# The guidelines manual

## Appendices

# The guidelines manual: appendices

## *Contents*



# Appendix A: Agreements and advice for Guideline Development Group members

## A1 Code of conduct for Guideline Development Group (GDG) members and others who attend GDG meetings

### *A1.1 Key principles of development*

NICE's clinical guideline development process:

- involves national patient and professional organisations (such as GDG members and stakeholders)
- involves companies that manufacture relevant medicines or devices, and the Department of Health and the Welsh Assembly Government (as stakeholders)
- uses sound and transparent methodologies
- produces guidance that is based on the clinical and economic evidence, and is clearly explained.

GDGs should incorporate into clinical guidelines recommendations that emerge from NICE's technology appraisal, interventional procedures and public health programmes, and should also take into account recommendations from appropriate national service frameworks (NSFs). In general, NICE clinical guidelines are concerned with the delivery of clinical care but not the configuration of services.

Each GDG should ensure that its guideline is developed in line with these requirements. It should also follow the principles set out in 'Social value judgements: principles for the development of NICE guidance (second edition)'[1] and adhere to the NICE equality scheme and action plan[2].

### *A1.2 Status of GDG members*

Members are appointed to a GDG either by virtue of their relevant experience (as in the case of patient and carer members and healthcare professional members) or because they have specific technical skills (as in the case of systematic reviewers and health economists). If members are from stakeholder organisations, NICE and the GDG assume that these members bring this perspective to the group, but they do not represent their organisations. GDG members are appointed for the duration of the development process for a clinical guideline.

People appointed to the GDG are co-authors of the guideline. They will respect the rights of NICE both to publish the final guideline documents and to

[1] Available from: www.nice.org.uk/aboutnice/howwework/socialvaluejudgements/socialvaluejudgements.jsp
[2] Available from: www.nice.org.uk/aboutnice/howwework/NICEEqualityScheme.jsp

receive notification of associated publications, as described in contracts with the National Collaborating Centres (NCCs).

### *A1.3 Mutual undertaking*

NICE, usually through one of its NCCs, undertakes to:

- ensure that the GDG is properly resourced to produce the guideline
- provide all members of the GDG with equal access to available resources and to the evidence used in the development of the guideline
- offer appropriate training to GDG members to enable them to play a full part in the development of the guideline
- provide technical support during the development of the guideline.

GDG members undertake to:

- make sufficient time available to attend meetings and properly inform the development of the guideline through their personal and professional knowledge and, where appropriate, their organisation's perspective
- provide the GDG, and subsequently (and only after failure to resolve the issue within the GDG) the NCC and NICE, with the opportunity to consider and resolve concerns or disagreements about either the process or the detail of the emerging guideline
- contribute positively to the work of the group and the development of the guideline.

### *A1.4 Transparency*

NICE believes that its guidance will be enhanced if those who are intended to benefit from it and those who have the responsibility for implementing it have been had the opportunity to be involved in its development.

For GDGs to operate successfully, they need to be able to develop and debate issues within the group before exposing them to wider comment. There is therefore a need for arrangements that protect the confidentiality of documents and discussions.

In order to provide the environment described above, NICE expects GDG members:

- to be aware that the Guidance Executive and Senior Management Team at NICE will not comment on the development of a guideline in progress, other than in the context of the formal consultation exercises
- to regard the views expressed by individual members of the GDG as confidential
- to regard the documents and discussions used by the GDG as confidential to the group until public consultation, as stipulated in the Confidentiality acknowledgement and undertaking agreement (see appendix A2)
- not to discuss commercial-in-confidence data outside the GDG if a technology appraisal is updated within a guideline

- to respect the confidentiality of documents supporting a published or unpublished technology appraisal and guidelines in development if such documents are received by the GDG
- to respect the confidentiality of documents relating to other unpublished NICE guidance (interventional procedures, public health guidance) if such documents are received by the GDG.

GDG members are also expected to adhere to NICE's policy for declaring conflicts of interests[3] (see also section A4.4).

[3] See: www.nice.org.uk/aboutnice/whoweare/policiesandprocedures/policies_and_procedures.jsp?domedia=1&mid=EEF24FBA-19B9-E0B5-D4ED345FBCECFBA1

# NATIONAL INSTITUTE FOR HEALTH AND CLINICAL EXCELLENCE

# Participation in NICE guidelines

## A2 Confidentiality acknowledgement and undertaking

Please complete the sections below and return by email to: [insert NCC email]

If email is not possible, please return by fax to: [insert NCC fax no.]

This agreement covers all those who have sight of documents, or are party to discussions, relating to a guideline before public consultation. This includes members of the Guideline Development Group (GDG), invited external experts, observers and participants in consensus exercises. Staff of national collaborating centres (NCCs) are covered by the contracts between NICE and the NCCs.

1) I undertake to NICE that I shall:

   (a) keep all confidential information strictly confidential
   (b) not use any confidential information for any purpose other than participating in the deliberations of the GDG (for GDG members and external experts)
   (c) not disclose any confidential information to any third party without the prior written consent of NICE
   (d) not disclose the deliberations of a GDG to any other person without the explicit consent of the Chair of the GDG and the Director of the NCC.

2) The undertakings set out in paragraph 1 above ('the undertakings') shall not apply to the use or disclosure of information that:

   (a) at or after the time of disclosure or acquisition is in the public domain in the form supplied otherwise than through a breach of any of the undertakings; or
   (b) was lawfully within my possession before its disclosure to me by NICE, provided that the source of such information was not bound by, or subject to, a confidentiality agreement with NICE; or
   (c) I am required to disclose by any court of competent jurisdiction or any government agency lawfully requesting the same, provided that I notify NICE in advance of such disclosure; or
   (d) is approved for release by prior written authorisation from NICE.

Signed ……..Date ……..

Print name ……

Data Protection. The personal data submitted on this form will be used by the National Institute for Health and Clinical Excellence for work on its Guidelines Programme and will be held on the Institute's databases for future reference and in accordance with the Data Protection Act 1998.

# A3 Dealing with enquiries on GDG work

## *A3.1 Introduction*

As a member of a GDG you will be considered by some to be a source of information, to have important influence or to be a lobbying target. This guidance will help you to deal with any enquiries you receive from individuals or organisations about your work on the GDG.

Although NICE will not publish your contact details anywhere, your name and the organisation that supported your application (if appropriate) will be published on the NICE website. Thus it may be easy for those who want to contact you to do so.

Just like the guideline you are developing, this document is guidance, not 'must do', and has been developed to support you in handling enquiries so that you do not feel obliged to deal with them yourself.

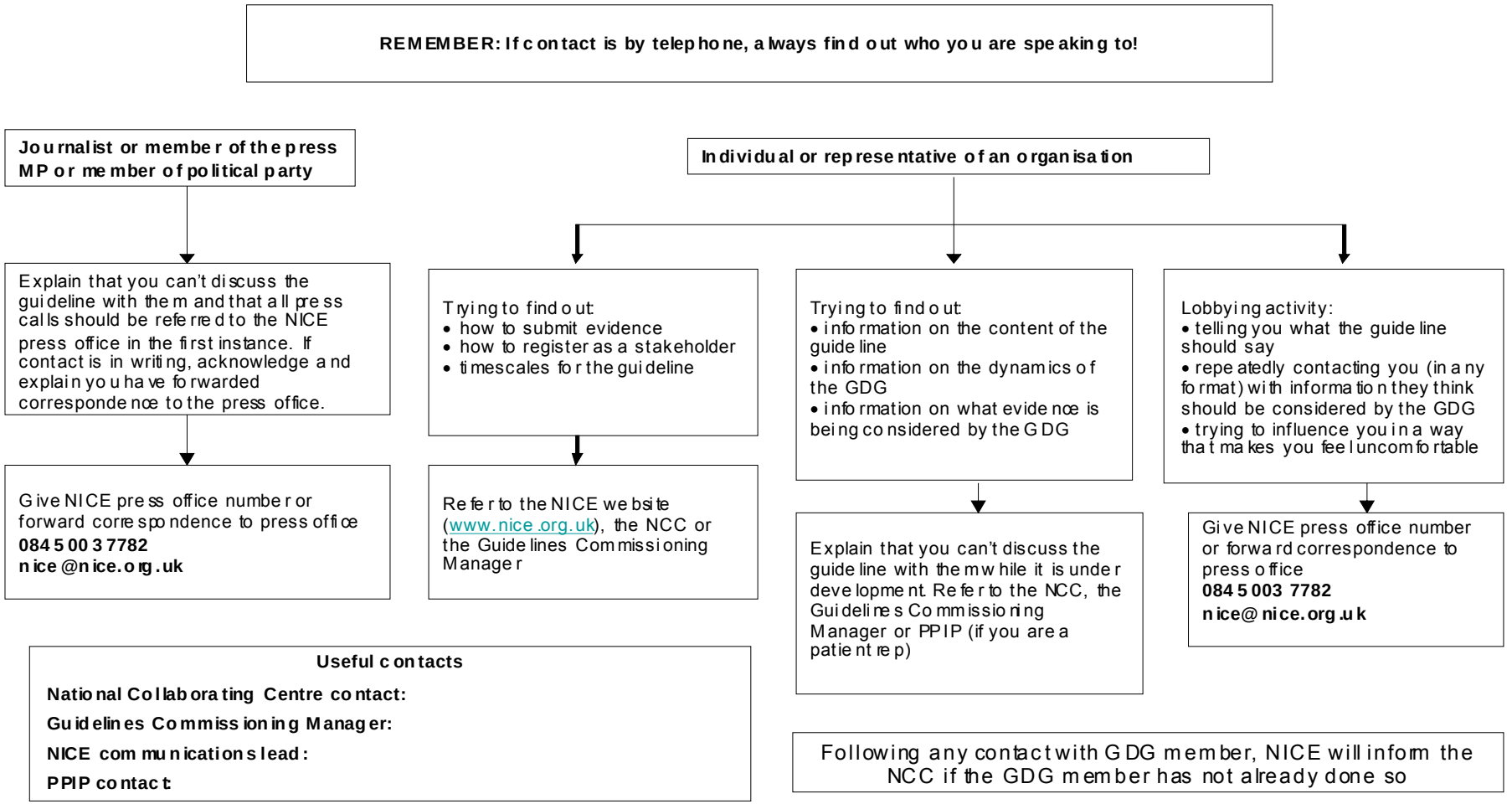
## *A3.2 Golden rules*

Some things to remember when talking to anyone about the guideline you are developing:

- Don't feel that you have to talk to anyone about the guideline: you can handle requests for information by offering to pass them on to someone who is able to deal with them (such as your lead contact at the NCC, the NICE Guidelines Commissioning Manager or the NICE communications lead).
- Don't speculate on the content of the guideline before it is finally published.
- Draft versions are just that: draft, not final; the content may change after consultation.
- Individuals and organisations can influence the outcome of the guideline only by submitting evidence that supports their point of view as part of the formal consultation process.
- You have not been selected to sit on the GDG to represent all patients, clinicians, nurses or other healthcare professionals. You are there to provide your own expert opinion to the group.
- In the unlikely event that you are contacted by telephone by an unpleasant or demanding caller, offer to pass the enquiry on to your NCC or NICE contact. Then tell the caller politely but firmly that you wish to end the call. If the caller persists, put the phone down.

## *A3.3 Flow chart for dealing with enquiries*



# NATIONAL INSTITUTE FOR HEALTH AND CLINICAL EXCELLENCE

## A4 Guidance for Guideline Development Group (GDG) members on attendance at NICE Appraisal Committee meetings

### *A4.1 Introduction*

Before invitations to participate in a NICE technology appraisal are sent out, the Guideline Commissioning Managers are contacted to establish whether a related clinical guideline is on the work programme, and if so, who should act as representative(s) of the guideline. Members of the GDG – usually the GDG Chair (or the expert with the most relevant experience) and one other member (usually the NCC Director, the Clinical Adviser or another senior clinician) – are invited to attend the appraisal consultation document (ACD) and final appraisal determination (FAD) meetings of the Appraisal Committee considering any appraisal that is relevant to the development of the guideline.

When the topic of the appraisal relates to a guideline on the work programme for which development has not yet begun (that is, there is no GDG in place), the NCC Director should be invited to attend on behalf of the future GDG.

Project managers from the appraisals and guidelines programmes at NICE will liaise at an early stage in each appraisal to determine appropriate links between the relevant committees and GDGs, and to provide operational support.

### *A4.2 Purpose of attendance and role*

The attendance of GDG members at the Appraisal Committee meeting allows them to participate fully in discussions about the technology. They can remain for the concluding discussions of the Appraisal Committee after the patient experts and clinical specialists have left the meeting.

The GDG members attending the Appraisal Committee meeting will also, in conjunction with their GDG as a whole, act as commentators on the documents produced. They will receive the ACD and FAD, and the GDG comments fed back via the NCC will be included in the review of the documents by the technical lead and the Appraisal Committee Chair, and be brought to the attention of the Appraisal Committee.

As is the case for other commentators, GDG members do not have the right of appeal.

### *A4.3 Voting*

In the event of a decision of the Appraisal Committee being taken by a vote, the GDG members attending the meeting will not have the right to vote.

## *A4.4 Conflicts of interest*

GDG members who attend the Appraisal Committee meetings will be expected to declare conflicts of interest[4] and to abide by the current rulings on these for full committee members if they wish to take part in all of the Committee's discussions.

GDG representatives who attend Appraisal Committee meetings do so as committee members (except for voting; see section A4.3). Therefore a GDG representative with a personal pecuniary interest will not be able to attend the Appraisal Committee discussion.

GDG members who have conflicts of interest that would have excluded Appraisal Committee members can be present at meetings only for the same period as the clinical specialists and patient experts, and must leave the room with them before the concluding discussions. This would require their nomination as clinical specialists by the Appraisals' consultees and commentators. Such specialists are only usually present at the first Appraisal Committee meeting (that is, the meeting to develop the ACD).

## *A4.5 GDG members as clinical specialists or patient experts*

GDG members may occasionally also be nominated as clinical specialists or patient experts for an appraisal. They may then act in both capacities, but must leave the meeting with the other specialists and experts if they have a conflict of interest (see section A4.4). GDG members may wish to avoid this dual role in order to maximise their attendance at Appraisal Committee meetings. Exclusion from the second half of the meeting because of conflicts of interest does not preclude GDG members who have attended from providing written comments during consultation.

## *A4.6 GDG comments on the ACD and FAD*

It is expected that the GDG comments on the ACD and FAD will represent a consensus view, expressed in a single document, preferably submitted via the GDG Chair and coordinated by the NCC.

[4] See: www.nice.org.uk/aboutnice/whoweare/policiesandprocedures/policies_and_procedures.jsp?domedia=1&mid=EEF24FBA-19B9-E0B5-D4ED345FBCECFBA1

# Appendix B: Study design checklist

| | | | | |
|---|---|---|---|---|
| **Study identification**<br>*Include author, title, reference, year of publication* | | | | |
| **Guideline topic:** | **Review question no:** | | | |
| **Checklist completed by:** | | | | |
| **Was there a comparison:** | | | | |
| Between two or more groups of participants? | Yes | No | Unsure | N/A |
| Within the same group over time? | Yes | No | Unsure | N/A |
| **How were groups formed?** | | | | |
| Randomisation | Yes | No | Unsure | N/A |
| Quasi-randomisation | Yes | No | Unsure | N/A |
| Other action of researchers | Yes | No | Unsure | N/A |
| Time differences | Yes | No | Unsure | N/A |
| Location differences | Yes | No | Unsure | N/A |
| Treatment decision-makers | Yes | No | Unsure | N/A |
| Participant preferences | Yes | No | Unsure | N/A |
| On the basis of outcome | Yes | No | Unsure | N/A |
| **Which parts of the study were prospective?** | | | | |
| Identification of participants | Yes | No | Unsure | N/A |
| Assessment of baseline and intervention allocation | Yes | No | Unsure | N/A |
| Assessment of outcomes | Yes | No | Unsure | N/A |
| Generation of hypotheses | Yes | No | Unsure | N/A |
| **Which variables were used for comparing the groups assessed?** | | | | |
| Potential confounders | Yes | No | Unsure | N/A |
| Baseline assessment of outcome variables | Yes | No | Unsure | N/A |

## *Notes on the use of Study design checklist*

This checklist is taken from:

> Higgins JPT, Green S, editors (2008) Cochrane handbook for systematic reviews of interventions, version 5.0.1 [updated September 2008], chapter 13. The Cochrane Collaboration. Available from: www.cochrane-handbook.org

Tables 13.2.a and 13.2.b in the Cochrane handbook are lists of study design features for studies with allocation to interventions at the individual and group level respectively. Box 13.4.a provides useful notes for completing the checklist.

# Appendix C: Methodology checklist: systematic reviews and meta-analyses

| | | | |
|---|---|---|---|
| **Study identification**<br>*Include author, title, reference, year of publication* | | | |
| **Guideline topic:** | **Review question no:** | | |
| **Checklist completed by:** | | | |
| **SCREENING QUESTIONS** | | | |
| **In a well-conducted, relevant systematic review:** | ***Circle one option for each question*** | | |
| The review addresses an appropriate and clearly focused question that is relevant to the guideline review question | Yes | No | Unclear |
| The review collects the type of studies you consider relevant to the guideline review question | Yes | No | Unclear |
| The literature search is sufficiently rigorous to identify all the relevant studies | Yes | No | Unclear |
| Study quality is assessed and reported | Yes | No | Unclear |
| An adequate description of the methodology used is included, and the methods used are appropriate to the question | Yes | No | Unclear |

If the review does not meet some or all of these criteria, it may still be useful as a source of references, but should not be relied upon on its own to address a review question.

If you have insufficient information on the design or quality of individual studies, you should use the checklists for studies on interventions (see appendices D, E and F) to appraise each study. Each study should appear as a separate entry in the evidence table (see appendix K); the review should not appear in the evidence table.

If you plan to use the review as a whole, you will need to complete a row in an evidence table for the systematic review and input the results into an evidence profile as appropriate.

## *Notes on use of Methodology checklist: systematic reviews and meta-analyses*

A systematic review uses explicit and systematic methods to identify, appraise and summarise the literature according to predetermined criteria. If the methods and criteria used to do this are not described or are not sufficiently detailed, it is not possible to make a thorough evaluation of the quality of the review.

The terms ‘systematic review’ and ‘meta-analysis’ are often used interchangeably. The term ‘meta-analysis’ is often used incorrectly to describe a systematic review that has used quantitative methods to summarise the results. However, it should be noted that meta-analysis refers only to the statistical techniques used to combine studies; thus not all meta-analyses are systematic reviews.

This checklist is intended for use with systematic reviews of questions about interventions and questions about diagnosis. It can potentially be used for any other types of question, although it has been designed primarily for the first two.

The aim of this checklist is to consider the suitability of the systematic review to answer a guideline review question. This assessment has two aspects: firstly, whether the question addressed by the review (in terms of the populations, interventions, comparisons and outcomes considered) is appropriate to answer the review question addressed by the guideline, and secondly, whether the methodology used for the review is sufficiently robust to permit a valid conclusion.

For each question in this section, you should indicate whether or not it has been addressed in the review. Choose ‘unclear’ if this aspect of the review process was ignored, or is not described in the report.

### The review addresses an appropriate and clearly focused question that is relevant to the guideline review question

If the question addressed by the systematic review is not clearly stated, it will be difficult to determine whether the review is adequate to answer the guideline review question. If the question is not clear, the systematic review is unlikely to be a good one because it difficult to be systematic in addressing an unclear question. The review report should give a clear description of the population considered, the interventions, exposures or tests evaluated, comparators, and outcomes evaluated. Inclusion and exclusion criteria should be clearly described. Outcomes considered should be clearly described within the methodology, including a precise definition and acceptable methods of measuring. The appropriateness of the question addressed in the systematic review for answering the guideline review question can be determined by comparing these components. If the review does not consider all of the outcomes that are judged to be important to your guideline review question, you may still be able to use the outcome data but may need to review the individual studies to obtain other outcome data.

### The review collects the type of studies you consider relevant to the guideline review question

You should be clear about the characteristics of studies that you consider will adequately address your guideline review question. These may relate to minimum design or quality characteristics (for example, randomised trials only). Systematic reviews should report the types of studies they sought, including any inclusion/exclusion criteria used. You can use this information to quickly assess the review's suitability for your purpose.

### The literature search is sufficiently rigorous to identify all the relevant studies

Systematic and rigorous searches can help to minimise publication biases and identify as many relevant data as possible. Exact search terms depend on the review question, but there are core databases that should have been searched for every question. As a minimum, a well-conducted review should look at EMBASE and MEDLINE. For questions about interventions in particular, the Cochrane Database of Systematic Reviews (CDSR), the Database of Abstracts of Reviews of Effects (DARE) and the Cochrane Central Register of Controlled Trials (CENTRAL) should also be searched. The dates on which the searches were carried out should be given in the review. Good-quality reviews will also attempt to identify relevant studies by handsearching of key journals and examining reference lists of retrieved studies for further references.

If the methods used to locate studies are not clearly reported, it will be difficult to determine whether the review is likely to have missed important relevant studies. Ideally, the search strategy used should be reported in sufficient detail that the process could be replicated.

Any restrictions applied to the search (such as language or year of publication) should also be reported. You should consider how these might have influenced the findings of the review.

Advice from the information specialist (and/or other members of the Guideline Development Group) working on the guideline may be useful to determine whether any important search terms have been omitted.

If the search described in the review is judged to be inadequate to identify all relevant studies, it may be possible to expand the search by including additional databases or extra search terms within the search strategy, or by updating the search to identify more recently published studies. Any additional studies identified by this expanded search should be appraised for quality using the appropriate NICE checklist (see appendices D–J). They should appear individually in separate rows in an evidence table.

### Study quality is assessed and reported

The inclusion of poor-quality studies within a review can result in biased estimates of effect. A well-conducted systematic review should have used clear criteria to assess whether individual studies had been appropriately designed and conducted, before deciding whether to include or exclude them.

These criteria should be clearly described and should be reported for each study included. The quality appraisal checklists in appendices D–J, as appropriate for the type of question and study design, can be used as a guide to the types of quality criteria that should be considered.

If there is no indication of such a quality assessment, the review is unlikely to be reliable enough to be used in formulating guideline recommendations. It may be necessary to obtain and quality appraise the individual studies as part of your review.

**An adequate description of the methodology used is included, and the methods used are appropriate to the question**

In common with primary research, the approach used to analyse the data should be described and justified where appropriate. This may include the choice of statistical test used to analyse the outcome data, meta-analytical techniques and approaches to dealing with heterogeneity, including the specification of any subgroup analyses and sensitivity analyses.

# Appendix D: Methodology checklist: randomised controlled trials

| | | | | | |
|---|---|---|---|---|---|
| **Study identification**<br>*Include author, title, reference, year of publication* | | | | | |
| **Guideline topic:** | | **Review question no:** | | | |
| **Checklist completed by:** | | | | | |
| | | ***Circle one option for each question*** | | | |
| **A. Selection bias (systematic differences between the comparison groups)** | | | | | |
| A1 | An appropriate method of randomisation was used to allocate participants to treatment groups (which would have balanced any confounding factors equally across groups) | Yes | No | Unclear | N/A |
| A2 | There was adequate concealment of allocation (such that investigators, clinicians and participants cannot influence enrolment or treatment allocation) | Yes | No | Unclear | N/A |
| A3 | The groups were comparable at baseline, including all major confounding and prognostic factors | Yes | No | Unclear | N/A |
| Based on your answers to the above, in your opinion was selection bias present? If so, what is the likely direction of its effect? | | | | | |
| Low risk of bias | Unclear/unknown risk | High risk of bias | | | |
| Likely direction of effect: | | | | | |
| **B. Performance bias (systematic differences between groups in the care provided, apart from the intervention under investigation)** | | | | | |
| B1 | The comparison groups received the same care apart from the intervention(s) studied | Yes | No | Unclear | N/A |
| B2 | Participants receiving care were kept 'blind' to treatment allocation | Yes | No | Unclear | N/A |
| B3 | Individuals administering care were kept 'blind' to treatment allocation | Yes | No | Unclear | N/A |
| Based on your answers to the above, in your opinion was performance bias present? If so, what is the likely direction of its effect? | | | | | |
| Low risk of bias | Unclear/unknown risk | High risk of bias | | | |
| Likely direction of effect: | | | | | |

| | | | | | |
|---|---|---|---|---|---|
| **C. Attrition bias (systematic differences between the comparison groups with respect to loss of participants)** | | | | | |
| C1 | All groups were followed up for an equal length of time (or analysis was adjusted to allow for differences in length of follow-up) | Yes | No | Unclear | N/A |
| C2 | a. How many participants did not complete treatment in each group? | | | | |
| | b. The groups were comparable for treatment completion (that is, there were no important or systematic differences between groups in terms of those who did not complete treatment) | Yes | No | Unclear | N/A |
| C3 | a. For how many participants in each group were no outcome data available? | | | | |
| | b. The groups were comparable with respect to the availability of outcome data (that is, there were no important or systematic differences between groups in terms of those for whom outcome data were not available). | Yes | No | Unclear | N/A |
| Based on your answers to the above, in your opinion was attrition bias present? If so, what is the likely direction of its effect? | | | | | |
| Low risk of bias | Unclear/unknown risk | High risk of bias | | | |
| Likely direction of effect: | | | | | |
| **D. Detection bias (bias in how outcomes are ascertained, diagnosed or verified)** | | | | | |
| D1 | The study had an appropriate length of follow-up | Yes | No | Unclear | N/A |
| D2 | The study used a precise definition of outcome | Yes | No | Unclear | N/A |
| D3 | A valid and reliable method was used to determine the outcome | Yes | No | Unclear | N/A |
| D4 | Investigators were kept 'blind' to participants' exposure to the intervention | Yes | No | Unclear | N/A |
| D5 | Investigators were kept 'blind' to other important confounding and prognostic factors | Yes | No | Unclear | N/A |
| Based on your answers to the above, in your opinion was detection bias present? If so, what is the likely direction of its effect? | | | | | |
| Low risk of bias | Unclear/unknown risk | High risk of bias | | | |
| Likely direction of effect: | | | | | |

## *Notes on use of Methodology checklist: randomised controlled trials*

The studies covered by this checklist are designed to answer questions about the relative effects of interventions such as drugs, psychological therapies, operations or placebos. Such studies can include comparisons of 'test and treat strategies' involving a diagnostic test and subsequent management. The checklist does not cover comparisons of diagnostic test accuracy or questions about prognosis.

This checklist replaces the methodology checklist for randomised controlled trials from 'The guidelines manual' 2007 (appendix C).

Some of the items on this checklist may need to be filled in individually for different outcomes reported by the study. It is therefore important that the systematic reviewer has a clear idea of what the important outcomes are **before** appraising a study. You are likely to need input from the Guideline Development Group in defining the important outcomes.

Checklist items are worded so that a 'yes' response always indicates that the study has been designed/conducted in such a way as to minimise the risk of bias for that item. An 'unclear' response to a question may arise when the item is not reported or not clearly reported. 'N/A' should be used when a randomised controlled trial cannot give an answer of 'yes' no matter how well it has been done.

This checklist is designed to assess the internal validity of the study; that is, whether the study provides an unbiased estimate of what it claims to show. Internal validity implies that the differences observed between groups of participants allocated to different interventions may (apart from the possibility of random error) be attributed to the intervention under investigation. Biases are characteristics that are likely to make estimates of effect differ systematically from the truth.

*Recording the presence and direction of bias*

The checklist contains four sections (A–D), each of which addresses a potential source of bias relating to internal validity. At the end of each section you are asked to give your opinion on whether bias is present and to estimate the likely direction of this bias – that is, whether you think it will have increased or decreased the effect size reported by the study. It will not always be possible to determine the direction of bias, but thinking this through can help greatly in interpreting results.

