission the topic
- <sup>q</sup> The academic groups that prepare assessments for NICE are established through occasion tenders. The choice of which group does which assessment depends on capacity, conflicts of interest, expertise, and preferences of the groups
- <sup>r</sup> NICE works together with seven academic centres (<http://www.nchta.org/publicationspdfs/infoleaflets/nice.pdf>). No criteria are stated
- <sup>s</sup> Consultants and commentators are consulted using a scoping process (methods document [12], p. 2 and 7)
- <sup>t</sup> Comments are public, for example: ([http://www.nice.org.uk/nicemedia/pdf/Asthma\\_children\\_comments\\_on\\_draftscope.pdf](http://www.nice.org.uk/nicemedia/pdf/Asthma_children_comments_on_draftscope.pdf))
- <sup>u</sup> Protocol is public, for example: <http://www.nice.org.uk/guidance/index.jsp?action=byID&o=11711>
- <sup>v</sup> Assessment group develops protocol based on scoping process (methods document [12], p. 7 Sect. 2.1.3)
- <sup>w</sup> The process states 'authors are responsible for content and quality of assessment (Process document, Sect. 4.4.1.5). Authors' names are stated on the report, for example: (<http://www.nice.org.uk/nicemedia/pdf/AssessmentReportSenttoC&CAsthmaChildren.pdf>)
- <sup>x</sup> The report by the external experts can be commented on by stakeholders in preparation of the appraisal meeting. In addition, the draft final report undergoes peer review and review by editors of the journal Health Technology Assessment
- <sup>y</sup> Described in methods document [12]
- <sup>z</sup> For consultants only (described in the guide for manufacturers p. 15)
- <sup>aa</sup> Accessible via the Internet

comparator intervention(s) and the proposed outcome measures. This procedure will ensure that the assessment to be performed serves the need of the FJC. At the same time, the FJC commissions the institute for the assessment. The topic is posted on the FJC's website to inform potential relevant parties that an assessment procedure has been started and that they are invited to participate.

## 6.2 Tendering and awarding the commission

Pursuant to Sect. 139b, § 3 of the German Social Code Book V, the institute must assign scientific research commissions for the conduct of benefit assessments to external experts. Although public tenders for all individual assessments are to be preferred, for reasons of planning and timely processing the institute could implement public tenders with contracting for a fixed number of assessments. Comparable to the awarding practice of the German Institute for Medical Documentation and Information (DIMDI), contracts for processing a known number of benefit assessments could be bindingly tendered every year. Internationally, this procedure resembles the practice of NICE, which collaborates with seven university institutions [65]. NICE does not conduct its own assessments.

Independent of the tendering procedure, the criteria that are applied to the selection of the external experts as well as the tendering procedure need to be made public and posted on the Internet. The pool of experts should be as broad as possible, and therefore the application of language restrictions is not recommended. To ensure a high degree of procedural transparency, the names of the external experts must be posted on the Internet within 4 weeks of when the

commission is made as well as published in the preliminary and final report. All experts involved in the assessment should declare any conflicts of interest regarding the product to be evaluated as well as the included alternatives, and disclose any past relationships to associated manufacturing companies in the preceding 3 years. Adequate checks of the procedures must be installed.