### A: Selection bias

Selection bias may be introduced into a study when there are systematic differences between the participants in the different treatment groups. As a result, the differences in the outcome observed may be explained by pre-existing differences between the groups rather than because of the treatment itself. For example, if the people in one group are in poorer health, then they are more likely to have a bad outcome than those in the other group,

regardless of the effect of the treatment. The treatment groups should be similar at the start of the study – the only difference between the groups should be the intervention received.

*Randomisation*

There are two aspects to randomisation:

- generation of the random allocation sequence that results in groups that differ only randomly
- allocation concealment, so that both the participant and the investigator are unaware of which group the next participant will be allocated to when entering the study.

### *A1. An appropriate method of randomisation was used to allocate participants to treatment groups*

If an appropriate method of randomisation has been used, each participant should have an equal chance of ending up in any of the treatment groups. Examples of random allocation sequences include random numbers generated by computer, tables of random numbers, and drawing of lots or envelopes. The allocation sequence should not be related to outcome or prognosis, or be predictable, such as date of birth or admission date.

There are some more complicated ways of allocating people to treatment groups that minimise the differences between groups, such as block randomisation and minimisation. Although these are not truly random, they are usually considered to be adequate for the purpose. If a study does not report the method of randomisation used, this should be scored as ‘unclear’.

### *A2. There was adequate concealment of allocation*

If investigators are aware of the allocation group for the next participant being enrolled in the study, there is potential for participants to be enrolled in an order that results in imbalances in important characteristics. For example, a clinician might feel that participants who are more unwell are likely to do better on a new, experimental, treatment and be tempted to enrol such participants when they know they will be allocated to that group. This would result in the participants in the intervention group being, on average, more unwell. Concealment of treatment group may not always be feasible (as in, for example, a comparison of a surgical with a medical intervention), but concealment of allocation up until the point of enrolment in the study should always be possible.

The information presented within the paper should provide some assurance that allocations were not known until at least the point of enrolment. Centralised allocation, computerised allocation systems and the use of coded identical containers are all regarded as adequate methods of concealment. Sealed envelopes can be considered as adequate concealment if the envelopes are serially numbered, sealed and opaque, and allocation is performed by a third party. Poor methods of allocation concealment include

alternation, or the use of case record numbers, date of birth or day of the week.

If the method of allocation concealment used is regarded as poor, or relatively easy to subvert, the study must be given a lower quality rating. If a study does not report any concealment approach, this should be scored as 'unclear'.

### *A3. The groups were comparable at baseline, including all major confounding and prognostic factors*

Studies may report the distributions of potential prognostic and confounding factors in the comparison groups, or important differences in the distribution of these factors may be noted.

Formal tests comparing the groups are problematic – failure to detect a difference does not mean that a difference does not exist, and multiple comparisons of factors may falsely detect some differences that are not real.

Clinical input may be required to determine whether all likely confounders have been considered. Confounding factors may differ according to outcome, so you will need to consider potential confounding factors for all of the outcomes that are of interest to your review.

## B: Performance bias

Performance bias refers to systematic differences between the comparison groups in the care provided to the participants, other than the intervention under investigation.

This may consist of additional treatment, advice or counselling, rather than a physical intervention, or even simply a belief about the effects of an intervention. If performance bias is present, it can be difficult to attribute any observed effect to the experimental treatment rather than to the other factors.

### *B1. The comparison groups received the same care apart from the intervention(s) studied*

There should be no differences between the treatment groups apart from the intervention received. If some participants received additional treatment (known as 'co-intervention'), this treatment is a potential confounding factor that may compromise the results.

*Blinding*

Blinding (also known as masking) refers to the process of withholding information about treatment allocation or exposure status from those involved in the study who could potentially be influenced by this information. This can include participants, investigators, those administering care and those involved in data collection and analysis. If people are aware of the treatment allocation or exposure status ('unblinded'), this can bias the results of studies, either intentionally or unintentionally, through the use of other effective co-interventions, decisions about withdrawal, differential reporting of symptoms or influencing concordance with treatment. Blinding of those assessing outcomes is covered in section D on detection bias.

Blinding of participants and carers is not always possible, particularly in studies of non-drug interventions, and so performance bias may be a particular issue in these studies. It is important to think about the likely size and direction of bias caused by failure to blind.

The terms 'single blind', 'double blind' and even 'triple blind' are sometimes used in studies. Unfortunately, they are not always used consistently. Commonly, when a study is described as 'single blind', only the participants are blind to their group allocation. When both participants and investigators are blind to group allocation, the study is often described as 'double blind'. It is preferable to record exactly who was blinded, if reported, to avoid misunderstanding.

### *B2. Participants receiving care were kept 'blind' to treatment allocation*

The knowledge of assignment to a particular treatment group may affect outcomes, such as a study participant's reporting of symptoms, self-use of other known interventions or even dropping out of the study.

### *B3. Individuals administering care were kept 'blind' to treatment allocation*

If individuals who are administering the intervention and/or other care to the participant are aware of treatment allocation, they may treat participants receiving one treatment differently from those receiving the comparison treatment; for example, by offering additional co-interventions.

## C: Attrition bias

Attrition refers to the loss of participants during the course of a study. Attrition bias occurs when there are systematic differences between the comparison groups with respect to participants lost, or differences between participants lost to the study and those who remain. Attrition can occur at any point after participants have been allocated to their treatment groups. As such, it includes participants who are excluded after allocation (and may indicate a violation of eligibility criteria), those who do not complete treatment (whether or not they continue measurement) and those who do not complete outcome measurement (regardless of whether or not treatment was completed). Consideration should be given to why participants dropped out, as well as how many. Participants who dropped out of a study may differ in some significant way from those who remained as part of the study throughout. Drop-out rates and reasons for dropping out should be similar across all treatment groups. The proportion of participants excluded after allocation should be stated in the study report, and the possibility of attrition bias considered within the analysis; however, these are not always reported.

### *C1. All groups were followed up for an equal length of time (or analysis was adjusted to allow for differences in length of follow-up)*

If the comparison groups are followed for different lengths of time, then more events are likely to occur in the group followed up for longer, distorting the comparison. This may be overcome by adjusting the denominator to take the time into account; for example by using person-years.

***C2a. How many participants did not complete treatment in each group?***

A very high number of participants dropping out of a study should give concern. The drop-out rate may be expected to be higher in studies conducted over a longer period of time. The drop-out rate includes people who did not even start treatment; that is, they were excluded from the study after allocation to treatment groups.

***C2b. The groups were comparable for treatment completion (that is, there were no important or systematic differences between groups in terms of those who did not complete treatment)***

If there are systematic differences between groups in terms of those who did not complete treatment, consider both why participants dropped out and whether any systematic differences in those who dropped out may be related to the outcome under study, such as potential confounders. Systematic differences between groups in terms of those who dropped out may also result in treatment groups that are no longer comparable with respect to potential confounding factors.

***C3a. For how many participants in each group were no outcome data available?***

A very high number of participants for whom no outcome data were available should give concern.

***C3b. The groups were comparable with respect to the availability of outcome data (that is, there were no important or systematic differences between groups in terms of those for whom outcome data were not available)***

If there are systematic differences between groups in terms of those for whom no outcome data were available, consider both why the outcome data were not available and whether there are any systematic differences between participants for whom outcome data were and were not available.

**D: Detection bias (this section should be completed individually for each important relevant outcome)**

The way outcomes are assessed needs to be standardised for the comparison groups; failure to 'blind' people who are assessing outcomes can also lead to bias, particularly with subjective outcomes. Most studies report results for more than one outcome, and it is possible that detection bias may be present in a study for some, but not all, outcomes. It is therefore recommended that this section is completed individually for each important outcome that is relevant to the guideline review question under study. To avoid biasing your review, you should identify the relevant outcomes **before** considering the results of the study. Clinical input may be required to identify the most important outcomes for a review.

***D1. The study had an appropriate length of follow-up***

The follow-up of participants after treatment should be of an adequate length to identify the outcome of interest. This is particularly important when different outcomes of interest occur early and late after an intervention. For example,

after surgical interventions there is usually an early harm because of side effects, with benefits apparent later on. A study that is too short will give an unbalanced assessment of the intervention.

For events occurring later, a short study will give an imprecise estimate of the effect, which may or may not also be biased. For example, a late-occurring side effect will not be detected in the treatment arm if the study is too short.

***D2. The study used a precise definition of outcome***
***D3. A valid and reliable method was used to determine the outcome***

The outcome under study should be well defined. It should be clear how the investigators determined whether participants experienced, or did not experience, the outcome. The same methods for defining and measuring outcomes should be used for all participants in the study. Often there may be more than one way of measuring an outcome (for example, physical or laboratory tests, questionnaire, reporting of symptoms). The method of measurement should be valid (that is, it measures what it claims to measure) and reliable (that is, it measures something consistently).

***D4. Investigators were kept 'blind' to participants' exposure to the intervention***
***D5. Investigators were kept 'blind' to other important confounding and prognostic factors***

In this context the 'investigators' are the individuals who are involved in making the decision about whether a participant has experienced the outcome under study. This can include those responsible for taking physical measurements and recording symptoms, even if they are not ultimately responsible for determining the outcome. Investigators can introduce bias through differences in measurement and recording of outcomes, and making biased assessments of a participant's outcome based on the collected data. The degree to which lack of blinding can introduce bias will vary depending on the method of measuring an outcome, but will be greater for more subjective outcomes, such as reporting of pain.

Physical separation of the assessment from the participant (for example, sending samples off to a laboratory) can often be considered as blind if it can be assumed that the laboratory staff are unaware of the treatment assignment.

# Appendix E: Methodology checklist: cohort studies

| | | | | | |
|---|---|---|---|---|---|
| **Study identification**<br>*Include author, title, reference, year of publication* | | | | | |
| **Guideline topic:** | | **Review question no:** | | | |
| **Checklist completed by:** | | | | | |
| | | ***Circle one option for each question:*** | | | |
| **A. Selection bias (systematic differences between the comparison groups)** | | | | | |
| A1 | The method of allocation to treatment groups was unrelated to potential confounding factors (that is, the reason for participant allocation to treatment groups is not expected to affect the outcome(s) under study) | Yes | No | Unclear | N/A |
| A2 | Were any attempts made within the design or analysis to balance the comparison groups for potential confounders? | Yes | No | Unclear | N/A |
| A3 | The groups were comparable at baseline, including all major confounding and prognostic factors | Yes | No | Unclear | N/A |
| Based on your answers to the above, in your opinion was selection bias present? If so, what is the likely direction of its effect? | | | | | |
| Low risk of bias | Unclear/unknown risk | High risk of bias | | | |
| Likely direction of effect: | | | | | |
| **B. Performance bias (systematic differences between groups in the care provided, apart from the intervention under investigation)** | | | | | |
| B1 | The comparison groups received the same care apart from the intervention(s) studied | Yes | No | Unclear | N/A |
| B2 | Participants receiving care were kept 'blind' to treatment allocation | Yes | No | Unclear | N/A |
| B3 | Individuals administering care were kept 'blind' to treatment allocation | Yes | No | Unclear | N/A |
| Based on your answers to the above, in your opinion was performance bias present? If so, what is the likely direction of its effect? | | | | | |
| Low risk of bias | Unclear/unknown risk | High risk of bias | | | |
| Likely direction of effect: | | | | | |

| | | | | | |
|---|---|---|---|---|---|
| **C. Attrition bias (systematic differences between the comparison groups with respect to loss of participants)** | | | | | |
| C1 | All groups were followed up for an equal length of time (or analysis was adjusted to allow for differences in length of follow-up) | Yes | No | Unclear | N/A |
| C2 | a. How many participants did not complete treatment in each group? | | | | |
| | b. The groups were comparable for treatment completion (that is, there were no important or systematic differences between groups in terms of those who did not complete treatment) | Yes | No | Unclear | N/A |
| C3 | a. For how many participants in each group were no outcome data available? | | | | |
| | b. The groups were comparable with respect to the availability of outcome data (that is, there were no important or systematic differences between groups in terms of those for whom outcome data were not available) | Yes | No | Unclear | N/A |
| Based on your answers to the above, in your opinion was attrition bias present? If so, what is the likely direction of its effect? | | | | | |
| Low risk of bias | Unclear/unknown risk | High risk of bias | | | |
| Likely direction of effect: | | | | | |
| **D. Detection bias (bias in how outcomes are ascertained, diagnosed or verified)** | | | | | |
| D1 | The study had an appropriate length of follow-up | Yes | No | Unclear | N/A |
| D2 | The study used a precise definition of outcome | Yes | No | Unclear | N/A |
| D3 | A valid and reliable method was used to determine the outcome | Yes | No | Unclear | N/A |
| D4 | Investigators were kept 'blind' to participants' exposure to the intervention | Yes | No | Unclear | N/A |
| D5 | Investigators were kept 'blind' to other important confounding/prognostic factors | Yes | No | Unclear | N/A |
| Based on your answers to the above, in your opinion was detection bias present? If so, what is the likely direction of its effect? | | | | | |
| Low risk of bias | Unclear/unknown risk | High risk of bias | | | |
| Likely direction of effect: | | | | | |

## *Notes on use of Methodology checklist: cohort studies*

Cohort studies are designed to answer questions about the relative effects of interventions, such as drugs, psychological therapies, operations or placebos. Such studies can include comparisons of 'test and treat strategies' involving a diagnostic test and subsequent management. This checklist does not cover comparisons of diagnostic test accuracy or questions about prognosis.

This checklist replaces the methodology checklist for cohort studies from 'The guidelines manual 2007' (appendix D).

Some of the items on this checklist may need to be filled in individually for different outcomes reported by the study. It is therefore important that the systematic reviewer has a clear idea of what the important outcomes are **before** appraising a study. You are likely to need input from the Guideline Development Group in defining the important outcomes.

Checklist items are worded so that a 'yes' response always indicates that the study has been designed/conducted in such a way as to minimise the risk of bias for that item. An 'unclear' response to a question may arise when the item is not reported or is not reported clearly. 'N/A' should be used when a cohort study cannot give an answer of 'yes' no matter how well it has been done.

This checklist is designed to assess the internal validity of the study; that is, whether the study provides an unbiased estimate of what it claims to show. Internal validity implies that the differences observed between groups of participants allocated to different interventions may (apart from the possibility of random error) be attributed to the intervention under investigation. Biases are characteristics that are likely to make estimates of effect differ systematically from the truth.

*Recording the presence and direction of bias*

This checklist contains four sections (A–D), each of which addresses a potential source of bias relating to internal validity. At the end of each section you are asked to give your opinion on whether bias is present, and to estimate the likely direction of this bias – whether you think it will have increased or decreased the effect size reported by the study. It will not always be possible to determine the direction of bias, but thinking this through can help greatly in interpreting results.

### A: Selection bias

Selection bias can be introduced into a study when there are systematic differences between the participants in the different treatment groups. As a result, the differences in the outcome observed may be explained by pre-existing differences between the groups rather than because of the treatment itself. For example, if the people in one group are in poorer health, then they are more likely to have a bad outcome than those in the other group, regardless of the effect of the treatment. The treatment groups should be

similar at the start of the study – the only difference between the groups should be in terms of the intervention received.

The main difference between randomised trials and non-randomised studies is the potential susceptibility of the latter to selection bias. Randomisation should ensure that, apart from the intervention received, the treatment groups differ only because of random variation. However, care needs to be taken in the design and analysis of non-randomised studies to take account of potential confounding factors. There are two main ways of accounting for potential confounding factors within non-randomised studies. Firstly, participants can be allocated to treatment groups to ensure that the groups are equal with respect to the known confounders. Secondly, statistical techniques can be used within the analysis to take into account known differences between groups. Neither of these approaches is able to address unknown or unmeasurable confounding factors, and it is important to remember that measurement of known confounders is subject to error. It can rarely, if ever, be assumed that all important factors relevant to prognosis and responsiveness to treatment are known. Hence, considerable judgement is needed to assess the internal validity of non-randomised studies; clinical input may be needed to identify potential confounding factors that should be taken into consideration.

### *A1. The method of allocation to treatment groups was unrelated to potential confounding factors*

In non-randomised studies, there will usually be a reason why participants are allocated to the treatment groups (often as a result of clinician and/or patient choice). If this reason is linked to the outcome under study, this can result in confounding by indication (where the decision to treat is influenced by some factor that is related in turn to the treatment outcome). For example, if the participants who are the most ill are selected for the treatment, then the treatment group may experience worse outcomes because of this difference between the groups at baseline. It will not always be possible to determine from the report of a study which factors influenced the allocation of participants to treatment groups.

### *A2. Were any attempts made within the design or analysis to balance the comparison groups for potential confounders?*

This represents an attempt when designing the study to ensure that the groups are similar in terms of known confounding or prognostic factors, in order to optimise comparability between the treatment groups. For example, in a matched design, the controls are deliberately chosen to be equivalent to the treatment group for any potential confounding variables, such as age and sex.

An alternative approach is to use statistical techniques to adjust for known confounding factors in the analysis.

### *A3. The groups were comparable at baseline, including all major confounding and prognostic factors*

Studies may report the distributions of potential prognostic and confounding factors in the comparison groups, or important differences in these factors may be noted.

Formal tests comparing the groups are problematic – failure to detect a difference does not that mean a difference does not exist, and multiple comparisons of factors may falsely detect some differences that are not real.

Clinical input may be needed to determine whether all likely confounders have been considered. Confounding factors may differ according to outcome, so you will need to consider potential confounding factors for each of the outcomes that are of interest to your review.

## B: Performance bias

Performance bias refers to systematic differences in the care provided to the participants in the comparison groups, other than the intervention under investigation.

This may consist of additional treatment, advice or counselling, rather than a physical intervention, or even simply a belief about the effects of an intervention. If performance bias is present, it can be difficult to attribute any observed effect to the experimental treatment rather than to the other factors.

Performance bias can be more difficult to determine within non-randomised than within randomised studies, because the latter are likely to have been better planned and executed according to strict treatment protocols that specify standardised interventions and care. It may be particularly difficult to determine performance bias for retrospective studies, where there is usually no control over standardisation.

### *B1. The comparison groups received the same care apart from the intervention(s) studied*

There should be no differences between the treatment groups apart from the intervention received. If some participants received additional treatment (known as ‘co-intervention’), this treatment is a potential confounding factor that may compromise the results.

*Blinding*

Blinding (also known as masking) refers to the process of withholding information about treatment allocation or exposure status from those involved in the study who could potentially be influenced by this information. This can include participants, investigators, those administering care and those involved in data collection and analysis. If people are aware of the treatment allocation or exposure status (‘unblinded’), this can bias the results of studies, either intentionally or unintentionally, through the use of other effective co-interventions, decisions about withdrawal, differential reporting of symptoms

or influencing concordance with treatment. Blinding of those assessing outcomes is covered in section D on detection bias.

Blinding of participants and carers is not always possible, particularly in studies of non-drug interventions, and so performance bias may be a particular issue in these studies. It is important to think about the likely size and direction of bias caused by failure to blind.

The terms 'single blind', 'double blind' and even 'triple blind' are sometimes used in studies. Unfortunately, they are not always used consistently. Commonly, when a study is described as 'single blind', only the participants are blind to their group allocation. When both participants and investigators are blind to group allocation the study is often described as 'double blind'. It is preferable to record exactly who was blinded, if reported, to avoid misunderstanding.

### *B2. Participants receiving care were kept 'blind' to treatment allocation*

The knowledge of assignment to a particular treatment group may affect outcomes such as a study participant's reporting of symptoms, self-use of other known interventions or even dropping out of the study.

### *B3. Individuals administering care were kept 'blind' to treatment allocation*

If individuals who are administering the intervention and/or other care to the participant are aware of treatment allocation, they may treat participants receiving one treatment differently from those receiving the comparison treatment; for example, by offering additional co-interventions.

## C: Attrition bias

Attrition refers to the loss of participants during the course of a study. Attrition bias occurs when there are systematic differences between the comparison groups with respect to participants lost, or differences between the participants lost to the study and those who remain. Attrition can occur at any point after participants have been allocated to their treatment groups. As such, it includes participants who are excluded after allocation (and may indicate a violation of eligibility criteria), those who do not complete treatment (whether or not they continue measurement) and those who do not complete outcome measurement (regardless of whether or not treatment was completed). Consideration should be given to why participants dropped out, as well as how many. Participants who dropped out of a study may differ in some significant way from those who remained as part of the study throughout. Drop-out rates and reasons for dropping out should be similar across all treatment groups. The proportion of participants excluded after allocation should be stated in the study report and the possibility of attrition bias considered within the analysis; however, these are not always reported.

### *C1. All groups were followed up for an equal length of time (or analysis was adjusted to allow for differences in length of follow-up)*

If the comparison groups are followed for different lengths of time, then more events are likely to occur in the group followed up for longer, distorting the

comparison. This may be overcome by adjusting the denominator to take the time into account; for example by using person-years.

***C2a. How many participants did not complete treatment in each group?***

A very high number of participants dropping out of a study should give concern. The drop-out rate may be expected to be higher in studies conducted over a longer period of time. The drop-out rate includes people who did not even start treatment; that is, they were excluded from the study after allocation to treatment groups.

***C2b. The groups were comparable for treatment completion (that is, there were no important or systematic differences between groups in terms of those who did not complete treatment)***

If there are systematic differences between groups in terms of those who did not complete treatment, consider both why participants dropped out and whether any systematic differences in those who dropped out may be related to the outcome under study, such as potential confounders. Systematic differences between groups in terms of those who dropped out may also result in treatment groups that are no longer comparable with respect to potential confounding factors.

***C3a. For how many participants in each group were no outcome data available?***

A very high number of participants for whom no outcome data were available should give concern.

***C3b. The groups were comparable with respect to the availability of outcome data (that is, there were no important or systematic differences between groups in terms of those for whom outcome data were not available)***

If there are systematic differences between groups in terms of those for whom no outcome data were available, consider both why the outcome data were not available and whether there are any systematic differences between participants for whom outcome data were and were not available.

**D: Detection bias (this section should be completed individually for each important relevant outcome)**

The way outcomes are assessed needs to be standardised for the comparison groups; failure to 'blind' people who are assessing the outcomes can also lead to bias, particularly with subjective outcomes. Most studies report results for more than one outcome, and it is possible that detection bias may be present for some, but not all, outcomes. It is therefore recommended that this section is completed individually for each important outcome that is relevant to the guideline review question under study. To avoid biasing your review, you should identify the relevant outcomes **before** considering the results of the study. Clinical input may be required to identify the most important outcomes for a review.

### *D1. The study had an appropriate length of follow-up*

The follow-up of participants after treatment should be of an adequate length to identify the outcome of interest. This is particularly important when different outcomes of interest occur early and late after an intervention. For example, after surgical interventions there is usually early harm because of side effects, with benefits apparent later on. A study that is too short will give an unbalanced assessment of the intervention.

For events occurring later, a short study will give an imprecise estimate of the effect, which may or may not also be biased. For example, a late-occurring side effect will not be detected in the treatment arm if the study is too short.

### *D2. The study used a precise definition of outcome*
### *D3. A valid and reliable method was used to determine the outcome*

The outcome under study should be well defined and it should be clear how the investigators determined whether participants experienced, or did not experience, the outcome. The same methods for defining and measuring outcomes should be used for all participants in the study. Often there may be more than one way of measuring an outcome (for example, physical or laboratory tests, questionnaire, reporting of symptoms). The method of measurement should be valid (that is, it measures what it claims to measure) and reliable (that is, it measures something consistently).

### *D4. Investigators were kept 'blind' to participants' exposure to the intervention*
### *D5. Investigators were kept 'blind' to other important confounding and prognostic factors*

In this context the 'investigators' are the individuals who are involved in making the decision about whether a participant has experienced the outcome under study. This can include those responsible for taking physical measurements and recording symptoms, even if they are not ultimately responsible for determining the outcome. Investigators can introduce bias through differences in measurement and recording of outcomes, and making biased assessments of a participant's outcome based on the collected data. The degree to which lack of blinding can introduce bias will vary depending on the method of measuring an outcome, but will be greater for more subjective outcomes, such as reporting of pain.

Physical separation of the assessment from the participant (for example, sending samples off to a laboratory) can often be considered as blind if it can be assumed that the laboratory staff are unaware of the treatment assignment.

# Appendix F: Methodology checklist: case–control studies

| | | | |
|---|---|---|---|
| **Study identification**<br>*Include author, title, reference, year of publication* | | | |
| **Guideline topic:** | | **Review question no:** | |
| **Checklist completed by:** | | | |
| **Section 1: Internal validity** | | | |
| | | ***Circle one option for each question*** | |
| 1.1 | The study addresses an appropriate and clearly focused question. | Well covered<br>Adequately addressed<br>Poorly addressed | Not addressed<br>Not reported<br>Not applicable |
| Selection of participants | | | |
| 1.2 | The cases and controls are taken from comparable populations | Well covered<br>Adequately addressed<br>Poorly addressed | Not addressed<br>Not reported<br>Not applicable |
| 1.3 | The same exclusion criteria are used for both cases and controls | Well covered<br>Adequately addressed<br>Poorly addressed | Not addressed<br>Not reported<br>Not applicable |
| 1.4 | What was the participation rate for each group (cases and controls)? | Cases:<br>Controls: | |
| 1.5 | Participants and non-participants are compared to establish their similarities or differences | Well covered<br>Adequately addressed<br>Poorly addressed | Not addressed<br>Not reported<br>Not applicable |
| 1.6 | Cases are clearly defined and differentiated from controls | Well covered<br>Adequately addressed<br>Poorly addressed | Not addressed<br>Not reported<br>Not applicable |
| 1.7 | It is clearly established that controls are not cases | Well covered<br>Adequately addressed<br>Poorly addressed | Not addressed<br>Not reported<br>Not applicable |
| Assessment | | | |
| 1.8 | Measures were taken to prevent knowledge of primary exposure influencing case ascertainment | Well covered<br>Adequately addressed<br>Poorly addressed | Not addressed<br>Not reported<br>Not applicable |

| | | | |
|---|---|---|---|
| 1.9 | Exposure status is measured in a standard, valid and reliable way | Well covered<br>Adequately addressed<br>Poorly addressed | Not addressed<br>Not reported<br>Not applicable |
| Confounding factors | | | |
| 1.10 | The main potential confounders are identified and taken into account in the design and analysis | Well covered<br>Adequately addressed<br>Poorly addressed | Not addressed<br>Not reported<br>Not applicable |
| Statistical analysis | | | |
| 1.11 | Have confidence intervals been provided? | | |

| **Section 2: Description of the study**<br>**(This information is required for evidence tables to facilitate cross-study comparisons. Please complete all sections for which information is available).**<br>***Please print clearly*** | | |
|---|---|---|
| 2.1 | How many people participated in the study?<br>*List the numbers of cases and controls separately.* | |
| 2.2 | What are the main characteristics of the study population?<br>*Include all characteristics used to identify both cases and controls – for example, age, sex, social class, disease status.* | |
| 2.3 | What environmental or prognostic factor is being investigated? | |
| 2.4 | What comparisons are made?<br>*Normally only one factor will be compared, but in some cases the extent of exposure may be stratified – for example, non-smokers vs light, moderate or heavy smokers. Note all comparisons here.* | |
| 2.5 | For how long are participants followed up?<br>*This is the length of time over which participant histories are tracked in the study.* | |
| 2.6 | What outcome measure(s) is/are used?<br>*List all outcomes that are used to assess the impact of the chosen environmental or prognostic factor.* | |
| 2.7 | What size of effect is identified?<br>*Effect size should be expressed as an odds ratio. If any other measures are included, note them as well. Include p-values and any confidence intervals that are provided.* | |
| 2.8 | How was the study funded?<br>*List all sources of funding quoted in the article, whether government, voluntary sector or industry.* | |
| 2.9 | Does this study help to answer your guideline review question?<br>*Summarise the main conclusions of the study and indicate how it relates to the review question.* | |

## *Notes on use of the Methodology checklist: case–control studies*

Case–control studies are designed to answer questions of the type 'What are the factors that caused this event?'. They involve comparison of individuals who have an outcome with other individuals from the same population who do not have the outcome. These studies start after the outcome of an event, and can be used to assess multiple causes of a single event. They are generally used to assess the causes of a new problem but they may also be useful for the evaluation of population-based interventions such as screening.

The questions in **section 1** are aimed at establishing the internal validity of the study under review – that is, making sure that it has been carried out carefully, and that any link between events and outcomes is clearly established. Each question covers an aspect of methodology that has been shown to make a significant difference to the conclusions of a study.

Case–control studies need to be designed very carefully,– the complexity of their design is often not appreciated by investigators, and so many poor-quality studies are conducted. The questions in this checklist are designed to identify the main features that should be present in a well-designed study. There are few criteria that should, alone and unsupported, lead to rejection of a study. However, a study that fails to address or report on more than one or two of the questions in the checklist should almost certainly be rejected.

For each question in this section you should choose one of the following categories to indicate how well it has been addressed in the study:

- well covered
- adequately addressed
- poorly addressed
- not addressed (not mentioned, or this aspect of study design was ignored)
- not reported (mentioned, but with insufficient detail to allow assessment to be made)
- not applicable.

### *1.1 The study addresses an appropriate and clearly focused question*

Unless a clear and well-defined question is specified, it will be difficult to assess how well the study has met its objectives or how relevant it is to the question you are trying to answer.

### Selection of participants

### *1.2 The cases and controls are taken from comparable populations*

Study participants may be selected from the target population (all individuals to which the results of the study could be applied), from the source population (a defined subset of the target population from which participants are selected) or from a pool of eligible people (a clearly defined and counted group selected from the source population). A study that does not include clear definitions of the source population should be rejected.

### *1.3 The same exclusion criteria are used for both cases and controls*

All selection and exclusion criteria should be applied equally to cases and controls. Failure to do so may introduce a significant degree of bias into the results of the study.

### *1.4 What was the participation rate for each group (cases and controls)?*

Differences between the eligible population and the study participants are important because they may influence the validity of the study. A participation rate can be calculated by dividing the number of study participants by the number of people who are eligible to participate. It is more useful if it is calculated separately for cases and controls. If the participation rate is low, or there is a large difference in rate between cases and controls, the study results may be invalid because of differences between participants and non-participants. In these circumstances the study should be downgraded, and rejected if the differences are very large.

### *1.5 Participants and non-participants are compared to establish their similarities or differences*

Even if participation rates are comparable and acceptable, it is still possible that the participants selected to act as cases or controls may differ from other members of the source population in some significant way. A well-conducted case–control study will look at samples of those not participating among the source population to ensure that the participants are a truly representative sample.

### *1.6 Cases are clearly defined and differentiated from controls*

The method of selection of cases is of critical importance to the validity of the study. Investigators have to be certain that cases are truly cases, but must balance this with the need to ensure that the cases admitted into the study are representative of the eligible population. The issues involved in case selection are complex, and should ideally be evaluated by someone with a good understanding of the design of case–control studies. If there is no information on how cases were selected it is probably safest to reject the study as a source of evidence.

### *1.7 It is clearly established that controls are not cases*

Just as it is important to be sure that cases are true cases, it is important to be sure that controls do not have the outcome under investigation. Controls should be chosen so that information on exposure status can be obtained or assessed in a similar way to that used for the selection of cases. If the methods of control selection are not described, the study should be rejected. If different methods of selection are used for cases and controls, the study should be evaluated by someone with a good understanding of the design of case–control studies.

## Assessment

### *1.8 Measures were taken to prevent knowledge of primary exposure from influencing case ascertainment*

If there is a possibility that case ascertainment was influenced by knowledge of exposure status, assessment of any association is likely to be biased. A well-conducted study should take this into account in the design of the study.

### *1.9 Exposure status is measured in a standard, valid and reliable way*

The inclusion of evidence from other sources or previous studies that demonstrate the validity and reliability of the assessment methods, or that the measurement method is a recognised procedure, should increase confidence in study quality.

## Confounding factors

### *1.10 The main potential confounders are identified and taken into account in the design and analysis*

Confounding is the distortion of a link between exposure and outcome by another factor that is associated with both exposure and outcome. The possible presence of confounding factors is one of the principal reasons why observational studies are not more highly rated as a source of evidence. The report of the study should indicate which potential confounders have been considered, and how they have been assessed or accounted for in the analysis. Clinical judgement should be used to consider whether all likely confounders have been taken into account. If the measures used to address the potential effects of confounders are considered inadequate, the study should be downgraded or rejected, depending on how serious the risk of confounding is considered to be. A study that does not address the possibility of confounding should be rejected.

## Statistical analysis

### *1.11 Have confidence intervals been provided?*

Confidence intervals are the preferred method for indicating the precision of statistical results, and can be used to differentiate between an inconclusive study and a study that shows no effect. Studies that report a single value with no assessment of precision should be treated with caution.

**Section 2** of the checklist asks you to summarise key points about the study that will be added to an evidence table (see appendix K) in the next stage of the process.

# Appendix G: Methodology checklist: the QUADAS tool for studies of diagnostic test accuracy[5]

| **Study identification**<br>*Including author, title, reference, year of publication* | |
|---|---|
| **Guideline topic:** | **Review question no:** |
| **Checklist completed by:** | |
| | ***Circle one option for each question*** |
| Was the spectrum of participants representative of the patients who will receive the test in practice? | Yes No Unclear N/A |
| Were selection criteria clearly described? | Yes No Unclear N/A |
| Was the reference standard likely to classify the target condition correctly? | Yes No Unclear N/A |
| Was the period between performance of the reference standard and the index test short enough to be reasonably sure that the target condition did not change between the two tests? | Yes No Unclear N/A |
| Did the whole sample or a random selection of the sample receive verification using the reference standard? | Yes No Unclear N/A |
| Did participants receive the same reference standard regardless of the index test result? | Yes No Unclear N/A |
| Was the reference standard independent of the index test? (that is, the index test did not form part of the reference standard) | Yes No Unclear N/A |
| Was the execution of the index test described in sufficient detail to permit its replication? | Yes No Unclear N/A |
| Was the execution of the reference standard described in sufficient detail to permit its replication? | Yes No Unclear N/A |
| Were the index test results interpreted without knowledge of the results of the reference standard? | Yes No Unclear N/A |
| Were the reference standard results interpreted without knowledge of the results of the index test? | Yes No Unclear N/A |
| Were the same clinical data available when the test results were interpreted as would be available when the test is used in practice? | Yes No Unclear N/A |
| Were uninterpretable, indeterminate or intermediate test results reported? | Yes No Unclear N/A |
| Were withdrawals from the study explained? | Yes No Unclear N/A |

[5] Adapted from: Whiting P, Rutjes AW, Dinnes J et al. (2004) Development and validation of methods for assessing the quality of diagnostic accuracy studies. Health Technology Assessment 8: 1–234

## *Notes on use of Methodology checklist: studies of diagnostic test accuracy*

This checklist is designed for the evaluation of studies assessing the accuracy of specific diagnostic tests. It does **not** address questions of the usefulness of the test in practice, or how the test compares with alternatives. Such questions should be assessed using the checklists for studies on interventions (see appendices D, E and F).

The questions in this checklist are aimed at establishing the validity of the study under review – that is, making sure that it has been carried out carefully, and that the conclusions represent an unbiased assessment of the accuracy and reliability of the test being evaluated. Each question covers an aspect of methodology that is thought to make a difference to the reliability of a study.

Checklist items are worded so that a 'yes' response always indicates that the study has been designed and conducted in such a way as to minimise the risk of bias for that item. An 'unclear' response to a question may arise when the answer to an item is not reported, or not reported clearly. 'N/A' should be used when a study of diagnostic test accuracy cannot give an answer of 'yes' no matter how well it has been done.

### Was the spectrum of participants representative of the patients who will receive the test in practice?

#### *What is meant by this item*

Differences between populations in demographic and clinical features may produce measures of diagnostic accuracy that vary considerably; this is known as spectrum bias. Reported estimates of diagnostic test accuracy may have limited clinical applicability (generalisability) if the spectrum of participants tested is not representative of the patients on whom the test will be used in practice. The spectrum of participants takes into account not only the severity of the underlying target condition but also demographic features and the presence of differential diagnoses and/or comorbidities.

#### *How to score this item*

Studies should score 'yes' for this item if you believe, based on the information reported, that the spectrum of participants included in the study was representative of those in whom the test will be used in practice. This judgement should be based on both the method for recruitment and the characteristics of those recruited. Studies that recruited a group of healthy controls and a group known to have the target disorder will be coded as 'no' on this item in nearly all circumstances. Reviewers should pre-specify what spectrum of participants would be acceptable, taking into account factors such as disease prevalence and severity, age and sex. Clinical input may be required from the Guideline Development Group (GDG). If you think that the population studied does not fit into what you specified as acceptable, the study should be scored as 'no'. If there is insufficient information available to make a judgement, this item should be scored as 'unclear'.

**Were selection criteria clearly described?**

***What is meant by this item***

This refers to whether studies have reported criteria for entry into the study.

***How to score this item***

If you think that all relevant information regarding how participants were selected for inclusion in the study has been provided, then this item should be scored as 'yes'. If study selection criteria are not clearly reported, then this item should be scored as 'no'. In situations where selection criteria are partially reported and you feel that you do not have enough information to score this item as 'yes', then it should be scored as 'unclear'.

**Was the reference standard likely to classify the target condition correctly?**

***What is meant by this item***

The reference standard is the method used to determine the presence or absence of the target condition. Indicators of diagnostic test accuracy are calculated by comparing the results of the index test with the results of the reference standard. Estimates of test performance are based on the assumption that the index test is being compared with a reference standard that is 100% sensitive and specific. If there are any disagreements between the reference standard and the index test, it is assumed that the index test is incorrect. Thus the use of an inappropriate reference standard can bias estimation of the diagnostic accuracy of the index test.

***How to score this item***

Making a judgement about the accuracy of the reference standard may not be straightforward. You may need to consult a member of the GDG to determine whether a test is an appropriate reference standard. If a combination of tests is used, you may have to consider carefully whether these were appropriate.

If you believe that the reference standard is likely to classify the target condition correctly, then this item should be scored as 'yes'. If you do not think that the reference standard is likely to have classified the target condition correctly, then this item should be scored as 'no'. If there is insufficient information to make a judgement, then it should be scored as 'unclear'.

**Was the period between performance of the reference standard and the index test short enough to be reasonably sure that the target condition did not change between the two tests?**

***What is meant by this item***

Ideally, the results of the index test and the reference standard are collected on the same participants at the same time. If this is not possible and there is a delay, misclassification may occur because of either spontaneous recovery or progression of the disease. This is known as disease progression bias. The length of the period that may cause such bias will vary between conditions. For example, a delay of a few days is unlikely to be a problem for chronic

conditions. However, for infectious diseases a delay of only a few days between performance of the index test and the reference standard may be important. This type of bias may also occur in chronic conditions in which the reference standard involves clinical follow-up of several years.

You will have to make judgements about what is considered 'short enough'. You should think about this **before** beginning your review, and define what you consider to be short enough for the specific topic area that you are reviewing. You may need clinical input to decide this.

### *How to score this item*

When to score this item as 'yes' is related to the target condition. For conditions that progress rapidly, a delay of a even few days may be important. For such conditions this item should be scored as 'yes' if the delay between the performance of the index test and the reference standard is very short – a matter of hours or days. However, for chronic conditions, disease status is unlikely to change in a week, a month or even longer. For such conditions, longer delays between performance of the index test and reference standard may be scored as 'yes'. If you think that the period between the performance of the index test and the reference standard was sufficiently long that disease status may have changed between the performance of the two tests, then this item should be scored as 'no'. If insufficient information is provided, it should be scored as 'unclear'.

## Did the whole sample or a random selection of the sample receive verification using the reference standard?

### *What is meant by this item*

Partial verification bias (also known as work-up bias, [primary] selection bias or sequential ordering bias) occurs when not all of the study group receive confirmation of the diagnosis by a reference standard. If the results of the index test influence the decision to perform the reference standard, then biased estimates of test performance may arise. If participants are randomly selected to receive the reference standard, the overall diagnostic performance of the test is, in theory, unchanged. However, in most cases this selection is not random, possibly leading to biased estimates of the overall diagnostic accuracy. Partial verification bias generally only occurs in diagnostic cohort studies in which participants are tested using the index test before the reference standard.

### *How to score this item*

If it is clear from the study that all participants (or a random selection) who received the index test went on to receive verification of their disease status using a reference standard, even if this reference standard was not the same for all participants, then this item should be scored as 'yes'. If some of the participants who received the index test did not receive verification of their true disease state (or the selection was not random), then this item should be scored as 'no'. If this information is not reported, this item should be scored as 'unclear'.

**Did participants receive the same reference standard regardless of the index test result?**

***What is meant by this item***

Differential verification bias occurs when some of the index test results are verified by a different reference standard. This is a particular problem if these reference standards differ in their definition of the target condition; for example, histopathology of the appendix and natural history for the detection of appendicitis. This usually occurs when participants who test positive on the index test undergo a more accurate, often invasive, reference standard test than those with negative results on the index test. The link (correlation) between a particular (negative) test result and being verified by a less accurate reference standard can lead to biased estimates of test accuracy. Differential verification bias generally only occurs in diagnostic cohort studies in which all participants are tested using the index test before the reference standard is performed.

***How to score this item***

If it is clear that participants received verification of their true disease status using the same reference standard, then this item should be scored as 'yes'. If some participants received verification using a different reference standard, then this item should be scored as 'no'. If this information is not reported, this item should be scored as 'unclear'.

**Was the reference standard independent of the index test? (that is, the index test did not form part of the reference standard)**

***What is meant by this item***

When the result of the index test is used in establishing the final diagnosis, incorporation bias may occur. This incorporation will probably increase the amount of agreement between index test results and the outcome of the reference standard, and hence result in overestimation of the various measures of diagnostic accuracy. For example, a study investigating magnetic resonance imaging (MRI) for the diagnosis of multiple sclerosis could have a reference standard composed of clinical follow-up, cerebrospinal fluid analysis and MRI. In this case, the index test forms part of the reference standard. It is important to note that knowledge of the results of the index test does not automatically mean that these results are incorporated in the reference standard. This item will only apply when a composite reference standard is used to verify disease status. In such cases it is essential that a full definition of how disease status is verified and which tests form part of the reference standard is provided.

***How to score this item***

For studies in which a single reference standard is used, this item will not be relevant and should be scored as 'N/A'. If it is clear that the index test did not form part of the reference standard, then this item should be scored as 'yes'. If it appears that the index test formed part of the reference standard, then this item should be scored as 'no'. If this information is not reported, this item should be scored as 'unclear'.

**Was the execution of the index test described in sufficient detail to permit its replication?**
**Was the execution of the reference standard described in sufficient detail to permit its replication?**

***What is meant by these items***

A sufficiently detailed description of the execution of the index test and the reference standard is important for two reasons. Firstly, variation in measures of diagnostic accuracy can sometimes be traced back to differences in the execution of index tests and reference standards. Secondly, a clear and detailed description (or references) is needed to implement a certain test in another setting. If tests are executed in different ways then this would be expected to have an impact on test performance. The extent to which this would be expected to affect results depends on the type of test being investigated.

***How to score these items***

If the study reports sufficient details to permit replication of the index test and the reference standard, then these items should be scored as 'yes'. In other cases these items should be scored as 'no'. In situations where details of test performance are partially reported and you consider that you do not have enough information to score these items as 'yes', then they should be scored as 'unclear'.

**Were the index test results interpreted without knowledge of the results of the reference standard?**
**Were the reference standard results interpreted without knowledge of the results of the index test?**

***What is meant by these items***

This issue is similar to the blinding of the people who assess outcomes in intervention studies. Interpretation of the results of the index test may be influenced by knowledge of the results of the reference standard, and vice versa. This is known as review bias, and may lead to inflated measures of diagnostic test accuracy. The extent to which this can affect test results will be related to the degree of subjectivity in the interpretation of the test result – the more subjective the interpretation, the more likely that the interpreter can be influenced by the results of the index test in interpreting the results of the reference standard, and vice versa. It is therefore important to consider the topic area that you are reviewing and to determine whether interpretation of the results of the index test or the reference standard could be influenced by knowledge of the results of the other test.

***How to score these items***

If the study clearly states that the test results (index test or reference standard) were interpreted blind to the results of the other test, then these items should be scored as 'yes'. If this does not appear to be the case, then they should be scored as 'no'. If this information is not reported, these items should be scored as 'unclear'. If in the topic area that you are reviewing the index test is always performed first, then interpretation of the results of the

index test will usually be done without knowledge of the results of the reference standard. Similarly, if the reference standard is always performed first, then the results will be interpreted without knowledge of the results of the index test. In situations where one form of review bias does not apply, the item should be scored as 'N/A'. If interpretation of test results is entirely objective, then test interpretation is not susceptible to review bias and the item should be scored as 'N/A'. Another situation in which this form of bias may not apply is when test results are interpreted in an independent laboratory. In such situations it is unlikely that the person interpreting the test results will have knowledge of the results of the other test (either index test or reference standard).

**Were the same clinical data available when the test results were interpreted as would be available when the test is used in practice?**

***What is meant by this item***

The availability of information on clinical data during the interpretation of test results may affect estimates of test performance. In this context, clinical data are defined broadly to include any information relating to the participant that is obtained by direct observation, such as age, sex and symptoms. The knowledge of such factors can influence the diagnostic test result if the test involves an interpretative component. If clinical data will be available when the test is interpreted in practice, then these should also be available when the test is evaluated. However, if the index test is intended to replace other clinical tests, then clinical data should not be available. Thus, before assessing studies for this item it is important to determine what information will be available when test results are interpreted in practice. You should consult the GDG to identify this information.

***How to score this item***

If clinical data would normally be available when the test results are interpreted in practice and similar data were available when interpreting the index test results in the study, then this item should be scored as 'yes'. Similarly, if clinical data would not be available in practice and these data were not available when the index test results were interpreted, then this item should be scored as 'yes'. If this is not the case, then this item should be scored as 'no'. If this information is not reported, this item should be scored as 'unclear'. If interpretation of the index test is fully automated, this item may not be relevant and can be scored 'N/A'.

**Were uninterpretable, indeterminate or intermediate test results reported?**

***What is meant by this item***

A diagnostic test can produce an uninterpretable, indeterminate or intermediate result with varying frequency, depending on the test. These problems are often not reported in studies on diagnostic test accuracy, the uninterpretable results simply being removed from the analysis. This may lead to the biased assessment of the test characteristics. Whether bias will arise depends on the possible correlation between uninterpretable test results and

the true disease status. If uninterpretable results occur randomly and are not related to the true disease status of the individual then, in theory, these should not have any effect on test performance. It is important that uninterpretable results are reported so that the impact on test performance can be considered; however, poor quality of reporting means that this is not always the case.

### *How to score this item*

If it is clear that all test results, including uninterpretable, indeterminate or intermediate results, are reported, then this item should be scored as 'yes'. If the authors do not report any uninterpretable, indeterminate or intermediate results, and if the results are reported for all participants who were described as having been entered into the study, then this item should also be scored as 'yes'. If you think that such results occurred but have not been reported, then this item should be scored as 'no'. If it is not clear whether all study results have been reported, then this item should be scored as 'unclear'.

## Were withdrawals from the study explained?

### *What is meant by this item*

This occurs when participants withdraw from the study before the results of both the index test and the reference standard are known. If participants lost to follow-up differ systematically from those who remain, for whatever reason, then estimates of test performance may be biased. Poor quality of reporting of withdrawals may make the impact on estimates of test performance difficult to determine.

### *How to score this item*

If it is clear what happened to all participants who entered the study, for example if a flow diagram of study participants is reported, then this item should be scored as 'yes'. If the authors do not report any withdrawals and if results are available for all participants who were reported to have been entered into the study, then this item should also be scored as 'yes'. If it appears that some of the participants who entered the study did not complete the study (that is, did not receive both the index test and the reference standard), and these participants were not accounted for, then this item should be scored as 'no'. If it is not clear whether all participants who entered the study were accounted for, then this item should be scored as 'unclear'.

# Appendix H: Methodology checklist: economic evaluations

This checklist is designed to determine whether an economic evaluation provides evidence that is useful to inform the decision-making of the Guideline Development Group (GDG) (see chapter 7). It is not intended to judge the quality of the study per se or the quality of reporting.

| | | |
|---|---|---|
| **Study identification**<br>*Including author, title, reference, year of publication* | | |
| **Guideline topic:** | | **Question no:** |
| **Checklist completed by:** | | |
| | | |
| **Section 1: Applicability (relevance to specific guideline review question(s) and the NICE reference case[6])**<br>*This checklist should be used first to filter out irrelevant studies.* | **Yes/ Partly/ No /Unclear /NA** | **Comments** |
| 1.1 Is the study population appropriate for the guideline? | | |
| 1.2 Are the interventions appropriate for the guideline? | | |
| 1.3 Is the healthcare system in which the study was conducted sufficiently similar to the current UK NHS context? | | |
| 1.4 Are costs measured from the NHS and personal social services (PSS) perspective? | | |
| 1.5 Are all direct health effects on individuals included? | | |
| 1.6 Are both costs and health effects discounted at an annual rate of 3.5%? | | |
| 1.7 Is the value of health effects expressed in terms of quality-adjusted life years (QALYs)? | | |
| 1.8 Are changes in health-related quality of life (HRQoL) reported directly from patients and/or carers? | | |
| 1.9 Is the valuation of changes in HRQoL (utilities) obtained from a representative sample of the general public? | | |
| 1.10 **Overall judgement**: Directly applicable/Partially applicable/Not applicable<br>There is no need to use section 2 of the checklist if the study is considered 'not applicable'. | | |
| Other comments: | | |

[6] As detailed in the 'Guide to the methods of technology appraisal' (June 2008), box 5.1 (page 30). Section 5.2.3 of the guide states: 'There may be important barriers to applying reference-case methods. In these cases, the reasons for a failure to meet the reference case should be clearly specified and justified, and the likely implications should, as far as possible, be quantified.'

| **Section 2: Study limitations (the level of methodological quality)**<br>*This checklist should be used once it has been decided that the study is sufficiently applicable to the context of the clinical guideline*[7]. | **Yes/ Partly /No/ Unclear/ NA** | **Comments** |
|---|---|---|
| 2.1 Does the model structure adequately reflect the nature of the health condition under evaluation? | | |
| 2.2 Is the time horizon sufficiently long to reflect all important differences in costs and outcomes? | | |
| 2.3 Are all important and relevant health outcomes included? | | |
| 2.4 Are the estimates of baseline health outcomes from the best available source? | | |
| 2.5 Are the estimates of relative treatment effects from the best available source? | | |
| 2.6 Are all important and relevant costs included? | | |
| 2.7 Are the estimates of resource use from the best available source? | | |
| 2.8 Are the unit costs of resources from the best available source? | | |
| 2.9 Is an appropriate incremental analysis presented or can it be calculated from the data? | | |
| 2.10 Are all important parameters whose values are uncertain subjected to appropriate sensitivity analysis? | | |
| 2.11 Is there no potential conflict of interest? | | |
| 2.12 **Overall assessment**: Minor limitations/Potentially serious limitations/Very serious limitations | | |
| Other comments: | | |

[7] The items and notes in this checklist have been developed from guidance in the NICE 'Guide to the methods of technology assessment' (June 2008), Evers S, Goossens M, de Vet H et al. (2005) Criteria list for assessment of methodological quality of economic evaluations – CHEC. International Journal of Technology Assessment in Health Care 21:240–5 and Philips Z, Ginnelly L, Sculpher M et al. (2004) Review of guidelines for good practice in decision-analytic modelling in health technology assessment. Health Technology Assessment 8. Available from: www.hta.ac.uk/fullmono/mon836.pdf .

## *Notes on use of Methodology checklist: economic evaluations*

For all questions:

- answer ‘yes’ if the study fully meets the criterion
- answer ‘partly’ if the study largely meets the criterion but differs in some important respect
- answer ‘no’ if the study deviates substantively from the criterion
- answer ‘unclear’ if the report provides insufficient information to judge whether the study complies with the criterion
- answer ‘NA (not applicable)’ if the criterion is not relevant in a particular instance.

For ‘partly’ or ‘no’ responses, use the comments column to explain how the study deviates from the criterion.

### Section 1: applicability

#### *1.1 Is the study population appropriate for the guideline?*

The study population should be defined as precisely as possible and should be in line with that specified in the guideline scope and any related review protocols.

This includes consideration of appropriate subgroups that require special attention. For many interventions, the capacity to benefit will differ for participants with differing characteristics. This should be explored separately for each relevant subgroup as part of the base-case analysis by the provision of estimates of clinical and cost effectiveness. The characteristics of participants in each subgroup should be clearly defined and, ideally, should be identified on the basis of an a priori expectation of differential clinical or cost effectiveness as a result of biologically plausible known mechanisms, social characteristics or other clearly justified factors.

Answer ‘yes’ if the study population is fully in line with that in the guideline question(s) and if the study differentiates appropriately between important subgroups. Answer ‘partly’ if the study population is similar to that in the guideline question(s) but: (i) it differs in some important respects; or (ii) the study fails to differentiate between important subgroups. Answer ‘no’ if the study population is substantively different from that in the guideline question(s).

#### *1.2 Are the interventions appropriate for the guideline?*

All relevant alternatives should be included, as specified in the guideline scope and any related review protocols. These should include routine and best practice in the NHS, existing NICE guidance and other feasible options.

Answer ‘yes’ if the analysis includes all options considered relevant for the guideline, even if it also includes other options that are not relevant. Answer ‘partly’ if the analysis omits one or more relevant options but still contains

comparisons likely to be useful for the guideline. Answer 'no' if the analysis does not contain any relevant comparisons.

### *1.3 Is the healthcare system in which the study was conducted sufficiently similar to the current UK NHS context?*

This relates to the overall structure of the healthcare system within which the interventions were delivered. For example, an intervention might be delivered on an inpatient basis in one country whereas in the UK it would be provided in the community. This might significantly influence the use of healthcare resources and costs, thus limiting the applicability of the results to a UK setting. In addition, old UK studies may be severely limited in terms of their relevance to current NHS practice.

Answer 'yes' if the study was conducted within the UK and is sufficiently recent to reflect current NHS practice. For non-UK or older UK studies, answer 'partly' if differences in the healthcare setting are unlikely to substantively change the cost-effectiveness estimates. Answer 'no' if the healthcare setting is so different that the results are unlikely to be applicable in the current NHS.

### *1.4 Are costs measured from the NHS and personal social services (PSS) perspective?*

The decision-making perspective of an economic evaluation determines the range of costs that should be included in the analysis. NICE works in a specific context; in particular, it does not set the budget for the NHS. The objective of NICE is to offer guidance that represents an efficient use of available NHS and PSS resources. For these reasons, the perspective on costs used in the NICE reference case is that of the NHS and PSS. Productivity costs and costs borne by patients and carers that are not reimbursed by the NHS or PSS are not included in the reference case. The reference case also excludes costs to other government bodies, although these may sometimes be presented in additional analyses alongside the reference case.

Answer 'yes' if the study only includes costs for resource items that would be paid for by the NHS and PSS. Also answer 'yes' if other costs have been included in the study, but the results are presented in such a way that the cost effectiveness can be calculated from an NHS and PSS perspective. Answer 'partly' if the study has taken a wider perspective but the other non-NHS/PSS costs are small in relation to the total expected costs and are unlikely to change the cost-effectiveness results. Answer 'no' if non-NHS/PSS costs are significant and are likely to change the cost-effectiveness results.