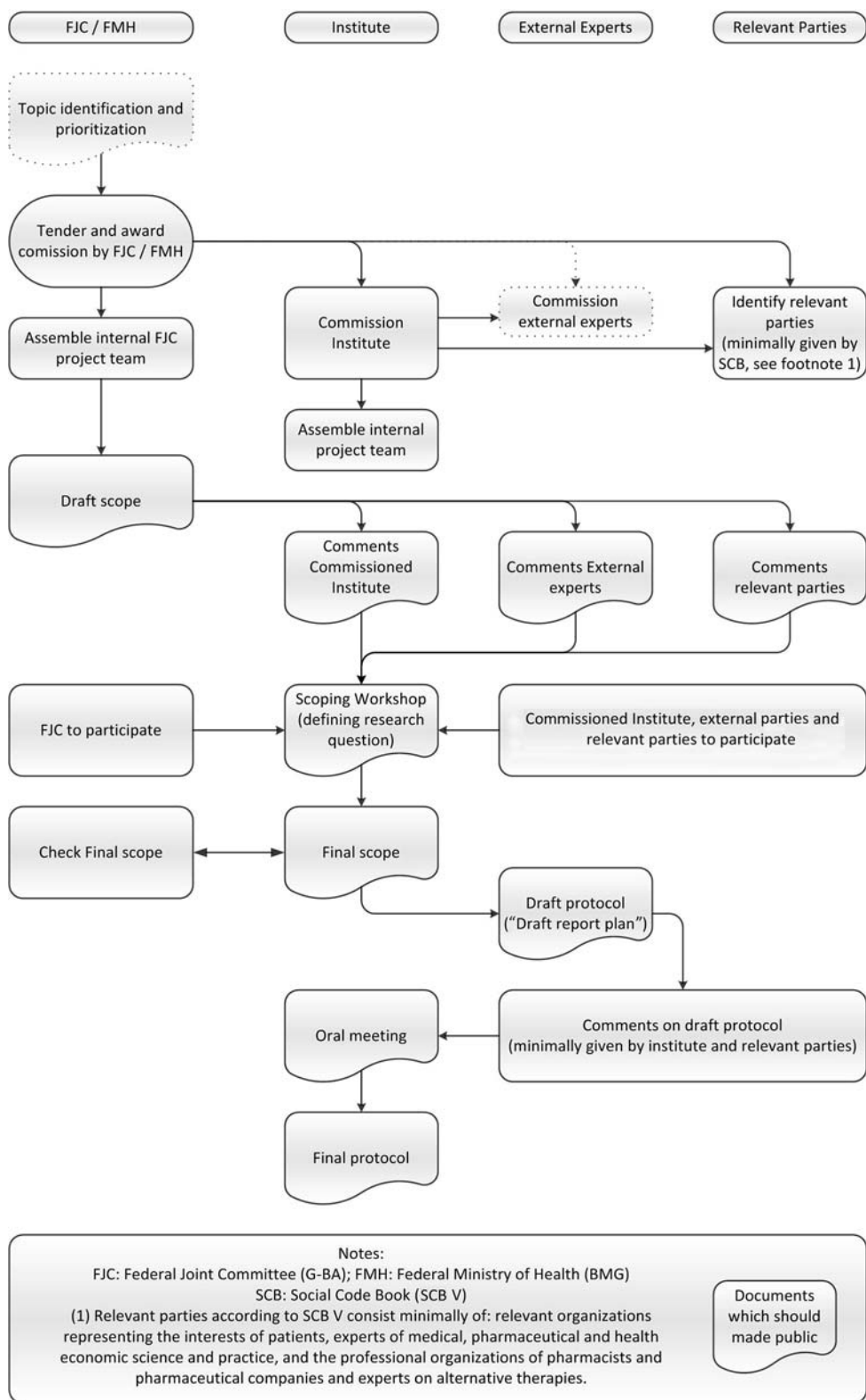
The commissioned institute should keep a log of each assessment on the Internet, as already implemented by IQWiG and NICE. Such a log would be the place to publish the procedure and all criteria in order to fulfil the requirement of high procedural transparency. It should contain the names of the external experts, as well as the time frame of the project, which should be agreed on beforehand by all participants of the project. It also should provide an up-to-date statement about the status of the project and all products that should be made public with a view to transparency (e.g. draft scope, final scope, draft protocol).