Some interventions may have a substantial impact on non-health outcomes or costs to other government bodies (for example, treatments to reduce illicit drug misuse may have the effect of reducing drug-related crime). In such situations, if the economic study includes non-health costs in such a way that they cannot be separated out from NHS/PSS costs, answer 'no' but consider retaining the study for critical appraisal. If studies containing non-reference-case costs are retained, use the comments column to note why.

### *1.5 Are all direct health effects on individuals included?*

In the NICE reference case, the perspective on outcomes should be all direct health effects, whether for patients or, when relevant, other people (principally carers). This is consistent with an objective of maximising health gain from available healthcare resources. Some features of healthcare delivery that are often referred to as 'process characteristics' may ultimately have health consequences; for example, the mode of treatment delivery may have health consequences through its impact on concordance with treatment. Any significant characteristics of healthcare technologies that have a value to people that is independent of any direct effect on health should be noted. These characteristics include the convenience with which healthcare is provided and the level of information available for patients.

This question should be viewed in terms of what is **excluded** in relation to the NICE reference case; that is, non-health effects.

Answer 'yes' if the measure of health outcome used in the analysis excludes non-health effects (or if such effects can be excluded from the results). Answer 'partly' if the analysis includes some non-health effects but these are small and unlikely to change the cost-effectiveness results. Answer 'no' if the analysis includes significant non-health effects that are likely to change the cost-effectiveness results.

### *1.6 Are both costs and health effects discounted at an annual rate of 3.5%?*

The need to discount to a present value is widely accepted in economic evaluation, although the specific rate varies across jurisdictions and over time. NICE considers it appropriate to discount costs and health effects at the same rate. The annual rate of 3.5%, based on the recommendations of the UK Treasury for the discounting of costs, applies to both costs and health effects.

Answer 'yes' if both costs and health effects (for example, quality-adjusted life years [QALYs]) are discounted at 3.5% per year. Answer 'partly' if costs and effects are discounted at a rate similar to 3.5% (for example, costs and effects are both discounted at 3% per year). Answer 'no' if costs and/or health effects are not discounted, or if they are discounted at a rate (or rates) different from 3.5% (for example, 5% for both costs and effects, or 6% for costs and 1.5% for effects). Note in the comments column what discount rates have been used. If all costs and health effects accrue within a short time (roughly a year), answer 'NA'.

### *1.7 Is the value of health effects expressed in terms of quality-adjusted life years (QALYs)?*

The QALY is a measure of a person's length of life weighted by a valuation of their health-related quality of life (HRQoL) over that period.

Given its widespread use, the QALY is considered by NICE to be the most appropriate generic measure of health benefit that reflects both mortality and effects on HRQoL. It is recognised that alternative measures exist (such as

the healthy-year equivalent), but few economic evaluations have used these methods and their strengths and weaknesses are not fully established.

NICE's position is that an additional QALY should be given the same weight regardless of the other characteristics of the patients receiving the health benefit.

Answer 'yes' if the effectiveness of the intervention is measured using QALYs; answer 'no' if not. There may be circumstances when a QALY cannot be obtained or where the assumptions underlying QALYs are considered inappropriate. In such situations answer 'no', but consider retaining the study for appraisal. Similarly, answer 'no' but retain the study for appraisal if it does not include QALYs but it is still thought to be useful for GDG decision-making: for example, if the clinical evidence indicates that an intervention might be dominant, and estimates of the relative costs of the interventions from a cost-minimisation study are likely to be useful. When economic evaluations not using QALYs are retained for full critical appraisal, use the comments column to note why.

### *1.8 Are changes in health-related quality of life (HRQoL) reported directly from patients and/or carers?*

In the NICE reference case, information on changes in HRQoL as a result of treatment should be reported directly by patients (and directly by carers when the impact of treatment on the carer's health is also important). When it is not possible to obtain information on changes in patients' HRQoL directly from them, data should be obtained from carers (not from healthcare professionals).

For consistency, the EQ-5D is NICE's preferred measure of HRQoL in adults. However, when EQ-5D data are not available or are inappropriate for the condition or the effects of treatment, other multi-attribute utility questionnaires (for example, SF6D, QWB or HUI) or mapping methods from disease-specific questionnaires may be used to estimate QALYs. For studies not reporting QALYs, a variety of generic or disease-specific methods may be used to measure HRQoL.

Answer 'yes' if changes in patients' HRQoL are estimated by the patients themselves. Answer 'partly' if estimates of patients' HRQoL are provided by carers. Answer 'no' if estimates come from healthcare professionals or researchers. Note in the comments column how HRQoL was measured (EQ-5D, QWB, HUI and so on). Answer 'NA' if the cost-effectiveness study does not include estimates of HRQoL (for example, studies reporting 'cost per life year gained' or cost-minimisation studies).

### *1.9 Is the valuation of changes in HRQoL (utilities) obtained from a representative sample of the general public?*

The NICE reference case specifies that the valuation of changes in HRQoL (utilities) reported by patients should be based on public preferences elicited using a choice-based method (such as the time trade-off or standard gamble) in a representative sample of the UK population.

Answer 'yes' if HRQoL valuations were obtained using the EQ-5D UK tariff. Answer 'partly' if the valuation methods were comparable to those used for the EQ-5D. Answer 'no' if other valuation methods were used. Answer 'NA' if the study does not apply valuations to HRQoL (for studies not reporting QALYs). In the comments column note the valuation method used (such as time trade-off or standard gamble) and the source of the preferences (such as patients or healthcare professionals).

### *1.10 Overall judgement*

Classify the applicability of the economic evaluation to the clinical guideline, the current NHS situation and the context for NICE guidance as one of the following:

- **Directly applicable** – the study meets all applicability criteria, or fails to meet one or more applicability criteria but this is unlikely to change the conclusions about cost effectiveness.
- **Partially applicable** – the study fails to meet one or more applicability criteria, and this could change the conclusions about cost effectiveness.
- **Not applicable** – the study fails to meet one or more applicability criteria, and this is likely to change the conclusions about cost effectiveness. Such studies would be excluded from further consideration and there is no need to continue with the rest of the checklist.

## Section 2: study limitations

### *2.1 Does the model structure adequately reflect the nature of the health condition under evaluation?*

This relates to the choice of model and its structural elements (including cycle length in discrete time models, if appropriate). Model type and its structural aspects should be consistent with a coherent theory of the health condition under evaluation. The selection of treatment pathways, whether health states or branches in a decision tree, should be based on the underlying biological processes of the health issue under study and the potential impact (benefits and adverse consequences) of the intervention(s) of interest.

Answer 'yes' if the model design and assumptions appropriately reflect the health condition and intervention(s) of interest. Answer 'partly' if there are aspects of the model design or assumptions that do not fully reflect the health condition or intervention(s) but that are unlikely to change the cost-effectiveness results. Answer 'no' if the model omits some important aspect of the health condition or intervention(s) and this is likely to change the cost-effectiveness results. Answer 'NA' for economic evaluations based on data from a clinical study which do not extrapolate treatment outcomes or costs beyond the study context or follow-up period.

### *2.2 Is the time horizon sufficiently long to reflect all important differences in costs and outcomes?*

The time horizon is the period of analysis of the study: the length of follow-up for participants in a trial-based evaluation, or the period of time over which the costs and outcomes for a cohort are tracked in a modelling study. This time

horizon should always be the same for costs and outcomes, and should be long enough to include all relevant costs and outcomes relating to the intervention. A time horizon shorter than lifetime could be justified if there is no differential mortality effect between options, and the differences in costs and HRQoL relate to a relatively short period (for example, in the case of an acute infection).

Answer 'yes' if the time horizon is sufficient to include all relevant costs and outcomes. Answer 'partly' if the time horizon may omit some relevant costs and outcomes but these are unlikely to change the cost-effectiveness results. Answer 'no' if the time horizon omits important costs and outcomes and this is likely to change the cost-effectiveness results.

### *2.3 Are all important and relevant health outcomes included?*

All relevant health outcomes should include direct health effects relating to harms from the intervention (adverse effects) as well as any potential benefits.

Answer 'yes' if the analysis includes all relevant and important harms and benefits. Answer 'partly' if the analysis omits some harms or benefits but these would be unlikely to change the cost-effectiveness results. Answer 'no' if the analysis omits important harms and/or benefits that would be likely to change the cost-effectiveness results.

### *2.4 Are the estimates of baseline health outcomes from the best available source?*

The estimate of the overall net treatment effect of an intervention is determined by the baseline risk of a particular condition or event and/or the relative effects of the intervention compared with the relevant comparator treatment. The overall net treatment effect may also be determined by other features of the people comprising the population of interest.

The process of assembling evidence for economic evaluations should be systematic – evidence must be identified, quality assessed and, when appropriate, pooled, using explicit criteria and justifiable and reproducible methods. These principles apply to all categories of evidence that are used to estimate clinical and cost effectiveness, evidence for which will typically be drawn from a number of different sources.

The sources and methods for eliciting baseline probabilities should be described clearly. These data can be based on 'natural history' (patient outcomes in the absence of treatment or with routine care), sourced from cohort studies. Baseline probabilities may also be derived from the control arms of experimental studies. Sometimes it may be necessary to rely on expert opinion for particular parameters.

Answer 'yes' if the estimates of baseline health outcomes reflect the best available evidence as identified from a recent well-conducted systematic review of the literature. Answer 'partly' if the estimates are not derived from a systematic review but are likely to reflect outcomes for the relevant group of patients in routine NHS practice (for example, if they are derived from a large

UK-relevant cohort study). Answer 'no' if the estimates are unlikely to reflect outcomes for the relevant group in routine NHS practice.

### *2.5 Are the estimates of relative treatment effects from the best available source?*

The objective of the analysis of clinical effectiveness is to produce an unbiased estimate of the mean clinical effectiveness of the interventions being compared.

The NICE reference case indicates that evidence on outcomes should be obtained from a systematic review, defined as the systematic location, inclusion, appraisal and synthesis of evidence to obtain a reliable and valid overview of the data relating to a clearly formulated question.

Synthesis of outcome data through meta-analysis is appropriate provided that there are sufficient relevant and valid data obtained using comparable measures of outcome.

Head-to-head randomised controlled trials (RCTs) provide the most valid evidence of relative treatment effect. However, such evidence may not always be available. Therefore, data from non-randomised studies may be required to supplement RCT data. Any potential bias arising from the design of the studies used in the assessment should be explored and documented.

Data from head-to-head RCTs should be presented in the base-case analysis, if available. When head-to-head RCTs exist, evidence from indirect or mixed treatment comparison analyses may be presented if it is considered to add information that is not available from the head-to-head comparison. This indirect or mixed treatment comparison must be fully described and presented as additional to the base-case analysis. (A 'mixed treatment comparison' estimates effect sizes using both head-to-head and indirect comparisons.)

If data from head-to-head RCTs are not available, indirect treatment comparison methods should be used. (An 'indirect comparison' is a synthesis of data from a network of trials that compare the interventions of interest with other comparators.)

When multiple interventions are being assessed that have not been compared within a single RCT, data from a series of pairwise head-to-head RCTs should be presented. Consideration should also be given to presenting a combined analysis using a mixed treatment comparison framework if it is considered to add information that is not available from the head-to-head comparison.

Only indirect or mixed treatment comparison methods that preserve randomisation should be used. The principles of good practice for standard meta-analyses should also be followed in mixed and indirect treatment comparisons.

The methods and assumptions that are used to extrapolate short-term results to final outcomes should be clearly presented and there should be documentation of the reasoning underpinning the choice of survival function.

Evidence for the evaluation of diagnostic technologies should normally incorporate evidence on diagnostic accuracy. It is also important to incorporate the predicted changes in health outcomes and costs resulting from treatment decisions based on the test result. The general principles guiding the assessment of the clinical and cost effectiveness of diagnostic interventions should be the same as for other technologies. However, particular consideration of the methods of analysis may be required, particularly in relation to evidence synthesis. Evidence for the effectiveness of diagnostic technologies should include the costs and outcomes for people whose test results lead to an incorrect diagnosis, as well as for those who are diagnosed correctly.

As for other technologies, RCTs have the potential to capture the pathway of care involving diagnostic technologies, but their feasibility and availability may be limited. Other study designs should be assessed on the basis of their fitness for purpose, taking into consideration the aim of the study (for example, to evaluate outcomes, or to evaluate sensitivity and specificity) and the purpose of the diagnostic technology.

Answer ‘yes’ if the estimates of treatment effect appropriately reflect all relevant studies of the best available quality, as identified through a recent well-conducted systematic review of the literature. Answer ‘partly’ if the estimates of treatment effect are not derived from a systematic review but are similar in magnitude to the best available estimates (for example, if the economic evaluation is based on a single large study with treatment effects similar to pooled estimates from all relevant studies). Answer ‘no’ if the estimates of treatment effect are likely to differ substantively from the best available estimates.

### *2.6 Are all important and relevant costs included?*

Costs related to the condition of interest and incurred in additional years of life gained as a result of treatment should be included in the base-case analysis. This should include the costs of handling non-adherence to treatment and treating side effects. Costs that are considered to be unrelated to the condition or intervention of interest should be excluded. If introduction of the intervention requires additional infrastructure to be put in place, consideration should be given to including such costs in the analysis.

Answer ‘yes’ if all important and relevant resource use and costs are included given the perspective and the research question under consideration. Answer ‘partly’ if some relevant resource items are omitted but these are unlikely to affect the cost-effectiveness results. Answer ‘no’ if important resource items are omitted and these are likely to affect the cost-effectiveness results.

### *2.7 Are the estimates of resource use from the best available source?*

It is important to quantify the effect of the interventions on resource use in terms of physical units (for example, days in hospital or visits to a GP) and valuing those effects in monetary terms using appropriate prices and unit costs. Evidence on resource use should be identified systematically. When

expert opinion is used as a source of information, any formal methods used to elicit these data should be clearly reported.

Answer ‘yes’ if the estimates of resource use appropriately reflect all relevant evidence sources of the best available quality, as identified through a recent well-conducted systematic review of the literature. Answer ‘partly’ if the estimates of resource use are not derived from a systematic review but are similar in magnitude to the best available estimates. Answer ‘no’ if the estimates of resource use are likely to differ substantively from the best available estimates.

### *2.8 Are the unit costs of resources from the best available source?*

Resources should be valued using the prices relevant to the NHS and PSS. Given the perspective of the NICE reference case, it is appropriate for the financial costs relevant to the NHS/PSS to be used as the basis of costing, although these may not always reflect the full social opportunity cost of a given resource. A first point of reference in identifying costs and prices should be any current official listing published by the Department of Health and/or the Welsh Assembly Government.

When the acquisition price paid for a resource differs from the public list price (for example, pharmaceuticals and medical devices sold at reduced prices to NHS institutions), the public list price should be used in the base-case analysis. Sensitivity analysis should assess the implications of variations from this price. Analyses based on price reductions for the NHS will only be considered when the reduced prices are transparent and can be consistently available across the NHS, and if the period for which the specified price is available is guaranteed.

National data based on healthcare resource groups (HRGs) such as the Payment by Results tariff can be used when they are appropriate and available. However, data based on HRGs may not be appropriate in all circumstances (for example, when the definition of the HRG is broad, or the mean cost probably does not reflect resource use in relation to the intervention(s) under consideration). In such cases, other sources of evidence, such as micro-costing studies, may be more appropriate. When cost data are taken from the literature, the methods used to identify the sources should be defined. When several alternative sources are available, a justification for the costs chosen should be provided and discrepancies between the sources explained. When appropriate, sensitivity analysis should have been undertaken to assess the implications for results of using alternative data sources.

Answer ‘yes’ if resources are valued using up-to-date prices relevant to the NHS and PSS. Answer ‘partly’ if the valuations of some resource items differ from current NHS/PSS unit costs but this is unlikely to change the cost-effectiveness results. Answer ‘no’ if the valuations of some resource items differ substantively from current NHS/PSS unit costs and this is likely to change the cost-effectiveness results.

### *2.9 Is an appropriate incremental analysis presented or can it be calculated from the data?*

An appropriate incremental analysis is one that compares the expected costs and health outcomes of one intervention with the expected costs and health outcomes of the next-best non-dominated alternative.

Standard decision rules should be followed when combining costs and effects, and should reflect any situation where there is dominance or extended dominance. When there is a trade-off between costs and effects, the results should be presented as an incremental cost-effectiveness ratio (ICER): the ratio of the difference in mean costs to the difference in mean outcomes of a technology compared with the next best alternative. In addition to ICERs, expected net monetary or health benefits can be presented using values placed on a QALY gained of £20,000 and £30,000.

For cost-consequence analyses, appropriate incremental analysis can only be done by selecting one of the consequences as the primary measure of effectiveness.

Answer 'yes' if appropriate incremental results are presented, or if data are presented that allow the reader to calculate the incremental results. Answer 'no' if: (i) simple ratios of costs to effects are presented for each alternative compared with a standard intervention; or (ii) if options subject to simple or extended dominance are not excluded from the incremental analyses.

### *2.10 Are all important parameters whose values are uncertain subjected to appropriate sensitivity analysis?*

There are a number of potential selection biases and uncertainties in any evaluation (trial- or model-based) and these should be identified and quantified where possible. There are three types of bias or uncertainty to consider:

- Structural uncertainty – for example in relation to the categorisation of different states of health and the representation of different pathways of care. These structural assumptions should be clearly documented and the evidence and rationale to support them provided. The impact of structural uncertainty on estimates of cost effectiveness should be explored by separate analyses of a representative range of plausible scenarios.
- Source of values to inform parameter estimates – the implications of different estimates of key parameters (such as estimates of relative effectiveness) must be reflected in sensitivity analyses (for example, through the inclusion of alternative scenarios). Inputs must be fully justified, and uncertainty explored by sensitivity analysis using alternative input values.
- Parameter precision – uncertainty around the mean health and cost inputs in the model. Distributions should be assigned to characterise the uncertainty associated with the (precision of) mean parameter values. Probabilistic sensitivity analysis is preferred, as this enables the uncertainty associated with parameters to be simultaneously reflected in the results of the model. In non-linear decision models – when there is not a straight-line

relationship between inputs and outputs of a model (such as Markov models) – probabilistic methods provide the best estimates of mean costs and outcomes. Simple decision trees are usually linear.

The mean value, distribution around the mean, and the source and rationale for the supporting evidence should be clearly described for each parameter included in the model.

Evidence about the extent of correlation between individual parameters should be considered carefully and reflected in the probabilistic analysis. Assumptions made about the correlations should be clearly presented.

Answer 'yes' if an extensive sensitivity analysis was undertaken that explored all key uncertainties in the economic evaluation. Answer 'partly' if the sensitivity analysis failed to explore some important uncertainties in the economic evaluation. Answer 'no' if the sensitivity analysis was very limited and omitted consideration of a number of important uncertainties, or if the range of values or distributions around parameters considered in the sensitivity analysis were not reported.

### *2.11 Is there no potential conflict of interest?*

The BMJ defines competing interests for its authors as follows: "A competing interest exists when professional judgment concerning a primary interest (such as patients' welfare or the validity of research) may be influenced by a secondary interest (such as financial gain or personal rivalry). It may arise for the authors of a BMJ article when they have a financial interest that may influence, probably without their knowing, their interpretation of their results or those of others."

Whenever a potential financial conflict of interest is possible, this should be declared.

Answer 'yes' if the authors declare that they have no financial conflicts of interest. Answer 'no' if clear financial conflicts of interest are declared or apparent (for example, from the stated affiliation of the authors). Answer 'unclear' if the article does not indicate whether or not there are financial conflicts of interest.

### *2.12 Overall assessment*

The overall methodological study quality of the economic evaluation should be classified as one of the following:

- **Minor limitations** – the study meets all quality criteria, or the study fails to meet one or more quality criteria but this is unlikely to change the conclusions about cost effectiveness.
- **Potentially serious limitations** – the study fails to meet one or more quality criteria and this could change the conclusions about cost effectiveness.
- **Very serious limitations** – the study fails to meet one or more quality criteria and this is highly likely to change the conclusions about cost

effectiveness. Such studies should usually be excluded from further consideration.

## *Supporting references*

Guide to the methods of technology appraisal (June 2008). Available from: www.nice.org.uk/media/B52/A7/TAMethodsGuideUpdatedJune2008.pdf

Social value judgements: principles for the development of NICE guidance (July 2008). Second edition. Available from: www.nice.org.uk/aboutnice/howwework/socialvaluejudgements/socialvaluejudgements.jsp

Philips Z, Ginnelly L, Sculpher M et al. (2004) Review of guidelines for good practice in decision-analytic modelling in health technology assessment. Health Technology Assessment 8 (36). Available from: www.ncchta.org/project/1342.asp

Evers, S, Goossens M, de Vet H et al. (2005) Criteria list for assessment of methodological quality of economic evaluations: Consensus on Health Economic Criteria. International Journal of Technology Assessment in Health Care 21: 240–5. Available from: http://journals.cambridge.org/production/action/cjoGetFulltext?fulltextid=292675

Six workshops were held to enable NICE to explore and capture different perspectives on specific questions as part of the 2007 review of the 'Guide to the methods of technology appraisal'. Documents listed below include briefing papers that were produced to facilitate discussion at each of the workshops and working party meetings:

- costs
- diagnostic technologies
- evidence synthesis (indirect and mixed treatment comparisons)
- identifying subgroups and exploring heterogeneity
- threshold
- exploring uncertainty
- health-related utility measurement.

These documents are available from: www.nice.org.uk/aboutnice/howwework/devnicetech/technologyappraisalprocessguides/selectedfurtherreadingguidetothemethodsoftechnologyappraisal.jsp

# Appendix I: Methodology checklist: qualitative studies[8]

| | |
|---|---|
| **Study identification**<br>*Include author, title, reference, year of publication* | |
| **Guidance topic:** | **Key research question/aim:** |
| **Checklist completed by:** | |

| **Section 1: theoretical approach** | | |
|---|---|---|
| **1.1 Is a qualitative approach appropriate?**<br>*For example:*<br>• Does the research question seek to understand processes or structures, or illuminate subjective experiences or meanings?<br>• Could a quantitative approach better have addressed the research question? | ☐ Appropriate<br>☐ Inappropriate<br>☐ Not sure | Comments: |
| **1.2 Is the study clear in what it seeks to do?**<br>*For example:*<br>• Is the purpose of the study discussed – aims/objectives/research question(s)?<br>• Is there adequate/appropriate reference to the literature?<br>• Are underpinning values/assumptions/theory discussed? | ☐ Clear<br>☐ Unclear<br>☐ Mixed | Comments: |

[8] This checklist is based on checklists in:

Spencer L. Ritchie J, Lewis J, Dillon L (2003) Quality in qualitative evaluation: a framework for assessing research evidence. London: Government Chief Social Researcher's Office. Available from: www.strategy.gov.uk/downloads/su/qual/downloads/qqe_rep.pdf

Public Health Resource Unit England (2006) Critical Appraisal Skills Programme (CASP) – making sense of evidence: 10 questions to help you make sense of qualitative research . Available from: www.phru.nhs.uk/Doc_Links/Qualitative%20Appraisal%20Tool.pdf

National Training and Research Appraisal Group (NTRAG); contact: www.ntrag.co.uk

British Sociological Association (BSA); contact: www.britsoc.co.uk

| **Section 2: study design** | | |
|---|---|---|
| **2.1 How defensible/rigorous is the research design/methodology?**<br>*For example:*<br>• Is the design appropriate to the research question?<br>• Is a rationale given for using a qualitative approach?<br>• Are there clear accounts of the rationale/justification for the sampling, data collection and data analysis techniques used?<br>• Is the selection of cases/sampling strategy theoretically justified? | ☐ Defensible<br>☐ Not defensible<br>☐ Not sure | Comments: |

| **Section 3: data collection** | | |
|---|---|---|
| **3.1 How well was the data collection carried out?**<br>*For example:*<br>• Are the data collection methods clearly described?<br>• Were the appropriate data collected to address the research question?<br>• Was the data collection and record keeping systematic? | ☐ Appropriate<br>☐ Inappropriate<br>☐ Not sure/ inadequately reported | Comments: |

| **Section 4: validity** | | |
|---|---|---|
| **4.1 Is the role of the researcher clearly described?**<br>*For example:*<br>• Has the relationship between the researcher and the participants been adequately considered?<br>• Does the paper describe how the research was explained and presented to the participants? | ☐ Clear<br>☐ Unclear<br>☐ Not described | Comments: |
| **4.2 Is the context clearly described?**<br>*For example:*<br>• Are the characteristics of the participants and settings clearly defined?<br>• Were observations made in a sufficient variety of circumstances?<br>• Was context bias considered? | ☐ Clear<br>☐ Unclear<br>☐ Not sure | Comments: |
| **4.3 Were the methods reliable?**<br>*For example:*<br>• Were data collected by more than one method?<br>• Is there justification for triangulation, or for not triangulating?<br>• Do the methods investigate what they claim to? | ☐ Reliable<br>☐ Unreliable<br>☐ Not sure | Comments: |

| **Section 5: analysis** | | |
|---|---|---|
| **5.1 Is the data analysis sufficiently rigorous?**<br>*For example:*<br>• Is the procedure explicit – is it clear how the data were analysed to arrive at the results?<br>• How systematic is the analysis – is the procedure reliable/dependable?<br>• Is it clear how the themes and concepts were derived from the data? | ☐ Rigorous<br>☐ Not rigorous<br>☐ Not sure/not reported | Comments: |
| **5.2 Are the data 'rich'?**<br>*For example:*<br>• How well are the contexts of the data described?<br>• Has the diversity of perspective and content been explored?<br>• How well have the detail and depth been demonstrated?<br>• Are responses compared and contrasted across groups/sites? | ☐ Rich<br>☐ Poor<br>☐ Not sure/not reported | Comments: |
| **5.3 Is the analysis reliable?**<br>*For example:*<br>• Did more than one researcher theme and code transcripts/data?<br>• If so, how were differences resolved?<br>• Did participants feed back on the transcripts/data? (if possible and relevant)<br>• Were negative/discrepant results addressed or ignored? | ☐ Reliable<br>☐ Unreliable<br>☐ Not sure/not reported | Comments: |
| **5.4 Are the findings convincing?**<br>*For example:*<br>• Are the findings clearly presented?<br>• Are the findings internally coherent?<br>• Are extracts from the original data included?<br>• Are the data appropriately referenced?<br>• Is the reporting clear and coherent? | ☐ Convincing<br>☐ Not convincing<br>☐ Not sure | Comments: |
| **5.5 Are the findings relevant to the aims of the study?** | ☐ Relevant<br>☐ Irrelevant<br>☐ Partially relevant | Comments: |
| **5.6 Are the conclusions adequate?** | ☐ Adequate | Comments: |

| | | |
|---|---|---|
| *For example:*<br>• How clear are the links between data, interpretation and conclusions?<br>• Are the conclusions plausible and coherent?<br>• Have alternative explanations been explored and discounted?<br>• Does this study enhance understanding of the research subject?<br>• Are the implications of the research clearly defined?<br>• Is there adequate discussion of any limitations encountered? | ☐ Inadequate<br>☐ Not sure | |

| **Section 6: ethics** | | |
|---|---|---|
| **6.1 How clear and coherent is the reporting of ethical considerations?**<br>*For example,*<br>• Have ethical issues been taken into consideration?<br>• Are ethical issues discussed adequately – do they address consent and anonymity?<br>• Have the consequences of the research been considered; for example, raising expectations, changing behaviour?<br>• Was the study approved by an ethics committee? | ☐ Clear<br>☐ Not clear<br>☐ Not sure/not reported | Comments: |

## *Notes on use of Methodology checklist: qualitative studies*

There is considerable debate over which quality criteria should be used to assess qualitative studies. Quality in qualitative research can be assessed using the same broad concepts of validity (or trustworthiness) used for quantitative research, but these need to be put in a different contextual framework to take into account the aims of qualitative research.