## 6.3 Defining the research question

Defining the research question is the most important phase of a study, as this determines the boundaries of the assessment. Relevant parties should be involved in all important parts of the assessment and thus also in defining the research question. Patients can contribute meaningfully to this phase of research; they are essential, since patient benefit is to be assessed as intended by the legislator.

To fulfil the requirements of transparency, such involvement should be implemented in two steps: first

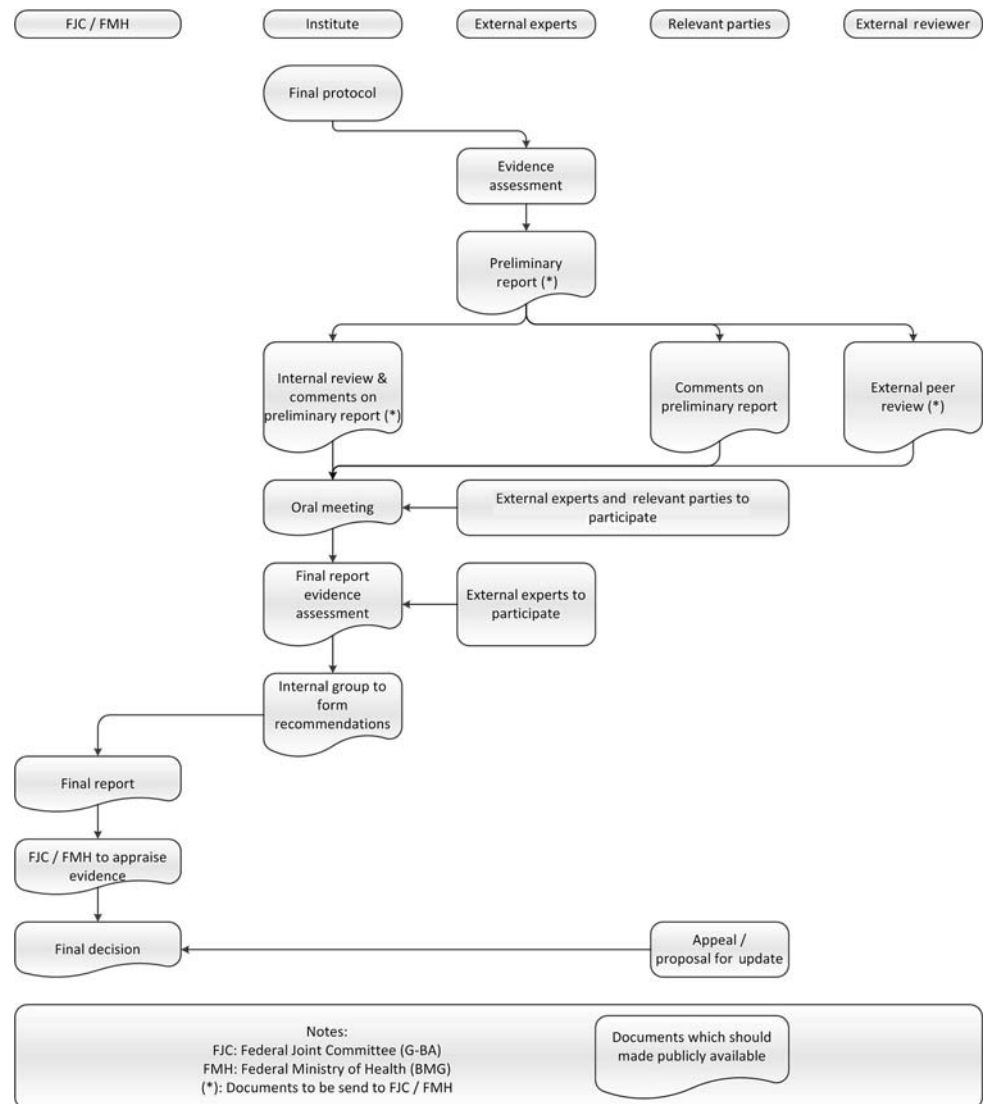
**Fig. 1** Process of benefit assessment up to publication of the final protocol



comments in writing and then a scoping workshop. The draft scope is sent to the institute, the relevant parties and the external experts, all of whom are invited to give written

comments within the agreed time frame. All submitted comments should be evaluated by the project team of the institute and commented on individually as to whether or

**Fig. 2** Process of benefit assessment from publication of the final protocol onward



not they are relevant. For transparency reasons all comments should be available publicly together with the reasons for rating certain comments as not relevant. Within the agreed time frame representatives of the comments as well as representatives from the FJC and the external experts who perform the assessment will be invited to the scoping workshop. This workshop has the following objectives:

- To evaluate and, if required, propose a revision of the problem
- To suggest clinically relevant comparative therapies
- To propose patient-relevant outcomes, including for each relevant party a definition and operationalisation of the term ‘benefit’
- To propose relevant subgroups that could benefit more or less from the intervention
- To suggest a commission-related methodology, including inclusion and exclusion criteria for the selection of literature
- To highlight relevant issues to the external experts in order to inform the development of the protocol and the appraisal

The scoping workshop should be headed by an independent person who will serve as a moderator. The workshop aims to generate discussion on the scope and the assessment from different perspectives so that an appropriate final scope can be produced, which leads to the development of a protocol by the commissioned institute. The moderator should create the broadest possible consensus with regard to the aforementioned objectives of the scoping workshop. A word-for-word transcript should be generated for the workshop, which will be part of the

protocol accessible over the Internet and the final report. The final scope should closely match the clinical problem that formed the basis for the commission of this benefit assessment. To ensure this, the FJC should be invited to participate in this phase.

#### 6.4 Developing the protocol

Based on the relevant suggestions given during the scoping workshop, a draft protocol will be developed. The protocol contains the following items: background information with research question(s), search strategy, study selection criteria and procedures, study quality assessment, data-extraction strategy, and synthesis of extracted evidence. The protocol also contains the time schedule for the assessment.

The research question is operationalised using the PICO criteria, which define in detail the patients, the intervention, the comparison intervention, and the outcomes. The results of the scoping workshop should form the basis for the research question.

Based on the results of the scoping workshop, the protocol should also identify which study designs ideally should be used to answer the question(s). This should not be used as a cut-off point to exclude studies of lower evidence levels, as the assessment should be made on the evidence that is currently available. The information should guide evidence appraisal with respect to the strength of evidence.

Within the agreed time frame, the protocol should be published and made accessible on the Internet. Subsequently, the relevant parties can submit their comments; an oral hearing to discuss the comments then follows. The comments, together with documentation about whether or not a suggestion was considered relevant and has been incorporated into the research question, should be published on the Internet and in the preliminary and the final report. If an oral hearing is not performed, this should be justified in the preliminary report.

To prevent reporting bias where possible, all relevant outcomes reported in the primary studies should be assessed. If outcomes are added to or deleted from the protocol once the assessment has started, this should be justified in an amendment to the protocol.

#### 6.5 Assessment of the evidence

To fulfil the legal requirement of transparency and to counteract the value judgments that are an inevitable part of such a process, all steps in the evidence assessment should be clearly documented. All steps in this process should be performed with equal transparency if the benefit assessment is conducted by internal staff of the institute.