This checklist is designed for people with a basic understanding of qualitative research methodology, and is based on the broadly accepted principles that characterise qualitative research and that may affect its validity. The following notes provide suggestions for completing the checklist. A list of publications on qualitative research is provided at the end of these notes for further reading on this topic.

The studies covered by this checklist are those that collect and analyse qualitative data – usually (but not exclusively) textual (written), spoken or observational data. Qualitative data are occasionally collected using structured questionnaires (for example, as thematically organised free-text comments), but such research needs to be scrutinised carefully, as it may not meet acceptable quality criteria for consideration as a qualitative study.

The questions in the checklist are framed to encompass the variety of ways in which qualitative research is conducted. Care must be taken to apply the checklist in a way that matches the research methodology.

*Note that the sub-questions given as examples under each question in the checklist are intended to highlight some of the key issues to be considered for that question – they are not intended to be exhaustive. Please add any additional considerations in the comments box.*

### Section 1: theoretical approach

This section deals with the underlying theory and principles applied to the research.

#### *1.1 Is a qualitative approach appropriate?*

A qualitative approach can be judged to be appropriate when the research sets out to investigate phenomena that are not easy to quantify or measure accurately, or where such measurement would be arbitrary and inexact. If clear numerical measures could reasonably have been put in place, then consider whether a quantitative approach may have been more appropriate.

Qualitative research in public health commonly measures:

- personal experiences (for example, of a condition, treatment or situation)
- processes (for example, action research, practitioner or patient views on the acceptability of using new technology)
- personal values and beliefs (for example, about death, birth, disability)
- interactions and relationships (for example, the quality of the GP–patient relationship, the openness of a psychotherapeutic relationship)

- service evaluations (for example, what was good or bad about patients' experiences of a smoking cessation group).

### *1.2 Is the study clear in what it seeks to do?*

The design of qualitative research tends to be 'theory generative' rather than 'theory testing'; it is therefore unlikely that a research question will be found in the form of a hypothesis or null hypothesis in the way that you would expect in traditional quantitative research. Nevertheless, the paper should still set out early and clearly what the study is investigating and what the parameters are. The research question should be set in context by the provision of an adequate summary of the background literature and the study's underpinning values and assumptions.

## Section 2: study design

This section considers the robustness of the design of the research project.

### *2.1 How defensible/rigorous is the research design/methodology?*

There are a large number of qualitative methodologies, and a tendency in healthcare studies to 'mix' aspects of different methodologies or to use a generic qualitative method. From a qualitative perspective, none of this compromises the quality of the study as long as the following criteria are fulfilled:

- The research design should capture appropriate data and have an appropriate plan of analysis for the subject under investigation. There should be a clear and reasonable justification for the methods chosen.
- The choice of sample and sampling method should be clearly set out (ideally including any shortcomings of the sample) and should be reasonable. It is important to remember that sampling in qualitative research can be purposive and should not be random. Qualitative research is not experimental and does not purport to be generalisable, and therefore does not require a large or random sample. People are usually 'chosen' for qualitative research based on being key informers.

## Section 3: data collection

### *3.1 How well was the data collection carried out?*

Were the methods of data collection used the most appropriate, given the aims of the research? Was the data collection robust, and are there details of:

- how the data were collected?
- how the data were recorded and transcribed? (if verbal data)
- how the data were stored?
- what records were kept of the data collection?

## Section 4: validity

Assessing the validity of qualitative research is very different from assessing that of quantitative research. Qualitative research is much more focused on demonstrating the causes of bias rather than eliminating them. It is therefore good practice to include sections in the report about the reflexive position of

the researcher (their 'role' in the research), the context in which the research was conducted and the reliability of the actual data.

### *4.1 Is the role of the researcher clearly described?*

The researcher should have considered their role in the research; for example, as a reader, interviewer or observer. This is often referred to as 'reflexivity'. The 'status' of the researcher can profoundly affect the data. For example, a middle-aged woman and an 18-year-old man are likely to get different responses to questions about sexual activity when interviewing a group of teenage boys. It is important to consider age, sex, ethnicity and 'insider' status (such as where the interviewer or researcher is part of the group being researched or has the same condition or illness). The researcher can also profoundly influence the data by use of questions, opinions, judgements and so on, so it is important to know what the researcher's position is in this regard, and how the researcher introduced and talked about the research with the participants.

### *4.2 Is the context clearly described?*

It is important when gauging the validity of qualitative data to engage with the data in a meaningful way, and to consider whether the data are plausible and realistic. To make an accurate assessment of this, it is important to have a good feeling for the context of the research in terms of the physical context (for example, youth club, GP surgery, gang headquarters) and who else was there (for example, participants are likely to position themselves very differently, and thus to respond very differently, in a discussion with parents present compared with a discussion with peers present). You should also feel that the participants are described in enough detail that the reader can have some sort of insight into their life and situation. Any potential context bias should be considered.

### *4.3 Were the methods reliable?*

It is important that the method used to collect the data is appropriate for the research question, and that the data generated map well to the aims of the study. Ideally, more than one method should have been used to collect data, or there should be some other kind of system of comparison that allows the data to be compared. This is referred to as 'triangulation'.

## Section 5: analysis

Qualitative data analysis is very different from quantitative analysis. This does not mean that it should not be systematic and rigorous; however, systematisation and rigour require different methods of assessment.

### *5.1 Is the data analysis sufficiently rigorous?*

The main way to assess this is by how clearly the analysis is reported and whether the analysis is approached systematically. There should be a clear and consistent method for coding and analysing data, and it should be clear how the coding and analytical strategies were derived. Above all, these must be reasonable in light of the evidence and the aims of the study. Transparency is the key to addressing the rigour of the analysis.

### *5.2 Are the data 'rich'?*

Qualitative researchers use the adjective 'rich' to describe data that are in-depth, convincing, compelling and detailed enough that the reader feels that they have achieved some level of insight into the research participants' experience. It is also important to know the 'context' of the data – where they came from, what prompted them, what they pertains to, and so on.

### *5.3 Is the analysis reliable?*

The analysis of data can be made more reliable by setting checks in place. It is good practice to have sections of data coded by another researcher, or at least to have a second researcher check the coding for consistency. Participants may also be allowed to verify the transcripts of their interview (or other data collection, if appropriate). Negative or discrepant results should always be highlighted and discussed.

### *5.4 Are the findings convincing?*

In qualitative research, the reader should find the results of the research convincing or credible. This means that the findings should be presented clearly and organised logically, they should not contradict themselves without explanation or consideration, and they should be clear and coherent.

Extracts from original data should be included where possible to give a fuller sense of the findings. These data should be appropriately referenced – although you would expect data to be anonymised, they still need to be referenced in relevant ways (for example, if gender differences were important, then you would expect extracts to be marked male/female).

### *5.5 Are the findings relevant to the aims of the study?*
### *5.6 Are the conclusions adequate?*

These sections are self explanatory.

### Section 6: ethics

### *6.1 How clear and coherent is the reporting of ethical considerations?*

All qualitative research involves ethical considerations, and these should be considered within any research report. Ideally there should be a full discussion of ethics, although this is rare because of space constraints in peer-reviewed journals. Important ethical issues that are raised by a particularly sensitive piece of research should be discussed in enough detail that the reader is convinced that every care was taken to protect research participants.

Any qualitative research should be approved by a research ethics committee, and this should be stated in the report.

## *Further reading*

Barbour RS (2001) Checklists for improving rigour in qualitative research: a case of the tail wagging the dog? British Medical Journal 322: 1115–7.

Daly J, Willis K, Small R et al. (2007) A hierarchy of evidence for assessing qualitative health research. Journal of Clinical Epidemiology 60: 43–9.

Mays N, Pope C (2000) Assessing quality in qualitative research. British Medical Journal 320: 50–2.

Miller G, Dingwall R, editors (1997) Context and method in qualitative research. London: Sage.

# Appendix J: Methodology checklist: prognostic studies

The criteria used in this checklist are adapted from: Hayden JA, Cote P, Bombardier C (2006) Evaluation of the quality of prognosis studies in systematic reviews. Annals of Internal Medicine 144: 427–37.

| | | | | |
|---|---|---|---|---|
| **Study identification**<br>*Include author, title, reference, year of publication* | | | | |
| **Guideline topic:** | | **Review question no:** | | |
| **Checklist completed by:** | | | | |
| | | ***Circle one option for each question*** | | |
| 1.1 | The study sample represents the population of interest with regard to key characteristics, sufficient to limit potential bias to the results | Yes | No | Unclear |
| 1.2 | Loss to follow-up is unrelated to key characteristics (that is, the study data adequately represent the sample), sufficient to limit potential bias | Yes | No | Unclear |
| 1.3 | The prognostic factor of interest is adequately measured in study participants, sufficient to limit potential bias | Yes | No | Unclear |
| 1.4 | The outcome of interest is adequately measured in study participants, sufficient to limit bias | Yes | No | Unclear |
| 1.5 | Important potential confounders are appropriately accounted for, limiting potential bias with respect to the prognostic factor of interest | Yes | No | Unclear |
| 1.6 | The statistical analysis is appropriate for the design of the study, limiting potential for the presentation of invalid results | Yes | No | Unclear |

## *Notes on use of Methodology checklist: prognostic studies*

The studies covered by this checklist are designed to answer questions about prognosis. Such questions address the likelihood of an outcome for patients from a population at risk for that outcome, based on the presence of a proposed prognostic factor. Prognostic factors may be disease-specific (for example, presence or absence of particular disease feature), demographic (for example, age, sex), or relate to the likely response to treatment or the presence of comorbidities.

This checklist is based on a checklist for the quality appraisal of studies about prognosis developed by Hayden and co-workers (2006).

Checklist items are worded so that a 'yes' response always indicates that the study has been designed and conducted in such a way as to minimise the risk of bias for that item. An 'unclear' response to a question may arise when the answer to an item is not reported or is not reported clearly.

### 1.1 The study sample represents the population of interest with regard to key characteristics, sufficient to limit potential bias to the results

Measures of prognosis can vary substantially when obtained from populations with different clinical or demographic features. Estimates of prognosis are not useful without information about the population from which they were obtained.

To minimise bias, the study population should be clearly defined and described and should represent the source population of interest. Points to consider include the following:

- Are the source population or the population of interest adequately described with respect to key characteristics?
- Are the sampling frame and recruitment adequately described, possibly including methods to identify the sample (number and type used; for example, referral patterns in healthcare), period of recruitment and place of recruitment (setting and geographical location)?
- Are inclusion and exclusion criteria adequately described (for example, including explicit diagnostic criteria or a description of participants at the start of the follow-up period)?
- Is participation in the study by eligible individuals adequate?
- Is the baseline study sample (that is, individuals entering the study) adequately described with respect to key characteristics?

### 1.2 Loss to follow-up is unrelated to key characteristics (that is, the study data adequately represent the sample), sufficient to limit potential bias

Attrition refers to the loss of participants during the course of a study. Consideration should be given to why participants dropped out, as well as

how many dropped out. Attrition bias occurs when there are systematic differences between participants lost to the study and those who remain.

To minimise bias, completeness of follow-up should be described and adequate. Points to consider include the following:

- Is the response rate (that is, proportion of study sample completing the study and providing outcome data) adequate?
- Are attempts to collect information on participants who dropped out of the study described?
- Are reasons for loss to follow-up provided?
- Are the key characteristics of participants lost to follow-up adequately described?
- Are there any important differences in key characteristics and outcomes between participants who completed the study and those who did not?

If your review addresses more than one outcome, you should score this item for each outcome individually.

### 1.3 The prognostic factor of interest is adequately measured in study participants, sufficient to limit potential bias

The prognostic factor under study should be well defined. It should be clear how the investigators determined whether participants were exposed or not to the factor. The same definition and measurement should be used for all participants in the study. Often there may be more than one way of determining the presence or absence of the factor (for example, physical or laboratory tests, questionnaire, reporting of symptoms). The method of measurement should be valid (that is, it measures what it is claimed to measure) and reliable (that is, it measures something consistently).

To minimise bias, prognostic factors should have been defined and measured appropriately. Points to consider include the following:

- Is a clear definition or description of the prognostic factor(s) measured provided (including dose, level, duration of exposure, and clear specification of the method of measurement)?
- Are continuous variables reported, or appropriate cut-off points (that is, not data-dependent) used?
- Are the prognostic factor measured and the method of measurement valid and reliable enough to limit misclassification bias? (This may include relevant outside sources of information on measurement properties, as well as characteristics such as blind measurement and limited reliance on recall.)
- Are complete data for prognostic factors available for an adequate proportion of the study sample?
- Are the method and setting of measurement the same for all study participants?
- Are appropriate methods employed if imputation is used for missing data on prognostic factors?

### 1.4 The outcome of interest is adequately measured in study participants, sufficient to limit potential bias

The outcome under study should be well defined. It should be clear how the investigators determined whether participants experienced, or did not experience, the outcome. The same methods for defining and measuring outcome should be used for all participants in the study. Often there may be more than one way of measuring an outcome (for example, physical or laboratory tests, questionnaire, reporting of symptoms). The method of measurement used should be valid and reliable.

To minimise bias, the outcome(s) of interest should be defined and measured appropriately. Points to consider include the following:

- Is a clear definition of the outcome of interest provided, including duration of follow-up?
- Are the outcome that was measured and the method of measurement valid and reliable enough to limit misclassification bias? (This may include relevant outside sources of information on measurement properties, as well as characteristics such as 'blind' measurement and limited reliance on recall.)
- Are the method and setting of measurement the same for all study participants?

If your review addresses more than one outcome, you should score this item for each outcome individually.

### 1.5 Important potential confounders are appropriately accounted for, limiting potential bias with respect to the prognostic factor of interest

Confounding can occur when there are differences between participants, apart from the presence or absence of the prognostic factor, that are related to both the outcome and the prognostic factor. An example of this is if the participants are recruited at different stages of disease progression. The design and analysis of prognostic studies are usually based on some conceptual model about how factors interact to lead to the outcome.

This question is not relevant where the study is being reviewed for the purposes of identifying the absolute risk of the outcome in the group with the prognostic factor.

To minimise bias, important confounders should be defined and measured, and confounding should be accounted for in the design or analysis. Points to consider include the following:

- Are all important confounders, including treatments (key variables in the conceptual model), measured? Are clear definitions of the important confounders measured (including dose, level and duration of exposures) provided?
- Is measurement of all important confounders valid and reliable? (This may include relevant outside sources of information on measurement

properties, as well as characteristics such as ‘blind’ measurement and limited reliance on recall.)
- Are the method and setting of measurement of confounders the same for all study participants?
- Are appropriate methods employed if imputation is used for missing data on confounders?
- Are important potential confounders accounted for in the study design (for example, matching for key variables, stratification or initial assembly of comparable groups)?
- Are important potential confounders accounted for in the analysis (that is, appropriate adjustment)?

If your review addresses more than one outcome, you should score this item for each outcome individually.

### 1.6. The statistical analysis is appropriate for the design of the study, limiting potential for the presentation of invalid results

Analysis undertaken within the study that is incorrect or inappropriate for the study design may result in false conclusions being drawn from the data.

To minimise bias, the statistical analysis undertaken should be clearly described and appropriate for the design of the study. Points to consider include the following:

- Is the presentation of data sufficient to assess the adequacy of the analysis?
- Where several prognostic factors are investigated, is the strategy for model building (that is, the inclusion of variables) appropriate and based on a conceptual framework or model?
- Is the selected model adequate for the design of the study?
- Is there any selective reporting of results?
- Are only pre-specified hypotheses investigated in the analyses?

In some circumstances it may be possible to reanalyse the data using the information supplied in the study report, in order to remove the bias.

# Appendix K: Evidence tables

## *K1: Evidence table for intervention studies*

This table is also suitable for diagnostic studies that compare the effectiveness of two or more tests. This only applies if the test is included as part of a test-and-treat strategy – otherwise the evidence table for studies of diagnostic test accuracy (K2) should be used.

| Bibliographic reference [1] | Study type [2] | Study quality [3] | Number of patients [4] | Patient characteristics [5] | Intervention [6] | Comparison [7] | Length of follow-up [8] | Outcome measures and effect size [9] | Source of funding [10] | Additional comments [11] |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | | | |

[1] Bibliographic reference: author(s), year, title, journal, volume, pages.
[2] Study type: for example, observational, cohort or case studies.
[3] Study quality: note particular strengths and weaknesses.
[4] Number of patients: total number of patients included in the study, including number of patients in each arm, with inclusion and exclusion criteria. Also record the numbers of patients who started and completed the study.
[5] Patient characteristics: characteristics relevant to the area of interest: age, sex, ethnic origin, comorbidity, disease status, community- or hospital-based.
[6] Intervention: treatment, procedure or test studied. If important for the study, specify duration of treatment. For diagnostic studies the intervention is the diagnostic test studied.
[7] Comparison: placebo or alternative treatment. For diagnostic studies, comparison of the test is with another test.
[8] Length of follow-up: the length of time that patients take part in the study, from first staging treatment until either a pre-specified end-point (for example, death, specified length of disease-free remission) or the end of the data-gathering phase is reached. If the study is halted earlier than originally planned for any reason, this should be noted here.
[9] Outcome measures: list all outcome measures, including associated harms. For studies with a diagnostic component there will be two interventions to consider – the diagnostic test used and the associated treatment. Use a separate line for each outcome.
Effect size: for example, absolute risk reduction and relative risk (reduction), number needed to treat, number needed to harm, odds ratios, as required. Give p-values and confidence intervals whenever possible.

[10] Source of funding: government funding (for example, NHS), voluntary charity (for example, Wellcome Trust), pharmaceutical company; and the role of funding organisations.

[11] Additional comments: additional characteristics and/or interpretations of the studies that the reviewer wishes to record. These might include important flaws in the study not identifiable from other data in the table, and additional questions or issues that will need to be considered but do not figure in the results tables in the study.

## *K2: Evidence table for studies of diagnostic test accuracy*

| Bibliographic reference [1] | Study type [2] | Study quality [3] | Number of patients [4] | Prevalence [5] | Patient characteristics [6] | Type of test [7] | Reference standard [8] | Sensitivity and specificity [9] | Positive and negative predictive values [10] | Source of funding [11] | Additional comments [12] |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | | | | |

[1] Bibliographic reference: author(s), year, article title, journal, volume, pages.
[2] Study type: for example, observational, cohort or case studies.
[3] Study quality: note particular strengths and weaknesses.
[4] Number of patients: total number of patients included in the study, with inclusion and exclusion criteria.
[5] Prevalence: proportion of people with the disease in the population at risk.
[6] Patient characteristics: characteristics relevant to the area of interest: age, sex, ethnic origin, comorbidity, disease status, community- or hospital-based.
[7] Type of test: description of the diagnostic test used in the study. Specify the test threshold where applicable.
[8] Reference standard: used as a measure of outcome. Specify if it is a 'gold standard' or 'current best practice'.
[9] Sensitivity: proportion of individuals classified as positive by the gold (or reference) standard who are correctly identified by the study test.
Specificity: proportion of individuals classified as negative by the gold (or reference) standard who are correctly identified by the study test.
[10] Positive predictive value: proportion of individuals with a positive test result who actually have the disease.
Negative predictive value: proportion of individuals with a negative test result who do not have the disease.
[11] Source of funding: government funding (for example, NHS), voluntary/charity (for example, Wellcome Trust), pharmaceutical company; and the role of funding organisations.
[12] Additional comments: additional characteristics and/or interpretations of the studies that the reviewer wishes to record. These might include important flaws in the study not identifiable from other data in the table, and additional questions or issues that will need to be considered but do not figure in the results tables in the study (for example, if a test is one of a sequence of tests; if its utility was determined).

## *K3: Evidence table for prognostic studies*

| Bibliographic reference<br><br>[1] | Study type<br><br>[2] | Study quality<br><br>[3] | Number of patients<br><br>[4] | Patient characteristics<br><br>[5] | Prognostic factor(s)<br><br>[6] | Length of follow-up<br><br>[7] | Outcome measures<br><br>[8] | Results<br><br>[9] | Source of funding<br><br>[10] | Additional comments<br><br>[11] |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | | | |

[1] Bibliographic reference: author(s), year, article title, journal, volume, pages.
[2] Study type: for example, cohort, nested cohort, case series.
[3] Study quality: note particular strengths and weaknesses.
[4] Number of patients: total number of patients included in the study, including number of patients in each arm, with inclusion and exclusion criteria. Also record numbers of patients who started and completed the study.
[5] Patient characteristics: characteristics relevant to the area of interest: age, sex, ethnic origin, comorbidity, disease status, community- or hospital-based. Include method used to select participants.
[6] Prognostic factor(s): include details of method of measurement.
[7] Length of follow-up: the length of time that patients take part in the study, from entry until either a pre-specified end-point (for example, death, specified length of disease-free remission) or the end of the data-gathering phase is reached. If the study is halted earlier than originally planned for any reason, this should be noted here.
[8] Outcome measures: all outcome measures should be listed, with each on a separate line.
[9] Results: relative risk or hazard associated with the prognostic factor of interest; absolute risk of event in baseline group.
[10] Source of funding: government funding (for example, NHS), voluntary charity (for example, Wellcome Trust), pharmaceutical company; and the role of funding organisations.
[11] Additional comments: additional characteristics and/or interpretations of the studies that the reviewer wishes to record. These might include important flaws in the study not identifiable from other data in the table, and additional questions or issues that will need to be considered but do not figure in the results tables in the study.

## *K4: Evidence table for qualitative studies*

### Title: (review question)

| Reference | Research parameters | | | | Population | Outcomes | Funding | Additional comments | |
|---|---|---|---|---|---|---|---|---|---|
| Bibliographic reference<br><br>[1] | Research question<br><br>[2] | Theoretical approach<br><br>[3] | Data collection<br><br>[4] | Method and process of analysis<br><br>[5] | Population and sample collection<br><br>[6] | Key themes<br><br>[7] | Source of funding<br><br>[8] | Limitations<br><br>[9] | Evidence gap<br><br>[10] |
| | | | | | | | | | |

[1] Bibliographic reference: author(s), year, article title, journal, volume, pages.
[2] Research question: what was/were the research question(s)?
[3] What theoretical approach (for example, grounded theory, interpretive phenomenological analysis) does the study take (if specified)?
[4] How were the data collected? Give details of:
- method(s)
- by whom
- setting(s)
- when.

[5] Method and process of analysis: what methods were used to analyse the data (for example, constant comparative method)?
[6] Population and sample collection: what population was the sample recruited from? Include the following information:
- how they were recruited (for example, specify the type of purposive sampling)
- how many participants were recruited
- specific exclusion criteria
- specific inclusion criteria.

[7] Key themes: list all relevant to this review (with illustrative quotes if available).
[8] Source of funding: government funding (for example, NHS), voluntary charity (for example, Wellcome Trust), pharmaceutical company; and the role of funding organisations.
[9] Limitations: both those identified by the author(s) and those identified by the reviewer.
[10] Evidence gap and/or recommendations for future research.

# Appendix L: Modified GRADE profile

## *L1 Example of evidence profile*

**Review: Omega-3 acid ethyl ester supplements vs control in people within 3 months of an acute myocardial infarction**

| Quality assessment | | | | | | | Summary of findings | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **Clinical evidence** | | | | | | | | | | | |
| *No. of studies* | *Design* | *Limitations* | *Inconsistency* | *Directness* | *Imprecision* | *Other considerations* | *Intervention* | *Control* | *Relative risks* | *Risk difference* | *Quality* |
| All-cause mortality | | | | | | | | | | | |
| 3 | RCT | Serious [a] | No important inconsistency | Some uncertainty [b] | No serious imprecision | None | 581/6830 | 755/6830 | 0.83 (0.75 to 0.93) | −0.02 (−0.03 to −0.01) | Low |
| Combined cardiovascular events | | | | | | | | | | | |
| 3 | RCT | Serious [a] | No important inconsistency | Some uncertainty [b] | No serious imprecision | None | 755/6830 | 839/6826 | 0.90 (0.82 to 0.99) | −0.01 (−0.02 to 0.00) | Low |
| Cancers | | | | | | | | | | | |
| 3 | RCT | Serious [a] | No important inconsistency | Some uncertainty [b] | No serious imprecision | None | 150/6830 | 138/6826 | 1.09 (0.86 to 1.36) | 0.00 (0.00 to 0.01) | Low |

| **Economic evidence** | | | | | | | |
|---|---|---|---|---|---|---|---|
| *Study* | *Limitations* | *Applicability* | *Other comments* | *Incremental cost (2006 £)* | *Incremental effects* | *ICER* | *Uncertainty* |
| Franzosi 2001 | Potentially serious limitations [c] | Partially applicable [d] | Based only on measured resource use and survival in 3.5 years follow-up in GISSI-P. | £871 [e] | 0.0332 LYs | £26,243 per LY gained | £16,769 to £56,025 per LY gained (best/worst case) |
| Lamotte 2006 | Very serious limitations [f] | Partially applicable [g] | Based on measured resource use and survival over 3.5 years in GISSI-P, plus longer-term survival benefits attributed to non-fatal events using Canadian database. Belgian results presented. | £1090 [h] | 0.282 LYs | £3860 per LY gained | >98% probability ICER less than €20,000 per QALY gained |
| NCC analysis | Minor limitations [i] | Directly applicable [j] | Based on morbidity and mortality estimated from Markov model using pooled effectiveness data from GISSI-P and DART. Results were sensitive to the size of treatment effects and over their assumed duration. | £1073 | 0.09 QALYs | £12,480 per QALY gained | £3912 to £130,705 per QALY gained (range in one-way sensitivity analyses) |

[a] Increase in statin use over follow-up in GISSI-P differed between the groups (from 4.4% to 46.0% in the omega-3 group and from 5.1% to 44.4% in the control group).
[b] High baseline rate of fish consumption in GISSI (more than 70%).
[c] This study is relatively conservative, as it does not impute any quality–of-life or longer-term survival benefit to supplements. Conversely, it omits gastrointestinal side effects.
[d] Some uncertainty over the applicability of Italian trial data to the UK. May be differences in population risk and diet as well as healthcare use and unit costs.
[e] Converted from 1999 Italian Euros using a PPP exchange rate of 0.797 (www.oecd.org/std/ppp) then uprated by inflation factor of 133.8% (www.pssru.ac.uk/pdf/uc/uc2006/uc2006.pdf).
[f] Methods and data used to estimate life expectancy are questionable, and were not subjected to sensitivity analysis. This is likely to have biased the results.

[g] Some uncertainty over the applicability of Italian trial data to the UK. May be differences in population risk and diet as well as healthcare use. Unit costs may also differ for UK.
[h] Converted from 2004 Belgian Euros using a PPP conversion rate of 0.706 (www.oecd.org/std/ppp) then uprated by inflation factor of 107.3% (www.pssru.ac.uk/pdf/uc/uc2006/uc2006.pdf).
[i] Some limitations in reporting (for example, for inputs taken from NICE statins appraisal). However, analysis is based on best-available effectiveness estimates and follows NICE methodological guidance. The robustness of results is also well tested through sensitivity analysis and comparison with other study results.
[j] Some uncertainty over applicability of trial data to UK because of differences in population risk and diet. However, resource use and unit costs are UK-specific and the perspective and discount rates follow the NICE reference case.