##### 6.5.1 Search strategy

Based on the search strategy as specified in the protocol, a search to identify relevant studies will be conducted by the external experts. The objective of the literature search is to identify all studies that, at the time of the search, may be suitable to answer the questions posed in the assessment. The search should be sufficiently broad as to identify literature on all the research questions of interest. Alternatively, separate searches should be performed for each question, as the studies and study types to be retrieved may differ.

A range of databases should be searched, as no single database is comprehensive enough to record all publications from all medical journals. General databases such as Medline, Embase and the databases in the Cochrane Library form a good basis from which to start. Apart from these, other databases that could provide additional literature references regarding the problem should be included. Since the benefit assessment must be conducted primarily in view of the German health-care situation, it must be ensured that studies especially relevant for Germany be identified. This can be done through an additional search of the most important databases of German publishing companies and a supplementary manual search of the relevant trade journals.

The search should be inclusive: neither language restrictions nor the exclusion of specific study designs (such as non-RCTs) or publication types (such as abstracts or unpublished studies) are recommended. Additionally, reference lists of relevant publications should be screened for missed studies. Relevant publications typically emerge during the search strategy, but these may also be identified during the scoping process. The list of relevant publications that were screened should be included in the preliminary and final report.

The search strategy should be clearly documented. The QUOROM statement recommends documenting the information sources (databases, registers, personal files, expert informants, agencies, hand-searching) and any restrictions (years considered, publication status, language of publication) in detail [7].

##### 6.5.2 Submission by manufacturers

As early as possible, the institute should inform the relevant manufacturers of the implementation of the benefit assessment and ask them to submit information on relevant published and unpublished studies within the agreed time frame. When it comes to the transmission of studies and data, the sample confidentiality agreement ratified by the IQWiG and the German Association of Research-Based Pharmaceutical Companies (VFA) can be used to protect

operative and business secrets of the pharmaceutical companies. All data used in the evidence assessment should be made public, at a time point to be agreed on by both the institute and the manufacturers.

### 6.5.3 Selection of studies

All publications identified in the literature search and submitted by the pharmaceutical companies will be independently examined by at least two scientists (external experts) with regard to their subject relevance by using the title and abstract information. The selection criteria should be clearly documented. All studies on the assessed medicine for indications as licensed should be included. Any co-intervention that is commonly used should be allowed, and thus co-interventions should not be a reason for excluding any of these studies.

Subsequently, the selected articles will be retrieved and evaluated in their full text version, also by at least two scientists. For each publication that is excluded, the reason for exclusion should be documented. The remaining studies are included in the assessment and their data will be extracted.

If systematic reviews were identified on a topic similar to the topic of the assessment, these should be examined closely. If systematic reviews have already been performed for a question, these must be reviewed based on the principles of EBM with the highest priority and evaluated and considered accordingly. The existence of a methodologically sound, recent systematic review on the same topic may make (parts of) the assessment redundant.

### 6.5.4 Data-extraction process

All selected papers should be read carefully and relevant data extracted. In case of missing data, authors should be contacted for additional information to make sure that no relevant information is lost owing to poor reporting. A standardised and transparent process of requests to the authors should be established for this purpose. The responses to all questions, including missing responses, should be made public.

An important element of the data extraction is the assessment of methodological quality. Such an assessment is useful only if it has consequences for the rest of the review. For example, the methodological quality may be used to explore quality differences as an explanation for heterogeneity in study results, or to guide the interpretation of findings by aiding the determination of the strength of inferences.

There are several quality-assessment instruments, and the choice of such a measure depends on the purpose of quality assessment. Therefore, the purpose of quality assessment

and the motivation for choosing a measure should be documented in the protocol, and consequences of quality assessment should be described. As the judgment of methodological quality requires subjective decisions, double data extraction should be employed for these assessments.

### 6.5.5 Data synthesis

The aim of the data synthesis is to summarise the results of included primary studies. It can be performed using descriptive, qualitative synthesis or, if possible, by quantitative methods. For more guidance on these methods, we refer to handbooks on systematic reviews published by the CRD [50] or the Cochrane Collaboration [6].

### 6.5.6 Subgroup analyses

Benefit assessments have the disadvantage that a resulting recommendation is not transferable to each individual patient. Factors such as age, sex, co-morbidities, and co-medications can influence the efficacy and safety of a therapy. In order to focus the results of a benefit assessment on specific target groups, where possible, the following strategy is suggested:

1. Subgroups that are relevant for the indication to be assessed should be defined during the scoping process.
2. The available data should be analysed for their statistical feasibility in the corresponding subgroup analyses.
3. The defined subgroups need to be published in a transparent way, as well as reasons for performing or not performing the corresponding subgroup analyses.
4. Results regarding the relevant subgroups should be published. When it is not possible to analyse an important subgroup due to statistical problems, trends should be published with appropriate advice concerning their lower significance.

Advantages of subgroup analyses are the optimisation of benefit and the reduction of harm. This can have a positive influence on costs. A technical advantage is that the heterogeneity of study design and results is less problematic in the course of the benefit assessment.

## 6.6 Publication of the preliminary report

The preliminary report as written by the external experts must be sent to the FJC and the relevant parties. It should also be published for access over the Internet. In addition, an internal review by the institute and an external peer review will be performed (see also “[Review of preliminary report](#)”). All involved parties are invited to submit written comments within the agreed time frame.

The submitted comments should focus on the methods of the assessment as laid out in the protocol compared with the subsequent implementation of the methods in the preliminary report as well as with a summary assessment of the evaluation for the benefit assessment.

The comments received are to be evaluated by the external experts and documented individually with regard to their relevance and to whether or not it they were incorporated. The submitted comments as well as the response of the external experts should be published within the agreed time frame and should always be included in the subsequent preliminary or final report.

Within the agreed time frame, the submitters of the comments as well as representatives from the FJC will meet in an oral hearing. The essence of the hearing is the scientific discussion of facts in dispute with the goal of improving the quality and acceptance of the assessment. The hearing will be chaired by a representative appointed by the institute management. If an oral hearing does not take place, this should be justified in the final report.

### 6.7 Review of preliminary report

An internal and an external review procedure should be implemented to support the external experts. The institute conducts its own internal review of the preliminary report, during which especially compliance with the formal requirements of the protocol are verified. In parallel, the content-related conclusions and the quality of the study assessments are evaluated by external peer reviewers. For transparency reasons, the criteria that external experts who peer review the report need to fulfil should be made public, as should the procedure for inviting and selecting them. Furthermore, as argued above, the best way to achieve transparency is open review, meaning that both the names of the reviewers and their comments are made public by posting them on the Internet. The review reports are forwarded to the contracting entity.

### 6.8 Recommendations and final report

Based on the comments of the relevant parties, the verbal contributions during the oral hearing, and the results of the internal and external reviews, the preliminary report is revised by the institute in consultation with the external experts. The institute then makes initial recommendations based on the evidence. The revised preliminary report and the recommendations constitute the final report. The recommendations need to be clearly marked for reasons of transparency. At the end of the agreed time period, the final report will be published on the institute's website.

The appraisal phase can then start. This should be performed by the FJC to separate the phase of evidence

assessment from the phase of evidence appraisal. In the appraisal phase, the evidence will be judged against the specific circumstances in Germany, and recommendations for further full compensation, ceiling costs or any other possible regulation of the medicine will be formulated. The factors taken into account will differ from appraisal to appraisal but may include economic efficiency, safety, social and ethical criteria and German health-care aspects. After the agreed time period, the final report, including evidence appraisal, will be published on the institute's website.

### 6.9 Planning update report and appeal decision

The benefit assessment represents the best available evidence on the topic at the time of the assessment. New evidence could have consequences for the conclusion of the assessment. A review of the assessment should be planned every 3–5 years.

Pursuant to Sect. 35b, § 2, clauses 2 and 3, regular verification of the assessment results is provided. The report must be regularly deliberated within the FJC. Independent of this fact, the manufacturers have the right to petition for an update when new scientific evidence is available. The FJC or FMH must decide on the petitions. The decision and its justification must be published on the Internet website of the FJC.

All relevant parties should have the right to appeal the decisions of the FJC based on the final report.

### 6.10 Conclusion

This paper describes the operational implementation of the legal requirements with regard to the benefit assessment of medicines in Germany. Legally, such assessments require participation of relevant parties, transparency and the use of EBM according to internationally accepted standards. To fulfil the requirement of active participation of relevant parties, a scoping workshop should be implemented to better involve these parties in defining the research question, which is a crucial phase in the assessment. Transparency of the whole process should be achieved by reporting of all procedures and criteria used in all phases of the assessment. Specifically, the input of all parties involved in defining the research question should be made clear by making all comments public, together with an evaluation of whether or not they led to changes in the research question. This procedure should also be applied to show the input of all involved parties in the development of the protocol and the preliminary report. Likewise, the comments of the peer reviewers should be made public. Furthermore, transparency requires a separation of the

phases of evidence assessment and evidence appraisal. Using the principles of EBM means that all evidence should be assessed in order to determine which is the best evidence that is available to answer each research question. Such a process ensures that the benefit assessments of medicines in Germany are performed according to the highest standards.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