## *L2 Blank evidence profile*

| Quality assessment | | | | | | | Summary of findings | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **Clinical evidence** | | | | | | | | | | | |
| *No. of studies* | *Design* | *Limitations* | *Inconsistency* | *Directness* | *Imprecision* | *Other considerations* | *Intervention* | *Control* | *Relative effect* | *Absolute effect* | *Quality* |
| Outcome | | | | | | | | | | | |
| | | | | | | | | | | | |
| Outcome | | | | | | | | | | | |
| | | | | | | | | | | | |
| Outcome | | | | | | | | | | | |
| | | | | | | | | | | | |
| **Economic evidence** | | | | | | | | | | | |
| *Study* | *Limitations* | *Applicability* | *Other comments* | | | | *Incremental cost (£)* | *Incremental effects* | *ICER* | *Uncertainty* | |
| | | | | | | | | | | | |
| | | | | | | | | | | | |
| | | | | | | | | | | | |

# Appendix M: Abbreviations and glossary

## *M1 Abbreviations*

| | |
|---|---|
| **CPHE** | Centre for Public Health Excellence |
| **DH** | Department of Health |
| **GDG** | Guideline Development Group |
| **GRADE** | Grading of recommendations assessment, development and evaluation |
| **GRP** | Guideline Review Panel |
| **HTA** | Health technology assessment |
| **ICER** | Incremental cost-effectiveness ratio |
| **IP** | Interventional procedure |
| **MeSH** | Medical subject headings |
| **MTA** | Multiple technology appraisal |
| **NCC** | National Collaborating Centre |
| **NIHR** | National Institute for Health Research |
| **PDG** | Programme Development Group |
| **PHIAC** | Public Health Interventions Advisory Committee |
| **PICO** | Patient, intervention, comparison and outcome |
| **QALY** | Quality-adjusted life year |
| **QUADAS** | Quality assessment of studies of diagnostic accuracy included in systematic reviews |
| **PPIP** | Patient and Public Involvement Programme |
| **RCT** | Randomised controlled trial |
| **STA** | Single technology appraisal |

## *M2 Glossary*

| | |
|---|---|
| **Absolute risk reduction (risk difference)** | The difference in risk between two groups (one subtracted from the other) in a comparative study. |
| **Abstract** | Summary of a study, which may be published alone or as an introduction to a full scientific paper. |
| **Adverse event** | An undesirable effect of an intervention. |
| **AGREE (appraisal of guidelines research and evaluation)** | An international collaboration of researchers and policy makers whose aim is to improve the quality and effectiveness of clinical practice guidelines (www.agreecollaboration.org). The AGREE instrument, developed by the group, is designed to assess the quality of clinical guidelines. |
| **Algorithm (in guidelines)** | A flow chart of the clinical decision pathway described in the guideline, where decision points are represented by boxes, linked with arrows. |
| **Allocation concealment** | The process used to prevent advance knowledge of group assignment in a randomised controlled trial. The allocation process should be uninfluenced by the person making the allocation, by being administrated by someone who is not responsible for recruiting participants. |
| **Applicability** | The degree to which the results of an observation, study or review are likely to hold true in a particular clinical practice setting. |
| **Appraisal Committee** | A standing advisory committee of NICE. Its members are drawn from the NHS, patient and carer organisations, relevant academic disciplines and the healthcare industries. |
| **Arm (of a clinical study)** | Subsection of participants within a study who receive one particular intervention (for example, the placebo arm). |
| **Assessment Group** | An independent group of researchers commissioned by NICE, as part of the technology appraisal process, to review the evidence on a group of treatments. |
| **Assessment report** | A review of the evidence about how well a group of similar treatments work, and whether they offer value for money. Assessment reports are produced for treatments being assessed using the multiple technology appraisal process. |
| **Association** | Statistical relationship between two or more events, characteristics or other variables. The relationship may or may not be causal. |

| | |
|---|---|
| **Audit support** | The provision of ready-to-use criteria, including exceptions, definitions and data source suggestions, in order to make the process of developing clinical audit projects easier. NICE provides audit support for all clinical guidelines. |
| **Audit trail** | Records of action to assess practice against standards. Also a record of actions (for example, changes to a draft guideline) so that the reasons are apparent to a third party. |
| **Baseline** | The initial set of measurements at the beginning of a study (after the period before the study starts when no treatment is given [the 'run-in' period], where applicable), with which subsequent results are compared. |
| **Bespoke implementation tools** | Tools produced in addition to the implementation support tools that are produced routinely. Bespoke implementation tools are tailored to needs that are identified in the implementation planning meeting or in other discussions with stakeholders. Examples include: implementation advice, templates for referral letters, flow charts, fact sheets and checklists. |
| **Bias** | Systematic (as opposed to random) deviation of the results of a study from the 'true' results, which is caused by the way the study is designed or conducted. |
| **Blinding (masking)** | The practice of keeping the investigators or participants of a study unaware of the group to which a participant has been assigned. |
| **'Burden of disease' study** | A study investigating the overall impact of diseases and injuries at the individual level, at the societal level or on the economic costs of diseases. |
| **Carer** | Someone other than a healthcare professional who is involved in caring for a person with a medical condition. |
| **Case–control study** | Comparative observational study in which the investigator selects people who have experienced an event (for example, developed a disease) and others who have not (controls), and then collects data to determine previous exposure to a possible cause. |
| **Case series** | Report of a number of cases of a given disease, usually covering the course of the disease and the response to treatment. There is no comparison (control) group of patients. |
| **Centre for Clinical Practice** | The department at NICE that manages the development of clinical guidelines. It commissions one of the National Collaborating Centres to develop each clinical guideline. |

| | |
|---|---|
| **Centre for Health Technology Evaluation** | The department at NICE that is responsible for producing technology appraisals and interventional procedures guidance. The guidance is developed by independent committees – the Appraisal Committee and the Interventional Procedures Advisory Committee. |
| **Centre for Public Health Excellence** | The department at NICE that is responsible for producing public health guidance. |
| **Class (of drugs)** | A group of drugs with the same or similar mechanism of action; these drugs may or may not have the same basic chemical structure. However, there may be differences between drugs within a class (for example, in side-effect profile). |
| **Clinical Adviser** | A member of the Guideline Development Group who works closely with the National Collaborating Centre technical team to provide expert topic-specific support. Responsibilities of the clinical adviser may include working with the systematic reviewer on the detail of the evidence reviews or checking clinical and technical terminology in the full guideline. |
| **Clinical audit** | A quality-improvement process that seeks to improve patient care and outcomes through systematic review of care against explicit criteria and the implementation of change. |
| **Clinical effectiveness** | The extent to which an intervention produces an overall health benefit in routine clinical practice. |
| **Clinician** | A healthcare professional providing healthcare (for example, a doctor, nurse or physiotherapist). |
| **Cochrane Library** | A regularly updated electronic collection of evidence-based medicine databases, including the Cochrane Database of Systematic Reviews. |
| **Cochrane Review** | A systematic review of the evidence from randomised controlled trials relating to a particular health problem or healthcare intervention, produced by the Cochrane Collaboration. Available electronically as part of the Cochrane Library. |
| **Code of conduct (of the GDG)** | A code of conduct developed by NICE for Guideline Development Group (GDG) members and other people who attend GDG meetings. This code sets out the responsibilities of NICE and the GDG, and the principles of transparency and confidentiality. |

| | |
|---|---|
| **Cohort study** | A retrospective or prospective follow-up study. People to be followed up are grouped on the basis of whether or not they have been exposed to a suspected risk factor or intervention. A cohort study can be comparative, in which case two or more groups are selected on the basis of differences in their exposure to the intervention of interest. |
| **Commentator** | Organisations that engage in the technology appraisal process but that are not asked to prepare a submission dossier, and that receive the final appraisal determination for information only, without right of appeal. |
| **Commercial in confidence** | See 'In confidence material' |
| **Commissioning guide** | An implementation tool available on NICE's website to help senior healthcare professionals and health service managers with decisions when they are commissioning services in clinical areas for which NICE has issued guidance. |
| **Comorbidity** | Co-existence of a disease or condition, or more than one disease or condition, in a person in addition to the disease or condition being studied or treated. |
| **Comparability** | Similarity of groups in terms of characteristics likely to affect study results (such as health status or age). |
| **Comparator** | The standard intervention against which an intervention is compared in a study. The comparator can be no intervention (for example, best supportive care). |
| **Complementary therapy** | Practices not generally recognised by the medical community as standard or conventional medical approaches, which are used to enhance or complement standard treatments. |
| **Conceptual framework** | A theoretical structure of assumptions, principles, and rules that holds together the ideas comprising a broad concept. |
| **Conceptual model** | A descriptive model of a system based on qualitative assumptions about its elements, their interrelationships, and system boundaries. |
| **Confidence interval (CI)** | A measure of uncertainty around the result of a statistical analysis. The 'confidence' interval means that if the method used to calculate the interval is repeated many times on different samples, then that proportion of intervals will actually contain the true value in the population. |
| **Conflict of interest** | An interest that might conflict, or be perceived to conflict, with duties and responsibilities to an organisation. |

| | |
|---|---|
| **Confounding** | In a study, confounding occurs when the effect of an intervention on an outcome is distorted as a result of an association between the population or intervention or outcome and another factor (the 'confounding variable' or 'confounder') that can influence the outcome independently of the intervention under investigation. |
| **Consensus methods** | Techniques that aim to reach an agreement on a particular issue. Formal consensus methods include Delphi and nominal group techniques, and consensus development conferences. In the development of clinical guidelines, consensus methods may be used where there is a lack of strong research evidence on a particular topic. Expert consensus methods aim to reach agreement between experts in a particular field. |
| **Consultation** | A stage during clinical guideline development when organisations can comment on the draft guideline. These organisations must be registered with NICE as stakeholders. There is also a consultation on the draft scope. |
| **Consultee** | An organisation that accepts an invitation to participate in a technology appraisal. Consultees can comment on the draft scope, the assessment report or Evidence Review Group report, and the appraisal consultation document during the consultation process. All consultees are given the opportunity to appeal against the final appraisal determination. |
| **Control** | An explicitly defined comparator against which the effects of an intervention are compared in a clinical study. |
| **Cost–consequence analysis** | A form of economic evaluation where the costs and consequences of two or more interventions are compared, and the consequences are reported separately from costs. |
| **Cost-effectiveness analysis** | A form of economic evaluation in which consequences of different interventions are measured using a single outcome, usually in 'natural' units (for example, life-years gained, deaths avoided, heart attacks avoided, cases detected). Alternative interventions are then compared in terms of cost per unit of effectiveness. |
| **Cost-minimisation analysis** | A form of economic evaluation that compares the costs of alternative interventions that have equal effects. |

| | |
|---|---|
| **'Cost of illness' study** | A study that measures the economic burden of a disease or diseases and estimates the maximum amount that could potentially be saved or gained if a disease was eradicated. |
| **Cost–utility analysis** | A form of cost-effectiveness analysis in which the units of effectiveness are quality-adjusted life years (QALYs). |
| **Costing report and costing template** | Documents that help with analysing the overall costs of putting NICE guidance into practice and any savings. The costing report summarises the national costs. The costing template helps primary care trusts to work out how much it will cost to implement the guidance in their area, and how much they could save. |
| **Costing statement** | A short statement that is published if there are not likely to be significant costs involved in implementing a piece of NICE guidance. |
| **Cross-sectional study** | The observation of a defined set of people at a single point in time or time period. This type of study contrasts with a longitudinal study, which follows a set of people over a period of time. |
| **Decision(-analytic) model (and/or technique)** | A model of how decisions are or should be made. This could be one of several models or techniques used to help people to make better decisions (for example, when considering the trade-off between costs, benefits and harms of diagnostic tests or interventions). |
| **Decision tree** | A method for helping people to make better decisions in situations of uncertainty. It illustrates the decision as a succession of possible actions and outcomes. It consists of the probabilities, costs and health consequences associated with each option. The overall effectiveness or cost-effectiveness of different actions can then be compared. |
| **Delphi technique** | A technique used for the purpose of reaching an agreement on a particular issue, without the participants meeting or interacting directly. It involves sending participants a series of postal questionnaires asking them to record their views. After the first questionnaire, participants are asked to give further views in the light of the group feedback. The judgements of the participants are statistically aggregated, sometimes after weighting for expertise. |

| | |
|---|---|
| **Discounting** | Costs and perhaps benefits incurred today have a higher value than costs and benefits occurring in the future. Discounting health benefits reflects individual preference for benefits to be experienced in the present rather than the future. Discounting costs reflects individual preference for costs to be experienced in the future rather than the present. |
| **Discrete event simulation** | A method that can be used to model the course of a disease (for example, to predict disease progression for the purposes of cost-effectiveness analysis). |
| **Dominate (in cost-effectiveness analysis)** | A term used in health economics when a treatment option is both more clinically effective and less costly than an alternative option. This treatment is said to 'dominate' the less effective and more costly option. |
| **Dosage** | The prescribed amount of a drug to be taken, including the size and timing of the doses. |
| **Economic evaluation** | Comparative analysis of alternative health strategies (interventions or programmes) in terms of both their costs and their consequences. |
| **Effect (as in effect measure, treatment effect, estimate of effect, effect size)** | The observed association between interventions and outcomes or a statistic to summarise the strength of the observed association. |
| **Effectiveness** | See ‘Clinical effectiveness’. |
| **Efficacy** | The extent to which an intervention is active when studied under controlled research conditions. |
| **Epidemiology** | The study of a disease within a population, defining its incidence (the number of instances of people falling ill during a given time in a specified population) and prevalence (the proportion of people in a population with a particular characteristic) and examining the roles of external influences (for example, infection or diet) and interventions. |
| **Equity** | Fair distribution of resources or benefits. |
| **Evidence** | Information on which a decision or guidance is based. Evidence is obtained from a range of sources, including randomised controlled trials, observational studies and expert opinion (of healthcare professionals and/or patients). |
| **Evidence profile** | A table summarising, for each important clinical outcome, the quality of the evidence and the outcome data (part of the GRADE approach; see definition below). |
| **Evidence table** | A table summarising the results of a collection of studies which, taken together, represent the evidence supporting a particular recommendation or series of recommendations in a guideline. |

| | |
|---|---|
| **Evidence statement** | A brief summary of one finding from a review of evidence that a clinical guideline is based on. |
| **Exceptional update** | Review of existing guidance carried out sooner than originally planned because new data have become available. |
| **Exclusion criteria (clinical study)** | Criteria that define who is not eligible to participate in a clinical study. |
| **Exclusion criteria (literature review)** | Explicit standards used to decide which studies should be excluded from consideration as potential sources of evidence. |
| **Executive lead** | The NICE executive director who is responsible for a piece of guidance. Executive leads are not usually involved day-to-day in the production of guidance. |
| **Expert adviser** | A person who has specialist knowledge in a particular area related to a clinical guideline. The expert adviser attends Guideline Development Group meetings to give advice, but is not a full member of the group. |
| **Expert consensus** | See 'Consensus methods'. |
| **Extrapolation** | In data analysis, predicting the value of a parameter outside the range of observed values. |
| **Facilitator** | A person whose role is to promote the effective functioning of a group. |
| **Follow up** | Observation over a period of time of a person, group or initially defined population whose characteristics have been assessed in order to observe changes in health status or health-related variables. |
| **Free text terms** | Data entered into a field without any formal or pre-defined structure other than the normal use of grammar and punctuation. |
| **Full guideline** | The version of a clinical guideline that contains the recommendations, summaries of the evidence and an explanation of how the recommendations were developed. It is written by members of the National Collaborating Centre (NCC) responsible for the guideline, and a Guideline Development Group. It is published by the NCC. |
| **Generalisability** | The extent to which the results of a study based on measurement in a particular patient population and/or a specific context hold true for another population and/or in a different context. In this instance, this is the degree to which a guideline recommendation is applicable across both geographical and contextual settings. For instance, guidelines that recommend increases in numbers of staff should acknowledge that associated costs might vary across the country. |

| | |
|---|---|
| **Generic name** | The general non-proprietary name of a drug or device. |
| **GRADE (Grading of recommendations assessment, development and evaluation)** | A systematic and explicit approach to grading the quality of evidence and the strength of recommendations. |
| **Grading (of evidence)** | A code given to a study or other evidence, indicating the quality and generalisability of the research. The highest grade evidence will usually be obtained from randomised controlled trials. |
| **Grey literature** | Reports that are unpublished or have limited distribution, and are not included in the common bibliographic retrieval systems. |
| **Guidance Executive** | NICE executive directors who approve all NICE guidance for publication. |
| **Guideline consultation table** | A table of all the comments received by NICE during guideline consultation. The Guideline Development Group considers the comments received, and the National Collaborating Centre then responds to the comments in the table. |
| **Guideline Development Group (GDG)** | A group of healthcare professionals, patients and carers, and technical staff who develop the recommendations for a clinical guideline. The National Collaborating Centre (NCC) responsible for developing the guideline recruits a GDG to work on the guideline. NCC staff review the evidence and support the GDG. The group writes draft guidance, and then revises it after a consultation with stakeholders. |
| **Guideline Review Panel** | A panel of independent experts who comment on the draft scope for a clinical guideline and check the full guideline. The panel pays particular attention to how the Guideline Development Group has responded to comments received during consultation. The members include healthcare professionals, and representatives of the healthcare industry and patients. |
| **Guidelines Commissioning Manager** | The NICE staff member with responsibility for managing the development of a particular clinical guideline. |
| **Handsearch/handsearching** | The planned searching of a journal page-by-page (by hand) to identify reports of studies to answer review questions. |
| **Harms** | Adverse effects of an intervention. |

| | |
|---|---|
| **Health economist** | A member of the Guideline Development Group with skills in economic analysis whose role is to advise on economic aspects of the clinical issues or questions, review economic literature, prioritise topics for further analysis and carry out additional cost-effectiveness analyses. |
| **Health inequalities** | The gap in health status and in access to health services, between different social classes and ethnic groups, and between populations in different geographical areas. For more information, see Department of Health website: www.dh.gov.uk/healthinequalities |
| **Health-related quality of life** | A combination of a person's physical, mental and social well-being; not merely the absence of disease. |
| **Health technology** | Any method used by those working in health services to promote health, prevent and treat disease, and improve rehabilitation and long-term care. Technologies in this context are not confined to new drugs or pieces of sophisticated equipment. |
| **Health Technology Assessment review** | Independent research information about the effectiveness, costs and broader impact of healthcare treatments and tests for those who plan, provide or receive care in the NHS. The Health Technology Assessment (HTA) programme is part of the National Institute for Health Research (NIHR). |
| **Healthcare professional member** | A member of the Guideline Development Group with appropriate knowledge and skills to represent the perspective(s) of the healthcare professionals (and social care professionals where relevant) involved in the care of patients affected by the guideline topic. |
| **Heterogeneity** | Used in meta-analyses and systematic reviews when the results or estimates of effects of a treatment from separate studies seem to be very different (for example, the size of treatment effects may vary across studies, or some studies may indicate beneficial treatment effects where others suggest adverse treatment effects). Such results may occur because of differences between studies in terms of the patient populations, outcome measures or definition of variables. |
| **Hypothesis** | An unproven theory that can be tested by research. |
| **Implementation** | The process of putting guidance into practice. |
| **Implementation advice** | Advice on how to put into practice a specific piece of NICE guidance. This advice is for NHS staff whose job includes ensuring this happens. The advice includes an action plan that staff can use. |

| | |
|---|---|
| **Implementation tool** | Various 'tools' are produced to help the NHS put NICE guidance into practice. These include costing reports, costing templates, implementation advice and slide sets. |
| **In confidence material** | Information (for example, the findings of a research project) defined as 'confidential' as its public disclosure could have an impact on the commercial interests of a particular company ('commercial in confidence') or the academic interests of a research or professional organisation ('academic in confidence'). |
| **Inclusion criteria (literature review)** | Explicit criteria used to decide which studies should be considered as potential sources of evidence. |
| **Incremental analysis** | The analysis of additional costs and additional clinical outcomes with different interventions. |
| **Incremental cost-effectiveness ratio (ICER)** | The difference in the mean costs in the population of interest divided by the differences in the mean outcomes in the population of interest. |
| **Index** | In epidemiology and related sciences, this word usually means a rating scale (for example, a set of numbers derived from a series of observations of specified variables). Examples include the various health status indices, and scoring systems for severity or stage of cancer. |
| **Index test** | The test being evaluated in a study to compare it with the best available test (the reference standard). |
| **Indication (specific)** | The defined use of a technology as licensed by the Medicines and Healthcare products Regulatory Agency. |
| **Indirect comparison** | An analysis that compares interventions that have not been compared directly within a head-to-head, randomised trial. |
| **Information specialists** | Specialists, based either at NICE or within a National Collaborating Centre, who assess the suitability of topic suggestions for consideration by NICE for the standard and short clinical guideline programmes, and provide information to support the decision-making of the topic selection team and affiliated groups. |
| **Internal validity** | The degree to which the results of a study are likely to approximate the 'truth' for the participants recruited in a study (that is, are the results free from bias?). |

| | |
|---|---|
| **Interventional procedure** | Any surgery, test or treatment that involves entering the body through skin, muscle, a vein or artery, or a body cavity, or using electromagnetic radiation (which includes X-rays, lasers, gamma-rays and ultraviolet light). |
| **Interventional Procedures (IP) Advisory Committee (IPAC)** | The independent committee that advises NICE on whether an interventional procedure is safe enough and works well enough to be used in the NHS. |
| **Key clinical issues** | The most important aspects of care that a clinical guideline will cover in order to ensure that it focuses on areas in which the NHS most needs advice. Key clinical issues relate to the effectiveness and cost effectiveness of interventions or tests that are being considered for a given population. |
| **Key priorities for implementation** | Up to 10 recommendations from a clinical guideline that should be implemented first because they will have the biggest impact. They are chosen by the Guideline Development Group. |
| **Licence** | See 'Marketing authorisation'. |
| **Likelihood ratio** | The ratio of the probability that a person with a condition has a specified test result to the probability that a person without the condition has the same specified test result. |
| **Marketing authorisation** | An authorisation that covers all the main activities associated with the marketing of a medicinal product. Medicines that meet the standards of safety, quality and efficacy set by the Medicines and Healthcare products Regulatory Agency are granted a marketing authorisation (previously a product licence), which is normally necessary before they can be prescribed or sold. |
| **Markov modelling** | A decision-analytic technique that characterises the prognosis of a cohort of patients by assigning them to a fixed number of health states and then models transitions among health states. |
| **Medical devices** | All products, except medicines, used in healthcare for the diagnosis, prevention, monitoring or treatment of illness or disability. |
| **Medicines and Healthcare products Regulatory Agency (MHRA)** | The Executive Agency of the Department of Health responsible for protecting and promoting public health and patient safety by ensuring that medicines, healthcare products and medical equipment meet appropriate standards of safety, quality, performance and effectiveness, and are used safely. |

| | |
|---|---|
| **MeSH (medical subject headings)** | The US National Library of Medicine's controlled vocabulary thesaurus used for indexing articles from biomedical journals for databases such as MEDLINE. |
| **Meta-analysis** | A statistical technique for combining (pooling) the results of a number of studies that address the same question and report on the same outcomes to produce a summary result. The aim is to derive more precise and clear information from a large data pool. It is generally more likely to reliably confirm or refute a hypothesis than the individual trials. |
| **Meta-ethnography** | A process of identifying relevant findings or other statements from the literature and sorting them into a pattern of evidence on the subject being studied. |
| **Mixed treatment comparison** | An analysis that compares two or more interventions using a combination of direct evidence (from trials that directly compare the interventions of interest) and indirect evidence (trials that do not compare the interventions of interest directly). |
| **Model input** | Information required for economic modelling. For clinical guidelines, this may include information about prognosis, adverse effects, quality of life, resource use or costs. |
| **Multiple technology appraisal** | The name given to the NICE process in which appraisals of more than one technology, or a single technology for more than one indication, are conducted. |
| **Narrative summary** | Summary of findings given as a written description. |
| **National Collaborating Centre (NCC)** | A group set up by NICE to develop clinical guidelines for a particular disease area. Each NCC is based at one of the Royal Medical Colleges. Staff at the NCC review the evidence for a guideline and appoint a Guideline Development Group. |
| **Negative predictive value** | The proportion of people with a negative test result who do not have the disease or characteristic. |
| **Net benefit estimate** | An estimate of the amount of money remaining after all payments made are subtracted from all payments received. This is a source of information used in the economic evidence profile for a clinical guideline. |
| **NICE guideline** | The version of a clinical guideline that presents the recommendations from the full guideline in a format that focuses on implementation by healthcare professionals and NHS organisations. |

| | |
|---|---|
| **Nominal-group technique** | A technique used for the purpose of reaching an agreement on a particular issue. It uses a variety of postal and direct contact techniques, with individual judgements being aggregated statistically to derive the group judgement. |
| **Number needed to harm** | The average number of people from a defined population that would need to be treated with a specific intervention for a given period of time to cause one additional adverse outcome. |
| **Number needed to treat** | The number of patients that on average must be treated with a specific intervention for a given period of time to prevent a single extra occurrence of the outcome of interest. |
| **Observational study** | Retrospective or prospective study in which the investigator observes the natural course of events with or without control groups (for example, cohort studies and case–control studies). |
| **Odds ratio** | A measure of treatment effectiveness. The odds of an event happening in the treatment group, expressed as a proportion of the odds of it happening in the control group. The 'odds' is the ratio of non-events to events. |
| **Off-label** | A situation where a drug is used to treat a condition or disease for which it is not specifically licensed. |
| **Opportunity cost** | The opportunity cost of investing in a healthcare intervention is the other healthcare programmes that are displaced by its introduction. This may be best measured by the health benefits that could have been achieved had the money been spent on the next best alternative healthcare intervention. |
| **Outcome** | Measure of the possible results that may stem from exposure to a preventative or therapeutic intervention. Outcome measures may be intermediate or final endpoints. See ‘Intermediate outcome’. |
| **p-values** | The probability that an observed, or more extreme, difference could have occurred by chance, assuming that there is in fact no true difference between the measurements in the groups being compared. If the probability is less than 1 in 20, the p-value is less than 0.05. A result with a p-value of less than 0.05 is conventionally considered to be ‘statistically significant’. |
| **Patient and carer member** | A member of the Guideline Development Group with knowledge of the issues that are important to patients and carers. |

| | |
|---|---|
| **Patient and Public Involvement Programme** | A department at NICE that advises on involving patients and carers in the guidance programmes. It also supports patients and carers who are members of committees or groups that produce guidance, and patient and carer stakeholder organisations that are commenting on draft guidance. |
| **Peer review** | A process where research is scrutinised by experts who have not been involved in the design or execution of the studies. |
| **Personal social services** | Personal services normally provided for people related to their specific needs and circumstances, in contrast to standardised services provided to people as members of categories. People who are typically users of personal social services include elderly people and their carers, children and families, and people with disabilities and their carers. |
| **PICO (population, intervention, comparison and outcome) framework** | A structured approach for developing questions about interventions that divides each question into four components: the patients (the population under study); the interventions (what is being done); the comparators (other main treatment options); and the outcomes (measures of how effective the interventions have been). |
| **Placebo** | An inactive and physically identical medication or procedure used as a comparator in controlled clinical trials. |
| **Positive predictive value** | The proportion of people with a positive test result who actually have the disease or characteristic. |
| **Pre-publication check** | A new step (introduced in 2009) in the clinical guideline development process that enables registered stakeholders to raise concerns about factual errors that may exist in a guideline after consultation and before its publication. |
| **Primary research** | Study generating original data (see also 'Secondary research'). |
| **Prognosis** | A probable course or outcome of a disease. Prognostic factors are patient or disease characteristics that influence the course. Good prognosis is associated with low rate of undesirable outcomes; poor prognosis is associated with a high rate of undesirable outcomes. |

| | |
|---|---|
| **Prognostic factor** | Patient or disease characteristics (for example, age or co-morbidity) that influence the course of the disease under study. In a randomised trial to compare two treatments, chance imbalances in variables (prognostic factors) that influence patient outcome are possible, especially if the size of the study is fairly small. In terms of analysis, these prognostic factors become confounding factors. |
| **Programme Development Group** | A group of health and other professionals and researchers, brought together to write a particular piece of public health guidance that is being developed using the public health programme process. |
| **Project manager** | The National Collaborating Centre staff member who oversees and facilitates the clinical guideline development process, organising Guideline Development Group (GDG) meetings and providing administrative support to the GDG Chair and members. |
| **Proprietary name** | The brand name given by the manufacturer to a drug or device it produces. |
| **Prospective cohort study** | An observational study that takes a group (cohort) of patients and follows their progress over time in order to measure outcomes such as disease or mortality rates and make comparisons according to the treatments or interventions that patients received. Prospective cohorts are assembled in the present and followed into the future. |
| **Public health guidance** | Guidance on ways to help people reduce their risk of illness and lead a healthier life. |
| **Public Health Interventions Advisory Committee (PHIAC)** | An independent standing committee set up by NICE to write some of its public health guidance. |
| **QUADAS (quality assessment of diagnostic accuracy studies)** | A tool for the quality assessment of studies on the accuracy of diagnostic technologies. |
| **Qualitative research** | Research using qualitative data collection techniques and qualitative analysis. |
| **Quality-adjusted life year (QALY)** | An index of survival that is adjusted to account for the patient's quality of life during this time. QALYs have the advantage of incorporating changes in both quantity (longevity/mortality) and quality (morbidity, psychological, functional, social and other factors) of life. Used to measure benefits in cost–utility analysis. |
| **Quality of life** | See 'Health-related quality of life'. |

| | |
|---|---|
| **Quick reference guide** | A short, printed version of a clinical guideline designed for use by healthcare professionals and other staff who will be following the guidance. It contains the recommendations (or a summary of the recommendations) but not the supporting evidence. |
| **Quorum** | The smallest number of group members that must be present to constitute a valid meeting. The quorum of a Guideline Development Group is 50% of appointed members. No business relating to the formulation of guideline recommendations may be conducted unless the quorum is reached. |
| **Randomisation** | Allocation of participants in a research study to two or more alternative groups using a chance procedure, such as computer-generated random numbers. This approach is used in an attempt to ensure there is an even distribution of participants with different characteristics between groups and thus ensure that groups are comparable. |
| **Randomised controlled trial (RCT)** | A comparative study in which participants are randomly allocated to intervention and control groups and followed up to examine differences in outcomes between the groups. |
| **Receiver operating characteristic (ROC) curve** | A plot of test sensitivity versus (1 − specificity), used to summarise the results of studies of diagnostic test accuracy. |
| **Recommendations** | Formal, numbered paragraphs in NICE clinical guidelines that give specific advice on the appropriate treatment and care of people with specific diseases and conditions within the NHS. |
| **Referral (from the Department of Health)** | A remit that identifies the broad areas to be covered by a clinical guideline and is translated into the scope for the guideline. Topics for clinical guidelines are referred to NICE by the Department of Health, based on recommendations from topic selection consideration panels. |
| **Reference case** | When estimating clinical and cost effectiveness in a technology appraisal, the reference case specifies the methods that are considered by NICE to be the most appropriate for the Appraisal Committee's purpose and are also consistent with an NHS objective of maximising health gain from limited resources. |
| **Reference standard (or gold standard)** | An agreed standard (for example, for a test or treatment) against which other interventions can be compared. |

| | |
|---|---|
| **Relative risk (RR)** | The number of times more likely or less likely an event is to happen in one group compared with another (calculated as the risk of the event in group A divided by the risk of the event in group B). |
| **Relative risk reduction** | The proportional reduction in risk between experimental and control participants in a trial. |
| **Reliability** | The degree of agreement exhibited when a measurement is repeated under identical conditions. Reliability refers to the degree to which the results obtained by a measurement procedure can be replicated. |
| **Remit** | The brief given by the Department of Health at the beginning of the clinical guideline development process. This defines core areas of care that the guideline needs to address. |
| **Research recommendation** | Recommendations for future research covering questions relating to an uncertainty or evidence gap that has been identified during the guideline development process. |
| **Resource implication** | The likely impact in terms of finance, workforce or other NHS resources. |
| **Review of the literature** | An article that summarises the evidence contained in a number of different individual studies and draws conclusions about their findings. It may or may not be systematically researched and developed. |
| **Review protocol** | A document that outlines the background, objectives and planned methods for a systematic review. |
| **Review question** | A structured question about treatment and care that is formulated by the Guideline Development Group from a key clinical issue in the scope to guide the systematic review. A review question has four components:<br>• patients (the population under study)<br>• interventions (what is being done)<br>• comparisons (other main treatment options)<br>• outcomes. |
| **Scope** | Document created at the start of producing a piece of guidance outlining what the guidance will and will not cover. Organisations registered as stakeholders, can comment on the draft scope during a consultation period. The final version of the scope – taking into account comments from the consultation – is used as a starting point for developing the guidance. |
| **Scope consultation table** | A table of all the comments received by NICE during consultation on the guideline scope, which is published on the NICE website with the final scope. |

| | |
|---|---|
| **Scope notes (databases)** | Scope notes provide additional information about database indexing terms (for example, when the term was first used for indexing, how the term is applied in the database, and 'used-for' terms and 'see-related' terms). |
| **Scoping group** | A group led by the National Collaborating Centre with input from the Guideline Development Group (GDG) Chair (and the GDG Clinical Adviser if applicable), NICE, and patient and carer groups, whose role is to:<br>• identify the key areas for inclusion in a piece of guidance<br>• revise the key areas after the stakeholder scoping meeting<br>• draft the scope for consultation<br>• respond to stakeholders' comments<br>• finalise the scope after consultation. |
| **Scoping search** | A search of the literature undertaken at the scoping stage to identify previous clinical guidelines, health technology assessment reports, key systematic reviews and economic evaluations relevant to the guideline topic. |
| **Search filter** | A collection of search terms designed to retrieve selections of records (for example, records of research using a specific study design or on a specific topic). |
| **Secondary research** | Research analysing data from existing studies (see also 'Primary research'). |
| **Selection bias (also known as allocation bias)** | A systematic bias in selecting participants for study groups, so that the groups have differences in prognosis and/or therapeutic sensitivities at baseline. Randomisation (with concealed allocation) of patients protects against this bias. |
| **Sensitivity (of a test)** | The proportion of people classified as positive by the gold (or reference) standard, who are correctly identified by the study test. |

| | |
|---|---|
| **Sensitivity analysis** | A means of representing uncertainty in the results of economic evaluations. Uncertainty may arise from missing data, imprecise estimates or methodological controversy. Sensitivity analysis also allows for exploring the generalisability of results to other settings. The analysis is repeated using different assumptions to examine the effect on the results.<br>Deterministic sensitivity analysis: tests the impact of potential bias resulting from the selection of data sources for key model parameters.<br>One-way sensitivity analysis (univariate analysis): each parameter is varied individually in order to isolate the consequences of each parameter on the results of the study.<br>Probabilistic sensitivity analysis: probability distributions are assigned to the uncertain parameters and are incorporated into evaluation models based on decision analytical techniques (for example, Monte Carlo simulation). |
| **Service delivery guidance** | Recommendations on service delivery primarily aimed at health service commissioners. Service delivery guidance focuses on the broad configuration and provision of clinical services and addresses only those interventions that are likely to have implications for the configuration of services. |
| **Short clinical guideline** | A NICE clinical guideline produced using a 'fast track' process. Short clinical guidelines address only part of a care pathway. |
| **Sign-off** | The approval or acknowledgement of something by or as if by a signature. |
| **Single technology appraisal** | The name given to the NICE process in which appraisals of single technologies for one indication are conducted. |
| **Slide set** | A set of slides that local NHS staff can use to raise awareness of a piece of NICE guidance among healthcare professionals and managers. |
| **Specificity (of a test)** | The proportion of people classified as negative by the gold (or reference) standard, who are correctly identified by the study test. |

| | |
|---|---|
| **Stakeholder** | An organisation with an interest in a topic that NICE is developing a clinical guideline or piece of public health guidance on. Organisations that register as stakeholders can comment on the draft scope and the draft guidance. Stakeholders may be:<br>• manufacturers of drugs or equipment<br>• national patient and carer organisations<br>• NHS organisations<br>• organisations representing healthcare professionals<br>• organisations that fund or carry out research. |
| **Stakeholder scoping workshop** | Workshop attended by registered stakeholders before consultation on the guideline scope, to discuss the key clinical issues identified by the scoping group. |
| **Stochastic analysis** | A cost-effectiveness analysis where both costs and effects are determined from data sampled from the same patients in a study. |
| **Study quality** | The extent to which a study has conformed to recognised good practice in the design and execution of its research methods. |
| **Synthesis of evidence** | A generic term to describe methods used for summarising (comparing and contrasting) evidence into a clinically meaningful conclusion in order to answer a defined clinical question. This can include systematic review (with or without meta-analysis), and qualitative and narrative summaries. |
| **Systematic review** | Research that summarises the evidence on a clearly formulated question according to a predefined protocol using systematic and explicit methods to identify, select and appraise relevant studies, and to extract, collate and report their findings. It may or may not use statistical meta-analysis. |
| **Technical team** | A core technical team from the National Collaborating Centre (NCC) that supports Guideline Development Group members with technical experience and expertise. This team usually includes the NCC Director, an information specialist, a lead systematic reviewer (who can also be the project manager) and a health economist. |

| | |
|---|---|
| **Technology appraisal guidance** | NICE technology appraisal guidance makes recommendations on the use of new and existing drugs and treatments in the NHS. If NICE recommends a drug or treatment for a particular condition, the NHS has to make it available for patients with that condition if it is suitable for them. Usually, this has to be done within 3 months of the guidance being issued. |
| **Technology assessment** | The process of evaluating the clinical, economic and other evidence relating to use of a technology in order to formulate guidance on its most efficient use. |
| **'Test and treat' study** | Studies that compare outcomes of patients after a diagnostic test (in combination with a management strategy) with those of patients who receive the usual diagnostic or management strategy. |
| **Time horizon** | The time span used in a NICE technology appraisal that reflects the period over which the main differences between interventions in health effects and use of healthcare resources are expected to be experienced, taking into account the limitations of supportive evidence. |
| **Treatment allocation** | The process by which study participants are allocated to a treatment group. |
| **Treatment options** | The choices of intervention available. |
| **Understanding NICE guidance** | A summary of a clinical guideline in everyday language for patients, carers and the general public. |
| **Utility** | A measure of the strength of a person's preference for a specific health state in relation to alternative health states. The utility scale assigns numerical values on a scale from 0 (death) to 1 (optimal or 'perfect' health). Health states can be considered worse than death and thus have a negative value. |
| **Utility weight** | The utility weight of a certain health state is most often expressed as a value on a scale of 0 to 1, where 0 represents death and 1 represents perfect health. To measure utility weights of a certain health state (that is, the quality of life experienced when in that health state), large patient surveys are performed using questionnaires such as the EuroQol instrument. |
| **Value-of-information methods** | Formal methods that may be used as part of cost-effectiveness analysis to estimate the 'value for money' of additional research. |

| | |
|---|---|
| **Workplan** | A document prepared by the National Collaborating Centre (NCC) to set out the guideline development process for each guideline. Its purpose is to specify methods, timelines and costings. It is an internal document that becomes a formal agreement between the NCC and NICE, and constitutes the reference from which the progress of the work can be assessed. |

# Appendix N: Guide to the short clinical guideline process

## Contents



# Foreword

This appendix describes the process by which short clinical guidelines are developed. It should be read in conjunction with the rest of 'The guidelines manual' (2009) and, where relevant, with the other NICE documents on contributing to an individual clinical guideline:

- How NICE clinical guidelines are developed: an overview for stakeholders, the public and the NHS
- A guide for patients and carers: contributing to a NICE clinical guideline

# 1 Introduction

1.1 Short clinical guidelines are clinical guidelines that address only part of a care pathway. They are intended to allow the rapid (11–13-month timescale) development of guidance on aspects of care for which the NHS requires urgent advice. This document sets out the process, including timelines, that the National Institute for Health and Clinical Excellence (NICE) follows when developing a short clinical guideline. It describes an open and transparent process designed to achieve robust guidance for the NHS. The document provides guidance for organisations that are invited to contribute to short clinical guidelines, and has been developed to inform consultees and stakeholders and to facilitate their comments on this work programme.

1.2 The document highlights the key differences in the development process for short clinical guidelines compared with that for standard clinical guidelines. The latter is outlined in the chapters of 'The guidelines manual' and in 'How NICE clinical guidelines are developed: an overview for stakeholders, the public and the NHS' (appendix O). Cross-referral is made to the relevant sections of 'The guidelines manual'.

1.3 Each short clinical guideline is developed by an independent Guideline Development Group (GDG) supported by a technical team based within NICE (the Short Clinical Guidelines Team). This technical team is constituted in the same way and undertakes the same functions as the established National Collaborating Centre (NCC) technical teams. The Short Clinical Guidelines Team does not have voting rights on recommendations made by the GDG. The development and quality assurance of short clinical guidelines is overseen by a Guidelines Commissioning Manager, the Director of the Centre for Clinical Practice and an Executive Lead at NICE.

1.4 Occasionally, a short clinical guideline may be externally commissioned by NICE from one of the NCCs; this is decided on a case-by-case basis.

# 2 The short clinical guideline process

## *2.1 Overview*

2.1.1 The short clinical guideline process consists of four phases:

1. Referral of the topic to NICE by the Department of Health.
2. Scoping the short clinical guideline.
3. Development of the short clinical guideline. This begins with the first meeting of the GDG and ends when the draft guideline is submitted for stakeholder consultation.

4. Consultation and publication. This consists of consultation with stakeholders on the draft guideline, revising the guideline in the light of comments received during consultation, receiving advice from the Guideline Review Panel and expert reviewers, preparation of the final draft, carrying out the pre-publication check, sign off by NICE's Guidance Executive and publication.

2.1.2 Each phase of the short guidelines process (topic selection, drafting of and consultation on the scope, development of the short clinical guideline, and consultation and publication) follows the principles set out in 'Social value judgements: principles for the development of NICE guidance (2nd edition)'[9] and NICE's 'Equality scheme and action plan 2007–2010'[10]. These are taken into account when developing the remit and scope and defining the population and management areas to be covered by the guideline; identifying stakeholders and GDG members; developing the review questions; identifying, reviewing and appraising the evidence; developing the recommendations; and producing the guideline publications.

2.1.3 The total time from topic referral to publication is between 11 and 13 months, depending on the length of the development phase. Figure N1 sets out the timeline in more detail.

[9] www.nice.org.uk/aboutnice/howwework/socialvaluejudgements/socialvaluejudgements.jsp
[10] www.nice.org.uk/aboutnice/howwework/NICEEqualityScheme.jsp

## Figure N1 Short clinical guideline process timeline

| Phase | Action | Time taken (weeks) | Elapsed time from initiation of process (weeks) |
|---|---|---|---|
| **1 Topic referral** | | | |
| **2 Scoping** | Registration of stakeholders and invitations to scoping workshop | 0 | 0 |
| | Short Clinical Guidelines Team to draft scope and key clinical issues based on scoping searches | 5 | 5 |
| | Stakeholder scoping workshop | 3 | 8 |
| | Scope revised after workshop | 1 | 9 |
| | Advertisement and appointment of GDG members | | 13 |
| | Public consultation on scope | 4 | 13 |
| | Scope revised and signed off | 2 | 15 |
| | Final scope available on web | 1 | 16 |
| **3 Development** | Development of guideline | 16–26 | 32–42 |
| **4 Consultation and publication** | Public consultation on guideline (including 1 week for editing) | 4+1 | 37–47 |
| | Guideline revised | 3 | 40–50 |
| | Review by Guideline Review Panel | 1 | 41–51 |
| | Pre-publication check | 2 | 42–53 |
| | Guidance Executive sign off | 1 | 44–54 |
| | **Total** | | **11–13 months** |

## *2.2 Phase 1 – referral of topic*

2.2.1 Topics are referred to NICE by the Department of Health (for more details on the topic selection process, see the NICE website[11]). The criteria for the referral to NICE should include both suitability for a short clinical guideline and a judgement about the urgency of the requirement for the advice. The Department of Health is responsible for identifying topics for the short clinical guideline process; proposals for topics may be put forward by the topic selection consideration panels.

## *2.3 Phase 2 – scoping the short clinical guideline*

### 2.3.1 Drafting the scope

2.3.1.1 A draft scope, which defines the areas the guideline will and will not cover, is prepared by the Short Clinical Guidelines Team. It is based on the remit from the Department of Health, input from relevant experts, patients and carers, and a preliminary search of the literature to identify existing clinical practice guidelines, key systematic reviews and other relevant publications. The literature search facilitates an overview of the issues likely to be covered by the guideline – the clinical need for the guideline and the clinical management of the condition – and helps define key clinical issues. It also informs the Short Clinical Guidelines Team of the volume of literature likely to be available in the topic area, and therefore the amount of work required. The draft scope is tightly focused, covering a small number of key clinical issues.

2.3.1.2 **Comparison with the standard clinical guideline process.** The process for drafting the scope broadly follows that outlined for standard clinical guidelines (see chapter 2 of 'The guidelines manual').

### 2.3.2 The scope consultation process

2.3.2.1 Stakeholders are invited to register at the time of formal referral of the guideline topic by the Department of Health. Contact with stakeholders is important to ensure that they are included in the development of the guideline and support it.

2.3.2.2 The draft scope is presented at a stakeholder scoping workshop to a relevant group of stakeholders and professional groups. Attendees are identified in two ways: firstly, by inviting all registered stakeholder organisations to offer suggestions of possible workshop attendees; and secondly, by the Short Clinical Guidelines Team identifying key individuals who are active in the topic area in the UK. One person from each registered stakeholder organisation may attend. The scoping search is used to identify UK-based individuals who have led on recent national published guidelines

[11] www.nice.org.uk/aboutnice/howwework/howguidancetopicsarechosen

and/or recent key reviews in the topic area. Workshop attendees, including representatives of relevant patient and carer organisations, should have specific knowledge or experience in the topic area. The workshop consists of presentations and tightly facilitated parallel-running working groups. The aim is to obtain detailed feedback on the draft scope and agree core areas of care to be covered in the guideline, to seek input about the composition of the GDG and to raise awareness that NICE is publicly advertising for applications for GDG membership.

2.3.2.3 The draft scope is amended to address and/or include issues raised in the workshop. The scope is then subject to a 4-week consultation with stakeholders. Stakeholder comments are reviewed by the Short Clinical Guidelines Team, the GDG Chair and the Clinical Adviser (if one is appointed; see section 2.4.1.1). A revised scope is prepared, which is reviewed by the Guideline Review Panel (GRP). The GRP considers whether stakeholders' comments have been appropriately and adequately addressed by the developers, and the GRP Chair then prepares a report. Subject to any amendments agreed by NICE as a result of the Chair's report, the revised scope is signed off by the Director of the Centre for Clinical Practice at NICE. Stakeholders are notified once the final version of the scope is available on the NICE website.

2.3.2.4 **Comparison with the standard clinical guideline process.** The process for consulting on the scope follows that outlined for standard clinical guidelines (see chapter 2 of 'The guidelines manual').

## *2.4 Phase 3 – development of the short clinical guideline*

### 2.4.1 Forming and running the short clinical guideline GDG

2.4.1.1 Each short clinical guideline is developed by a unique GDG consisting of 10–12 members, supported by the Short Clinical Guidelines Team. Each GDG has a Chair, healthcare professional members and a minimum of two patient and carer members. Co-opted expert advisers are recruited, as appropriate. A Clinical Adviser, who has specific content expertise and additional responsibilities, may also be appointed depending on the topic. Recruitment of the GDG Chair and members is carried out in accordance with NICE's policy 'Appointments to guidance producing bodies advisory to NICE' (November 2006)[12]. Development of the guideline takes 4–6 months and the GDG meets approximately every 4–6 weeks.

2.4.1.2 NICE reserves the option of selecting the GDG Chair and some members for a short clinical guideline from a pool of suitable members. This pool will be recruited through a formal

[12] Available from: www.nice.org.uk/384476

advertisement and recruitment process to act as standing members for each guideline. They will be appointed on 3-year rolling contracts. Healthcare professional members and patient and carer members will be recruited using the standard process. The pool will consist of the following: a) experienced Chairs and b) methodological experts, such as epidemiologists, statisticians and health economists. This option will help foster consistency between the approaches taken with different topics, and will be a more efficient way of setting up GDGs. The system of a unique GDG for every guideline is resource intensive. There is also the risk of ineffective group working, given that the short timeframe requires the GDG to perform as a small group immediately.

2.4.1.3 The GDG makes its decisions using the best available evidence presented to it at GDG meetings by the Short Clinical Guidelines Team. The use of formal consensus methods within the GDG will be considered on a case-by-case basis (see section 3.5 of 'The guidelines manual'). However, formal consensus methods that seek the views of groups outside the GDG are unlikely to be used in the short clinical guideline process because of the short timeframe.

2.4.1.4 **Comparison with the standard clinical guideline process.** The process of forming and running the GDG outlined in section 2.4.1.1 is consistent with that for the standard clinical guideline programme (see chapter 3 of 'The guidelines manual'). However, the guideline development time is 4–6 months compared with up to 18 months in the standard process. The process outlined in section 2.4.1.2 is an adaptation of standard methods.

### 2.4.2 Developing review questions

2.4.2.1 A short clinical guideline has a narrow scope and covers only part of a care pathway. It addresses a maximum of three subject areas covering clinical management. This will result in a small number of key clinical issues (listed in the scope). These are broken down into a defined number of review questions – usually one or two per clinical management area. The exact number will be dictated by the size of the short clinical guideline remit and the amount of development time available. As with the standard clinical guideline programme, it is feasible to present a maximum of two systematic reviews at any one GDG meeting. These review questions are formulated and structured according to the process for standard clinical guidelines (see chapter 4 of 'The guidelines manual').

2.4.2.2 **Comparison with the standard clinical guideline process.** The tightly focused scope and short development phase (4–6 months) mean that between three and six review questions are considered, compared with 15–20 review questions in the standard clinical guideline process.

### 2.4.3 Identifying the evidence

2.4.3.1 The short clinical guideline process follows the standard process for identifying evidence (see chapter 5 of 'The guidelines manual').

### 2.4.4 Reviewing the evidence

2.4.4.1 The short clinical guideline process follows the standard process for assessing and summarising the evidence (see chapter 6 of 'The guidelines manual').

### 2.4.5 Incorporating health economics in the guideline and assessing health-economic impact

2.4.5.1 The short clinical guideline process in general follows the standard process for incorporating health economics in the guideline and assessing health-economic impact (see chapter 7 of 'The guidelines manual'). However, given the short overall timeframe, it will be necessary to consider identifying relevant topics for health-economic analysis during the scoping phase.

### 2.4.6 Creating guideline recommendations

2.4.6.1 Explicit methods of linking the evidence to recommendations are used for short clinical guidelines if the topic is suitable. This involves using the Grading of Recommendations Assessment, Development and Evaluation (GRADE) approach, which is also being implemented in the standard clinical guidelines programme (see section 9.1 of 'The guidelines manual').

2.4.6.2 The smaller number of review questions results in a smaller number of guideline recommendations. The number of recommendations in each short clinical guideline is likely to be between 5 and 20. In addition, because there are usually fewer than 20 recommendations, short clinical guidelines do not generally have key priorities for implementation.

2.4.6.3 Research recommendations are formulated for short clinical guidelines. Their number is dependent on the size of the short clinical guideline remit and the amount of development time available.

2.4.6.4 **Comparison with the standard clinical guideline process.** The short clinical guideline process broadly follows the standard process for creating guideline recommendations (see chapter 9 of 'The guidelines manual').

### 2.4.7 Writing the guideline

2.4.7.1 There are usually three versions of short clinical guidelines:

- The full guideline – all the recommendations, details of how they were developed and summaries of the evidence they are based on.

- The quick reference guide – a summary of the recommendations for healthcare professionals.
- 'Understanding NICE guidance' – a summary for patients and carers.

2.4.7.2 The full guideline is written by the Short Clinical Guidelines Team, following the principles in chapters 9 and 10 of 'The guidelines manual'. The quick reference guide and 'Understanding NICE guidance' are written by NICE editorial staff.

2.4.7.3 In cases where an NCC is commissioned by NICE to develop a short clinical guideline, the full guideline is produced by the NCC. NICE also produces a 'NICE guideline' that contains only the recommendations from the full guideline, without the information on methods and evidence.

### *2.5 Phase 4 – consultation and publication*

2.5.1 Following the development of the draft short clinical guideline, there is a 4-week consultation period for registered stakeholders to comment on the draft guideline.

2.5.2 The formal expert review process that has been established within the Centre for Clinical Practice for standard clinical guidelines is also used for short clinical guidelines (see section 11.2.2 of 'The guidelines manual').

2.5.3 Following consultation with stakeholders, the guideline is revised by the Short Clinical Guidelines Team working in collaboration with the GDG.

2.5.4 The revised short clinical guideline is reviewed by one of the existing GRPs, is subject to a pre-publication check of 10 working days, and is then signed off by NICE's Guidance Executive and published.

2.5.5 **Comparison with the standard clinical guideline process.** The consultation period for short clinical guidelines is 4 weeks, compared with 8 weeks for standard clinical guidelines. For the pre-publication check, the full guideline is posted on the NICE website for a period of 10 working days for short guidelines, compared with 15 working days for standard guidelines. The Short Clinical Guidelines Team works with the GDG in the same way that the NCCs work with their GDGs (see chapter 11 of 'The guidelines manual').

## 3 Linking short clinical guidelines to other NICE guidance

3.1.1 Short clinical guidelines are linked to other NICE guidance in the same way as standard clinical guidelines (see chapter 8 of 'The guidelines manual').

## 4 Updating short clinical guidelines

4.1 Short clinical guidelines are reviewed for consideration of updating by the Short Clinical Guidelines Team using the process for standard clinical guidelines (see chapter 14 of 'The guidelines manual').

# Appendix O: How NICE clinical guidelines are developed: an overview for stakeholders, the public and the NHS

This appendix is available as a separate file from the NICE website (www.nice.org.uk). For a printed copy, phone NICE publications on 0845 003 7783 or email publications@nice.org.uk and quote reference number N1739. This edition replaces the April 2007 edition of 'The guideline development process: an overview for stakeholders, the public and the NHS' (reference N1233).

Also available from the NICE website is 'A guide for patients and carers: contributing to a NICE clinical guideline', which explains how individual patients and carers, as well as patient organisations, can get involved (www.nice.org.uk/nicemedia/pdf/guidelinecontribute_how_to_CG.pdf).

# How NICE clinical guidelines are developed: an overview for stakeholders, the public and the NHS

**Fourth edition**

January 2009

This booklet summarises the process used for developing NICE clinical guidelines from January 2009 onwards, including:

- how Guideline Development Group members are selected
- how organisations can register as stakeholders
- the stages when registered stakeholders can contribute to the development of a clinical guideline.

This booklet is available from the NICE website (www.nice.org.uk). For a printed copy, phone NICE publications on 0845 003 7783 or email publications@nice.org.uk and quote reference number N1739. This edition replaces the April 2007 edition of 'The guideline development process: an overview for stakeholders, the public and the NHS' (reference N1233).

Other documents on the clinical guidelines process are available from the NICE website:

- 'The guidelines manual', which gives full details of the methods for developers of clinical guidelines (www.nice.org.uk/guidelinesmanual)
- 'A guide for patients and carers: contributing to a NICE clinical guideline', which explains how individual patients and carers, as well as patient organisations, can get involved (www.nice.org.uk/nicemedia/pdf/guidelinecontribute_how_to_CG.pdf).