## References

1. Federal Joint Committee (2006) Code of procedure. [Gemeinsamer Bundesausschuss: Verfahrensordnung] Version from 20.09.2005, last changes from 18.04.2006, enforced: 07.07.2006. <http://www.g-ba.de/institution/aufgaben-und-struktur/Verfahrensordnung/>
2. Cochrane Effective Practice and Organisation of Care Group (2008) Effective interventions unit. Evaluation guide 5: designing an economic evaluation. <http://www.epoc.cochrane.org/Files/Website/Reviewer%20Resources/draftps.pdf>
3. CRD (2008) What is an economic evaluation? <http://www.york.ac.uk/inst/crd/nfaq1.htm> CRD, York
4. von der Schulenburg, J., Vauth, C., Mittendorf, T., Greiner, W.: Methods for determining cost-benefit ratios for pharmaceuticals in Germany. *Eur J Health Econ* **8**(Suppl 1), S5–S31 (2007)
5. Göhlen, B., Rüter, A.: HTA beim DIMDI. *Z Arztl Fortbild Qualitatssich Gesundh Wes* **101**, 508–511 (2007)
6. Higgins JPT, Green S (eds) (2007) Cochrane handbook for systematic reviews of interventions 4.2.6 [updated September 2006]. <http://www.cochrane.org/resources/handbook/hbook.htm> Accessed: 8th November 2007
7. Moher, D., Cook, D.J., Eastwood, S., Olkin, I., Rennie, D., Stroup, D.F.: Improving the quality of reports of meta-analyses of randomised controlled trials: the QUOROM statement. Quality of reporting of meta-analyses. *Lancet* **354**, 1896–1900 (1999)
8. Velasco, M., Perleth, M., Drummond, M., Gurtner, F., Jorgensen, T., Jovell, A., Malone, J., Ruther, A., Wild, C.: Best practice in undertaking and reporting health technology assessments. Working group 4 report. *Int J Technol Assess Health Care* **18**, 361–422 (2002)
9. NICE: Guide to the technology appraisal process. National Institute for Clinical Excellence, London (2004)
10. Van Rooyen, S., Godlee, F., Evans, S., Black, N., Smith, R.: Effect of open peer review on quality of reviews and on reviewers' recommendations: a randomised trial. *BMJ* **318**, 23–27 (1999)
11. Godlee, F., Gale, C.R., Martyn, C.N.: Effect on the quality of peer review of blinding reviewers and asking them to sign their reports: a randomized controlled trial. *JAMA* **280**, 237–240 (1998)
12. NICE: Guide to the methods of technology appraisal. National Institute for Clinical Excellence, London (2004)
13. DUET (2008) Database of uncertainties about the effects of treatments (DUETs): a resource to help prioritise new research. <http://www.duets.nhs.uk/Default.asp> Accessed: Feb 12, 2008
14. Tallon, D., Chard, J., Dieppe, P.: Consumer involvement in research is essential. *BMJ* **320**, 380–381 (2000)
15. Goodare, H., Lockwood, S.: Involving patients in clinical research improves the quality of research. *BMJ* **319**, 724–725 (1999)
16. INVOLVE (2008) Promoting public involvement in NHS, public health and social care research. <http://www.invo.org.uk/> Accessed: Jan 10th 2008
17. Hanley, B., Truesdale, A., King, A., Elbourne, D., Chalmers, I.: Involving consumers in designing, conducting, and interpreting randomised controlled trials: questionnaire survey. *BMJ* **322**, 519–523 (2001)
18. Kahan, J.P., Park, R.E., Leape, L.L., Bernstein, S.J., Hilborne, L.H., Parker, L., Kamberg, C.J., Ballard, D.J., Brook, R.H.: Variations by specialty in physician ratings of the appropriateness and necessity of indications for procedures. *Med Care* **34**, 512–523 (1996)
19. Coulter, I., Adams, A., Shekelle, P.: Impact of varying panel membership on ratings of appropriateness in consensus panels—a comparison of a multidisciplinary and single-disciplinary panel. *Health Serv Res* **30**, 577–591 (1995)
20. Grant-Pearce, C., Miles, I., Hills, P.: Mismatches in priorities for health research between professionals and consumers. A report to the standing advisory group on consumer involvement in the NHS R&D Programme. PREST, University of Manchester, Manchester (1998)
21. Devereaux, P.J., Anderson, D.R., Gardner, M.J., Putnam, W., Flowerdew, G.J., Brownell, B.F., Nagpal, S., Cox, J.L.: Differences between perspectives of physicians and patients on anticoagulation in patients with atrial fibrillation: observational study. *BMJ* **323**, 1218–1222 (2001)
22. Montgomery, A.A., Fahey, T.: How do patients' treatment preferences compare with those of clinicians? *Qual Health Care* **10**(Suppl 1), i39–i43 (2001)
23. Chard, J., Dickson, J., Tallon, D., Dieppe, P.: A comparison of the views of rheumatologists, general practitioners and patients on the treatment of osteoarthritis. *Rheumatology (Oxford)* **41**, 1208–1210 (2002)
24. Kirwan, J.R., Minnock, P., Adebajo, A., Bresnihan, B., Choy, E., de Wit, M., Hazes, M., Richards, P., Saag, K., Suarez-Almazor, M., Wells, G., Hewlett, S.: Patient perspective: fatigue as a recommended patient centered outcome measure in rheumatoid arthritis. *J Rheumatol* **34**, 1174–1177 (2007)
25. Garland, A.F., Lewczyk-Boxmeyer, C.M., Gabayan, E.N., Hawley, K.M.: Multiple stakeholder agreement on desired outcomes for adolescents' mental health services. *Psychiatr Serv* **55**, 671–676 (2004)
26. Lee, T.T., Ziegler, J.K., Sommi, R., Sugar, C., Mahmoud, R., Lenert, L.A.: Comparison of preferences for health outcomes in schizophrenia among stakeholder groups. *J Psychiatr Res* **34**, 201–210 (2000)
27. Kwok, C.K., Ibrahim, S.A.: Rheumatology patient and physician concordance with respect to important health and symptom status outcomes. *Arthritis Rheum-Arthritis Care Res* **45**, 372–377 (2001)
28. Hubbard, G., Kidd, L., Donaghy, E., McDonald, C., Kearney, N.: A review of literature about involving people affected by cancer in research, policy and planning and practice. *Patient Educ Couns* **65**, 21–33 (2007)
29. Oliver, S., Milne, R., Bradburn, J., Buchanan, P., Kerridge, L., Wally, T., Gabbay, J.: Involving consumers in a needs-led research programme: a pilot project. *Health Expectations* **4**, 18–28 (2001)
30. Bradburn, J., Maher, J., Adewuyi-Dalton, R., Grunfeld, E., Lancaster, T., Mant, D.: Developing clinical trial protocols: the use of patient focus groups. *Psychooncology* **4**, 107–112 (1995)
31. Ali, K., Roffe, C., Crome, P.: What patients want: consumer involvement in the design of a randomized controlled trial of routine oxygen supplementation after acute stroke. *Stroke* **37**, 865–871 (2006)



32. Culyer, A.J.: Involving stakeholders in healthcare decisions—the experience of the National Institute for Health and Clinical Excellence (NICE) in England and Wales. *Healthcare Quart* **8**, 56–60 (2005)
33. The AGREE Collaboration: Development and validation of an international appraisal instrument for assessing the quality of clinical practice guidelines: the AGREE project. *Qual Saf Health Care* **12**, 18–23 (2003)
34. Marsden, J., Bradburn, J.: Patient and clinician collaboration in the design of a national randomized breast cancer trial. *Health Expect* **7**, 6–17 (2004)
35. Sackett, D.L., Straus, S.E., Richardson, W.S., Rosenberg, W., Haynes, R.B.: Evidence-based medicine. How to practice and teach EBM, 2nd edn. Churchill Livingstone, Edinburgh (2000)
36. Danish Institute for Health Technology Assessment: Health Technology Assessment Handbook, 1st edn. Danish Institute for Health Technology Assessment, Copenhagen (2001)
37. Fletcher, R.H., Fletcher, S.W., Wagner, E.H.: Clinical epidemiology—the essentials. Williams & Wilkins, Baltimore (1982)
38. Villar, J., Carroli, G., Belizan, J.M.: Predictive ability of meta-analyses of randomised controlled trials. *Lancet* **345**, 772–776 (1995)
39. LeLorier, J., Gregoire, G., Benhaddad, A., Lapierre, J., Derderian, F.: Discrepancies between meta-analyses and subsequent large randomized, controlled trials. *N Engl J Med* **337**, 536–542 (1997)
40. Cappelleri, J.C., Ioannidis, J.P., Schmid, C.H., de Ferranti, S.D., Aubert, M., Chalmers, T.C., Lau, J.: Large trials vs. meta-analysis of smaller trials: how do their results compare? *JAMA* **276**, 1332–1338 (1996)
41. Ioannidis, J.P., Cappelleri, J.C., Lau, J.: Meta-analyses and large randomized, controlled trials. *N Engl J Med* **338**, 59–62 (1998)
42. Ioannidis, J.P., Cappelleri, J.C., Lau, J.: Issues in comparisons between meta-analyses and large trials. *JAMA* **279**, 1089–1093 (1998)
43. Furukawa, T.A., Streiner, D.L., Hori, S.: Discrepancies among megatrials. *J Clin Epidemiol* **53**, 1193–1199 (2000)
44. Jadad, A.R., Cook, D.J., Browman, G.P.: A guide to interpreting discordant systematic reviews. *Can Med Assoc J* **156**, 1411–1416 (1997)
45. NICE: Guide to the methods of technology appraisal. Draft for consultation (Nov 2007). NHS, National Institute for Health and Clinical Excellence, England (2007)
46. Claxton, K., Sculpher, M., Drummond, M.: A rational framework for decision making by the National Institute for Clinical Excellence (NICE). *Lancet* **360**, 711–715 (2002)
47. Guyatt, G.H., Sackett, D.L., Sinclair, J.C., Hayward, R., Cook, D.J., Cook, R.J.: Users' guides to the medical literature. IX. A method for grading health care recommendations. Evidence-Based Medicine Working Group. *JAMA* **274**, 1800–1804 (1995)
48. Harris RP, Helfand M, Woolf SH, Lohr KN, Mulrow CD, Teutsch SM, Atkins D, for the Methods Work Group Third U.S. Preventive Services Task Force (2008) Current methods of the U.S. Preventive Services Task Force: a review of the process. <http://www.ahrq.gov/clinic/ajpmsuppl/harris1.htm> Accessed: Feb 12, 2008
49. Philips B, Ball C, Sackett D, Badenoch D, Straus S, Haynes B, Dawes M (2001) Oxford Centre for evidence-based medicine levels of evidence (may 2001). <http://www.cebm.net/index.aspx?o=1025> Accessed: November 8th, 2007
50. CRD (2001) Undertaking systematic reviews of research of effectiveness. CRD's guidance for those carrying out or commissioning reviews, CRD Report 4, 2nd edn. CRD, York
51. Egger, M., Dickersin, K., Davey Smith, G.: Problems and limitations in conducting systematic reviews. In: Egger, M., Davey Smith, G., Altman, D.G. (eds.) *Systematic reviews in health care. Meta-analysis in context*, pp. 43–68. BMJ Publishing Group, London (2001)
52. Jüni, P., Altman, D.G., Egger, M.: Assessing the quality of randomised controlled trials. In: Egger, M., Davey Smith, G., Altman, D.G. (eds.) *Systematic reviews in health care. Meta-analysis in context*, pp. 87–108. BMJ Publishing Group, London (2001)
53. Bucher, H.C., Guyatt, G.H., Griffith, L.E., Walter, S.D.: The results of direct and indirect treatment comparisons in meta-analysis of randomized controlled trials. *J Clin Epidemiol* **50**, 683–691 (1997)
54. Song, F., Altman, D.G., Glenny, A.M., Deeks, J.J.: Validity of indirect comparison for estimating efficacy of competing interventions: empirical evidence from published meta-analyses. *BMJ* **326**, 472 (2003)
55. CCOHTA (2003) Guidelines for authors of CCOHTA Health Technology Assessment reports. Canadian Coordinating Office for Health Technology Assessment, editor
56. Haute Autorité de santé (2007) General method for assessing health technologies. source: [http://www.has-sante.fr/portail/display.jsp?id=c\\_541195](http://www.has-sante.fr/portail/display.jsp?id=c_541195). Haute Autorité de santé
57. Staal, P.C., Ligtenberg, G.: Beoordeling stand van de wetenschap en praktijk, 254. College voor Zorgverzekeringen, Diemen (2007)
58. Martelli, F., Torre, G.L., Ghionno, E.D., Staniscia, T., Neroni, M., Cicchetti, A., Bremen, K.V., Ricciardi, W.: Health technology assessment agencies: an international overview of organizational aspects. *Int J Technol Assess Health Care* **23**, 414–424 (2007)
59. Bastian, H., Bender, R., Ernst, A.S., Kaiser, T., Kirchner, H., Kolominsky-Rabas, P., Lange, S., Sawicki, P.T., Weber, M.: Institute for quality and efficiency in health care. Methods (pre-ample). IQWiG, Cologne, Germany (2007)
60. NICE: Guide to the technology appraisal process. National Institute for Clinical Excellence, London (2001)
61. House of Commons HC (2008) National Institute for Health and Clinical Excellence. First report of session 2007–08. Volume 1. Report, together with formal minutes. HC 27-I. 10-1-2008. The Stationery Office Limited, London
62. Bastian, H., Bender, R., Ernst, A.S., Kaiser, T., Kirchner, H., Kolominsky-Rabas, P., Lange, S., Sawicki, P.T., Weber, M.: Institute for quality and efficiency in health care. Methods (version 2.0). IQWiG, Cologne (2007)
63. Antes G, Jöckel KH, Kohlmann T, Raspe H, Wasem J (2007) Commenting synopsis of expert positions for economic evaluations for pharmaceuticals (“Kommentierende Synopse der Fachpositionen zur Kosten-Nutzenbewertung für Arzneimittel”). [http://www.bmg.bund.de/cln\\_041/nn\\_1145310/DE/Themenschwerpunkte/Gesundheit/Arzneimittel/Fachtagung/kommentierte-synopse.templateId=raw.property=publicationFile.pdf/kommentierte-synopse.pdf](http://www.bmg.bund.de/cln_041/nn_1145310/DE/Themenschwerpunkte/Gesundheit/Arzneimittel/Fachtagung/kommentierte-synopse.templateId=raw.property=publicationFile.pdf/kommentierte-synopse.pdf) Federal Ministry of Health, Germany
64. IQWiG (2007) Allgemeine Methoden. Entwurf für Version 3.0 vom 15.11.2007. IQWiG, Cologne
65. Collaborative centres NICE (2008) <http://www.nccta.org/publicationspdfs/infoleaflets/nice.pdf> Accessed: January 16th, 2008