**National Institute for Health and Clinical Excellence**
MidCity Place
71 High Holborn
London, WC1V 6NA

www.nice.org.uk

ISBN: 1-84629-850-4

Published by the National Institute for Health and Clinical Excellence

January 2009

# Contents



### Abbreviations

| | |
|---|---|
| **GDG** | Guideline Development Group |
| **NCC** | National Collaborating Centre |
| **NICE** | National Institute for Health and Clinical Excellence |
| **PPIP** | Patient and Public Involvement Programme |

We welcome comments on this document. These should be emailed to: guidelinesmanual@nice.org.uk

# About NICE guidance

The National Institute for Health and Clinical Excellence (NICE) is the independent organisation responsible for providing national guidance on the promotion of good health and the prevention and treatment of ill health. We produce the following types of guidance:

- **Clinical guidelines** – recommendations about the treatment and care of people with specific diseases or conditions in the NHS in England and Wales[1].
- **Technology appraisal guidance** and **interventional procedures guidance** – guidance on the use of new and existing medicines, treatments and procedures in the NHS[2].
- **Public health guidance** – guidance on ways of helping people improve their health and reduce their risk of illness[3].

## Key point

NICE is committed to promoting equality, eliminating unlawful discrimination and actively considering the implications of our guidance for human rights. We aim to comply fully with all legal obligations to:

- promote race and disability equality, and equality of opportunity between men and women, and
- eliminate unlawful discrimination on grounds of race, disability, age, sex and gender, sexual orientation, and religion or belief in the way we carry out our functions and in our employment policies and practices.

Our equality scheme and action plan* sets out how we are meeting these obligations on equality and discrimination and what we still need to do.

*Available at www.nice.org.uk/aboutnice/howwework/NICEEqualityScheme.jsp

We encourage stakeholders to get involved in the development of our guidance at all stages. Stakeholders can include national organisations that represent patients and carers, local patient and carer organisations when there is no relevant national organisation, healthcare professionals, the NHS, organisations that fund or carry out research, and the healthcare industry.

[1] Clinical guidelines may also apply to Northern Ireland under special arrangements.

[2] Technology appraisal guidance applies to England and Wales; interventional procedures guidance applies to England, Wales, Scotland and Northern Ireland.

[3] Public health guidance applies to England only.

**Key point**

In this booklet we have used the terms 'patients' and 'carers' to cover all groups of lay people (people who are not healthcare professionals) who contribute to the development of our clinical guidelines. This includes:

- people who have the condition or disability
- people such as family and friends who provide unpaid care for them
- organisations representing patients and carers
- voluntary sector and non-governmental organisations.

We recognise that readers may use other terms such as 'consumer', 'service user', 'user representative' or 'patient representative'.

# NICE clinical guidelines

## What is a NICE clinical guideline?

NICE clinical guidelines give recommendations on how healthcare professionals should care for people with specific conditions. The recommendations are based on the best available evidence. Clinical guidelines are also important for health service managers and those who commission NHS services.

Our clinical guidelines can cover any aspect of a condition. This may include recommendations about:

- providing information, education and advice (for example, about self-care)
- prevention
- treatment in primary care (GPs and other community services)
- treatment in secondary care (provided by or in hospitals)
- treatment in specialised services.

The key principles underlying our clinical guidelines are given in box 1.

## Box 1 Key principles underlying NICE clinical guidelines

Our clinical guidelines:

- aim to improve the quality of care for patients
- assess how well different treatments and ways of managing a specific condition work
- assess whether treatments and ways of managing a condition are good value for money for the NHS
- set out the clinical care that is suitable for most patients with a specific condition using the NHS in England and Wales
- take account of the views of those who might be affected by the guideline (including healthcare professionals, patients and carers, health service managers, NHS trusts, the public, government bodies and the healthcare industry)
- are based on the best available research evidence and expert consensus
- are developed using a standard process, and standard ways of analysing the evidence, which are respected by the NHS and other stakeholders, including patients
- make it clear how each recommendation was decided on
- are advisory rather than compulsory, but should be taken into account by healthcare professionals when planning care for individual patients.

A clinical guideline applies to all patients with a particular condition, but there will be times when the recommendations are not appropriate for a particular patient. Healthcare professionals are expected to take our clinical guidelines fully into account when exercising their clinical judgement. However, the guidance does not override the responsibility of healthcare professionals to make decisions appropriate to the circumstances of each patient. These decisions should be made in consultation with, and with the agreement of, the patient and/or their guardian or carer. Healthcare professionals should record their reasons for not following clinical guideline recommendations.

Our clinical guidelines are developed for the NHS, but they may also be relevant to professionals working outside the NHS, such as those working in social care.

### *What are short clinical guidelines?*

Most published NICE clinical guidelines are standard clinical guidelines. A standard guideline covers broad aspects of clinical care and the management of specific conditions.

NICE short clinical guidelines, the first of which was published in 2007, address a smaller part of a care pathway. They are produced more quickly, and generally cover areas for which the NHS requires urgent advice.

The details of how standard and short clinical guidelines are developed differ in a number of ways. The development of a short clinical guideline is usually coordinated by the Short Clinical Guidelines Team at NICE, and not by one of the National Collaborating Centres.

The methods and processes described in 'The guidelines manual' and in this overview are those used for producing standard clinical guidelines. Any differences in the short clinical guideline development process are highlighted throughout this overview in boxes like this one. These differences are also described in more detail in the document 'Guide to the short clinical guideline process', which forms appendix N of 'The guidelines manual'.

## Different versions of NICE clinical guidelines

Four versions of each standard clinical guideline are published (see box 2). We also produce tools to support implementation of the guideline in the NHS.

### Box 2 Versions of the clinical guideline and support for implementation

**The full guideline** contains all the background details and evidence for the guideline, as well as the recommendations. This document is produced by the National Collaborating Centre that is responsible for the guideline (see pages O-11 to O-12).

**The 'NICE guideline'** contains only the recommendations from the full guideline, without the information on methods and evidence.

**The quick reference guide** summarises the recommendations in an easy-to-use format for healthcare professionals.

**'Understanding NICE guidance'** summarises the recommendations in everyday language. It is aimed at patients and their families and carers.

**Implementation support tools** (see page O-34) are produced by NICE to encourage and promote the uptake of guideline recommendations by the NHS. These may include:

- a costing report and costing template
- a slide set
- audit support
- other tools as required.

We publish all versions of the guideline, and the implementation tools, on our website (www.nice.org.uk). We also produce printed versions of the quick reference guide and 'Understanding NICE guidance', and anyone can get a copy.

### *Short clinical guidelines*

There are usually three versions of short clinical guidelines: the full guideline, the quick reference guide and 'Understanding NICE guidance'.

## How are NICE clinical guidelines developed?

Developing a standard NICE clinical guideline takes 18–24 months from the time we are asked to develop it by the Department of Health to its publication. Developing a short clinical guideline takes 11–13 months.

### Proposing and selecting topics for clinical guidelines

Anyone can suggest a guideline topic for consideration. Details of how to do this are on our website (go to www.nice.org.uk/getinvolved and click on 'Suggest a topic').

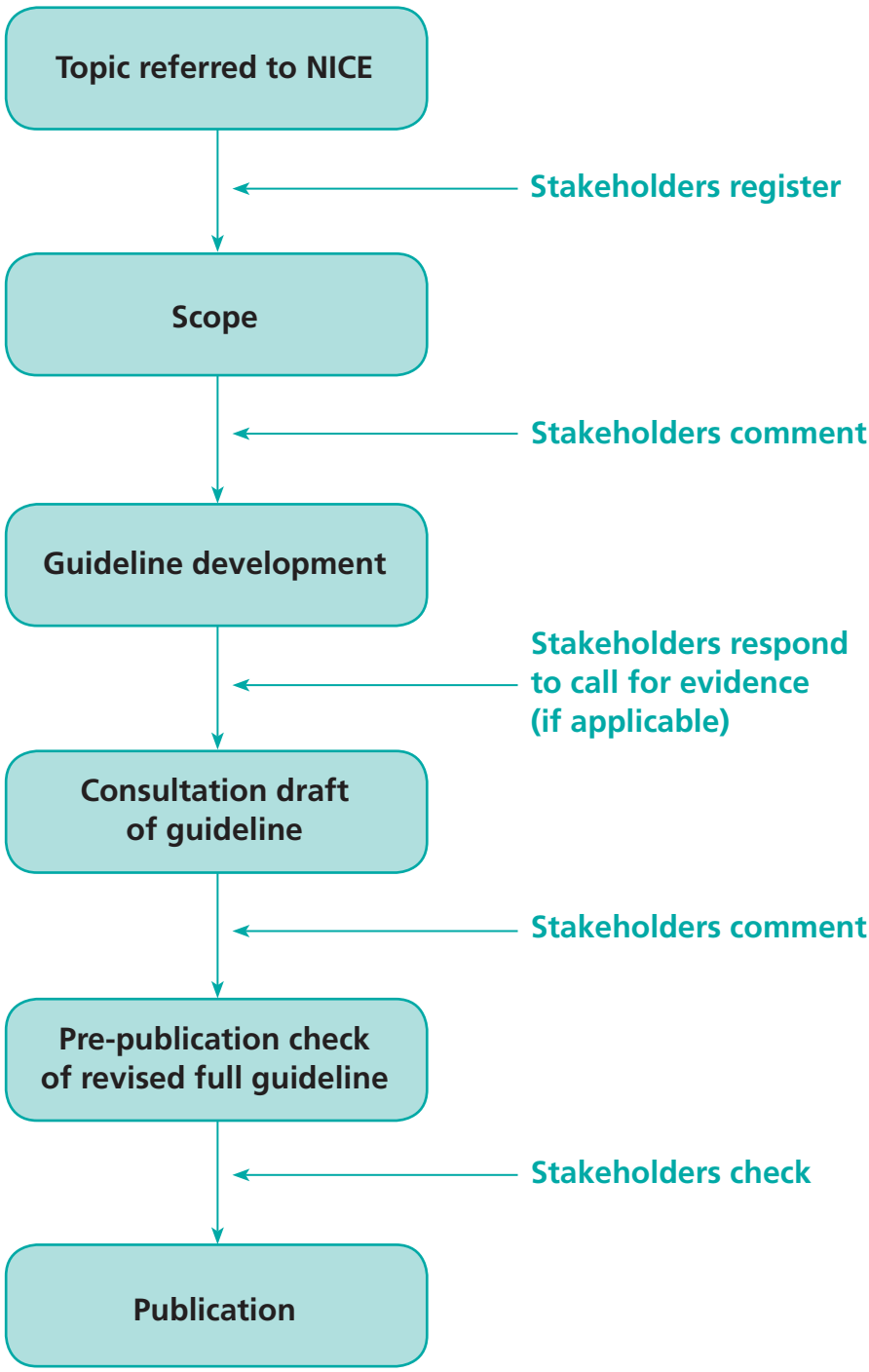
We look at each suggestion we receive to make sure that it is appropriate and that we aren't already producing a clinical guideline in that area. The suggestions are then filtered using a checklist based on selection criteria from the Department of Health. These criteria take into account:

- 'burden of disease' (this includes the number of people affected, the impact of the disease on them and the number of people dying because of it)
- resource impact of the proposed guideline (that is, the likely cost to the NHS, and to other public sector organisations if relevant)
- importance in relation to government policy (that is, whether the topic falls within a 'priority area')
- whether there is variation in clinical practice in different places
- any other reasons why the guideline is needed urgently.

Next, the suggestions are reviewed by 'topic selection consideration panels' composed of experts in the topic, other healthcare professionals with a good knowledge of the NHS, public health and the public sector, and patient and carer members. The recommendations of the topic selection panels go to the Department of Health. The Secretary of State for Health makes the final decision on which topics are referred to NICE for the development of clinical guidelines.

More details about the topic selection process are available on our website.

## Key stages of clinical guideline development



# Who is involved in developing NICE clinical guidelines?

The development of NICE standard clinical guidelines involves:

- NICE
- National Collaborating Centres (NCCs)
- Guideline Development Groups (GDGs)
- the Patient and Public Involvement Programme (PPIP) at NICE
- Guideline Review Panels
- expert reviewers
- stakeholders.

The following sections explain the roles of these various groups.

## NICE

When the Department of Health asks NICE to produce a clinical guideline on a particular topic, we commission one of the NCCs to coordinate the guideline's development. The guidelines team in the Centre for Clinical Practice at NICE supports and advises the NCC throughout the guideline's development.

NICE's 'Guidance Executive' is responsible for giving final approval of ('signing off') the guideline. The Guidance Executive confirms that the NCC has developed the guideline in accordance with the remit from the Department of Health (see page O-24), and by following the correct process and methods.

NICE publishes the 'NICE guideline', the quick reference guide and 'Understanding NICE guidance', as well as the implementation support tools (see box 2).

## The National Collaborating Centres (NCCs)

The NCCs were established by NICE to develop clinical guidelines. The NCCs bring together the expertise of the medical and nursing royal colleges, NHS trusts, professional organisations, and patient and carer organisations. They have the capacity, skills and expertise to produce high-quality clinical guidelines, working closely with the GDGs.

Each NCC has staff with:

- technical skills in:
  - guideline development
  - project management
  - health economics
  - reviewing evidence
  - using formal methods to reach consensus in areas where there is a lack of good-quality evidence
- experience in engaging with patients and with patient and carer groups.

Each NCC also has access to professional networks to support its activities.

## Role of the NCC

For each clinical guideline, the NCC:

- prepares the draft scope and refines it in response to comments received during consultation (see pages O-24 to O-28)
- establishes and works with the GDG to develop the clinical guideline
- undertakes systematic reviews of the literature and health economics analyses
- ensures that the processes described in 'The guidelines manual' are followed, and documents this
- together with the GDG, prepares the consultation draft of the guideline
- together with the GDG, makes changes to the guideline in response to comments received during consultation
- publishes the final full clinical guideline
- advises NICE on the publication, implementation and updating of the guideline.

There is more information about the NCCs on our website[4].

[4] www.nice.org.uk/aboutnice/howwework/developingniceclinicalguidelines/nationalcollaboratingcentres/national_collaborating_centres.jsp

### *Short clinical guidelines*

The Short Clinical Guidelines Team at NICE is responsible for establishing and providing technical support to the GDG for a short clinical guideline. NCCs are not usually involved in the development of short clinical guidelines. NICE publishes all versions of short clinical guidelines.

## Guideline Development Groups (GDGs)

One of the NCCs or the Short Clinical Guidelines Team sets up an independent GDG for each clinical guideline that is developed. GDG members include healthcare professionals, technical experts, and patients and carers who have relevant expertise and experience.

The role of the GDG in developing the clinical guideline is described in detail on pages O-16 to O-20 of this overview, and in chapter 3 of 'The guidelines manual'.

## The Patient and Public Involvement Programme (PPIP) at NICE

The PPIP is an integral part of NICE. Its main role is to work with our guidance-producing teams and with the NCCs so that patients, carers and the public can be fully involved in developing our guidance.

The PPIP team also works with patient and carer organisations, and provides training and support for the individual patient and carer members of GDGs.

### Advice and support to NICE

The PPIP team:

- advises the clinical guidelines team at NICE on patient and carer issues
- advises the Guideline Review Panels on patient and carer issues
- identifies potential patient and carer stakeholders for each clinical guideline topic
- helps in recruiting patient and carer GDG members by promoting vacancies and encouraging applications
- comments from a patient and carer perspective on the clinical guideline development process
- for each guideline, comments from a patient and carer perspective on the draft scope and the draft recommendations.

## Advice and support to the NCCs

The PPIP team:

- advises on ways of involving patients and carers in the work of the NCCs and the GDGs
- encourages and supports applications from patients and carers who want to get involved in the NCCs' activities – such as membership of GDGs and NCC Partners' Boards
- provides dedicated training for patients and carers who are involved in the NCCs' activities.

## Advice and support to patients and carers

The PPIP team:

- advises and supports patient and carer organisations, and individual patients and carers, who are interested in contributing to the development of NICE clinical guidelines
- advises and supports people who apply to become patient and carer GDG members during the application and selection process
- advises, supports and trains appointed patient and carer GDG members
- supports the lay members of Guideline Review Panels.

For information on involving patients and carers in clinical guideline development, see Kelson (2005)[5].

Factsheets accompanying this document explain in more detail how patients and carers, and the organisations that represent them, can get involved in developing our clinical guidelines[6].

[5] Kelson M (2005) The NICE Patient Involvement Unit. Evidence-based Healthcare and Public Health 9: 304–307.

[6] See www.nice.org.uk/getinvolved/patientandpublicinvolvement/patient_and_public_involvement.jsp

## The Guideline Review Panels

There are four independent Guideline Review Panels. Each has four or five members. The healthcare professions, NHS commissioners and managers and the healthcare industry are represented, and there is also a lay member on every panel.

Each clinical guideline is allocated to one of the Guideline Review Panels. The panel:

- comments on the draft scope and the draft guideline
- ensures that stakeholder comments on the draft scope and draft guideline have been responded to appropriately
- makes sure that it will be feasible for the NHS to implement the final recommendations.

There is more information about Guideline Review Panels on our website[7].

## Expert peer reviewers

We commission expert peer reviewers to carry out a statistical and health economics review of each clinical guideline. This takes place during the consultation period for the draft guideline (see pages O-30 to O-32).

## Stakeholders

Stakeholders play an integral part in the development of our clinical guidelines. This is described in detail on pages O-24 to O-35.

[7] www.nice.org.uk/aboutnice/howwework/developingniceclinicalguidelines/guidelinereviewpanels

# The Guideline Development Group (GDG)

## The role of the GDG

The GDG is established by the NCC or the Short Clinical Guidelines Team, and is responsible for developing the clinical guideline.

During development of the clinical guideline, the GDG:

- agrees the questions about treatment and management of the condition that will guide the search for evidence
- considers the evidence and reaches conclusions based on the evidence
- uses expert consensus to make decisions if evidence is poor or lacking
- formulates the guideline recommendations
- considers comments made by stakeholders during consultation
- agrees the necessary changes to the guideline after consultation.

### Key point

GDG members do not comment during the stakeholder consultation on the draft guideline (see page O-30) or during the pre-publication check of the revised full guideline (see page O-32).

There is more information on the role of the GDG in chapter 3 of 'The guidelines manual' (www.nice.org.uk/guidelinesmanual).

## GDG membership

All members of a GDG need to have:

- an interest in and commitment to developing the clinical guideline
- time to attend all meetings (usually 10–15 in total, held monthly)
- time to do the background reading and help formulate the recommendations
- good communication and team-working skills.

### *Short clinical guidelines*

There are usually three 2-day GDG meetings and one 1-day meeting; these are held at approximately 6-week intervals.

Each GDG is made up of healthcare professionals, technical experts and patients and/or carers. The membership reflects the range of stakeholders and groups whose professional activities or care will be covered by the guideline. Every GDG includes at least two members with direct personal experience or knowledge of patient and carer issues. As far as possible, the GDG will have an appropriate balance with regard to the principles of NICE's equality scheme[8]. Expert advisers may also be invited to attend GDG meetings for specific discussions.

NICE is not represented on the GDG, but the Guidelines Commissioning Manager who is responsible for overseeing the clinical guideline may attend meetings as an observer.

The healthcare industry is not represented on GDGs because of potential conflicts of interest. However, manufacturers have input into the clinical guideline development process through the Guideline Review Panels and as stakeholders.

All members of the GDG are expected to abide by the NICE code of conduct and the NICE equality scheme[8] and to declare potential conflicts of interest. On appointment, all GDG members are required to sign a confidentiality form.

GDG members are reimbursed for travel and subsistence. In addition, patient and carer members are offered an attendance allowance, and GPs are offered an allowance to enable them to provide locum cover at their surgeries.

## Becoming a GDG member

Adverts for all GDG vacancies are posted on our website. A brief job description and person specification are provided, together with additional information and details of how to apply. All applicants must complete a declaration of interests form and an equality monitoring form. For details of vacancies and application forms, visit www.nice.org.uk and click on 'Get involved' and then 'Join a NICE committee or working group'.

When selecting GDG members, both of the following are taken into account:

- the suitability of individual applicants, and
- the requirement for the best combination of people to maximise the range of skills and experience of the GDG.

[8] See www.nice.org.uk/aboutnice/howwework/NICEEqualityScheme.jsp

### *Short clinical guidelines*

We may select the GDG Chair and technical members of the GDG (for example, epidemiologists, statisticians and health economists) from a pool of suitable members. This pool will be recruited through a formal advertisement and recruitment process to act as standing members for each guideline.

## GDG Chair

The GDG Chair is appointed before work starts on the scope of the guideline (see page O-24). We inform registered stakeholder organisations about the vacancy. Applicants are required to submit a CV and a covering letter.

The GDG Chair is selected after interview. The selection panel includes the Director of the NCC, the Director of the Centre for Clinical Practice at NICE (or their representative) and a Non-Executive Director of NICE.

### *Short clinical guidelines*

The selection panel for the GDG Chair includes the Director of the Centre for Clinical Practice (or their representative), as well as an Executive Director and a Non-Executive Director of NICE.

## Clinical Adviser

Some GDGs have a Clinical Adviser who is an expert on the topic, and who provides extra support to the GDG. The Clinical Adviser is appointed in the same way as the GDG Chair, before work on the guideline scope begins.

## Patient and carer members of the GDG

A key role of patient and carer members is to ensure that patient issues are considered in everything that the GDG does.

The Patient and Public Involvement Programme (PPIP) team at NICE contacts patient and carer organisations that have registered an interest in the guideline topic to notify them of vacancies. Vacancies are also advertised on our website, and individual patients and carers who are not associated with a particular organisation can also apply.

Patients and carers do not need any formal qualifications to become GDG members, and they are not required to act as a representative of a patient organisation. However, they should meet the following criteria:

- Be familiar with the condition being covered by the guideline and the issues that are important to people with it. For example, they might:
  - have (or have had) the condition themselves
  - be related to and/or care for someone with the condition
  - be a member of a patient organisation.
- Understand the range of experiences of people with the condition. They should be willing to reflect these different experiences, rather than basing their views only on their own experience.
- Have some familiarity with medical and research language. For example, it is helpful if they can understand an abstract from the 'British Medical Journal'. However, training and help will be available.

When considering whether to apply, anyone interested in becoming a patient and carer GDG member should bear the following in mind.

- The clinical guideline will usually cover the entire 'patient journey', from the first time a person contacts a healthcare professional to treatments and long-term care. An understanding of the different stages of the condition is therefore useful. We encourage applications from people with a broad knowledge of the condition. GDG members need the confidence to consider and to discuss all findings from research studies.
- The guideline will cover many aspects of treatment and care. Anyone who is only interested in a specific aspect of care should consider carefully whether they want to apply. The time spent discussing any one issue may be limited, and issues discussed will be restricted to those listed in the guideline's scope. Ideally, applicants should have an interest in, and a willingness to consider the evidence on, a wide range of possible treatments. It is useful for potential applicants to look at the scope (which will be available on our website) to get a clear idea of what the guideline will cover.

### Selection of patient and carer members

Applicants should complete an application form describing how their skills and experience meet the specified requirements. The NCC and the GDG Chair shortlist applicants. Those on the shortlist are interviewed either in person or by telephone. The GDG Chair, with help from the NCC, makes the final decision on which patient and carer members to appoint, and is responsible for notifying both successful and unsuccessful applicants.

### *Short clinical guidelines*

The Short Clinical Guidelines Team usually carries out the tasks described in this section as being the responsibility of the NCC.

## Healthcare professional members of the GDG

Between six and eight members of the GDG should be healthcare professionals ('healthcare professional members') who either treat people with the condition directly or manage services. The NCC and NICE agree a list of professions that will be represented on the GDG to ensure the widest possible range of viewpoints on the topic. If relevant, members from the social care professions will be included.

Healthcare professional GDG members should:

- have an interest in and experience of the guideline topic, but this need not be as an 'expert' – GDGs need to include clinicians who treat patients on a day-to-day basis in the NHS
- be chosen based on their individual skills and experience – they should not be asked to act as a representative of their profession or a professional organisation.

### Selection of healthcare professional members

The NCC informs stakeholder organisations about vacancies for healthcare professional GDG members. Applicants are required to submit a CV and a covering letter.

Healthcare professional members of the GDG are selected by the Director of the NCC and the GDG Chair, subject to confirmation by the Director of the Centre for Clinical Practice at NICE. Applicants may be interviewed.

### Key point

All GDG members are recruited as individuals and not as representatives of particular organisations or professional groups.

# How to register as a stakeholder for a clinical guideline

Stakeholders play a vital role in the development of NICE clinical guidelines. Professional and government organisations, patient and carer groups and companies can all register as stakeholders for a clinical guideline.

### Key point

We encourage stakeholder organisations to register their interest in a particular clinical guideline as soon as possible after the topic is announced. This will enable you to participate in the early stages of the guideline's development (including commenting on the scope). However, you may register your organisation as a stakeholder at any time during the development process. You can then be involved in the remaining stages of the guideline's development.

## How NICE alerts potential stakeholders

We announce several new topics for clinical guidelines at the same time, after they are referred by the Department of Health. This usually happens three times a year. We publicise these new topics by:

- issuing a press release
- listing the topics on our website, with details of how to register as a stakeholder
- contacting organisations that registered as stakeholders for previous clinical guidelines to alert them to the new topics
- writing to other patient and carer and professional organisations that may have an interest in a new guideline topic
- writing to relevant consultees for a technology appraisal if the clinical guideline may update the appraisal (for further details, see section 8.1.2 of 'The guidelines manual').

## Organisations that can register as stakeholders

The following can register as stakeholders for NICE clinical guidelines:

- national patient and carer organisations that represent the interests of people whose care will be covered by the guideline ('patient and carer stakeholders')
- local patient and carer organisations, but only if there is no relevant national organisation
- national organisations that represent the healthcare professionals who provide the services described in the guideline ('professional stakeholders')
- companies that manufacture drugs or devices used in treatment of the condition covered by the guideline and whose interests may be significantly affected by the guideline ('commercial stakeholders')
- providers and commissioners of health services in England and Wales
- statutory organisations, including the Department of Health, the Welsh Assembly Government, NHS Quality Improvement Scotland, the Healthcare Commission and the National Patient Safety Agency
- research organisations that have carried out nationally recognised research in the area.

NICE clinical guidelines are produced for the NHS in England and Wales, so a 'national' organisation is defined as one that represents England and/or Wales, or has a commercial interest in England and/or Wales.

## Organisations that cannot register as stakeholders

For reasons of capacity, local patient and carer and professional groups cannot register as stakeholders unless there is no national organisation representing the group's specific interests.

Individuals cannot register as stakeholders. However, we encourage anyone with an interest in the topic to participate by contacting a registered stakeholder and expressing their views to them. The registered stakeholders for each guideline are listed on our website.

## How to register

To register an interest in a particular clinical guideline, you should complete the stakeholder registration form. This can be done via our website[9],or you can ask us for a printed copy of the form.

The form asks potential stakeholders to:

- provide a brief description of their organisation
- indicate who the organisation represents
- describe the contribution that the organisation can make to the guideline
- provide contact details of the person who will be the stakeholder contact for the organisation.

If an organisation fits the definition of a stakeholder, we will confirm the registration. If you have not received a confirmation within 28 days of submitting the form, contact the NICE guidelines team (guidelines@nice.org.uk).

We cannot guarantee that all organisations that may have an interest in a particular clinical guideline topic will be notified about new topics. We strongly encourage potential stakeholders to visit our website regularly to check the list of guideline topics and register for appropriate guidelines.

## Once an organisation has registered as a stakeholder

We encourage registered stakeholder organisations to check the summary pages about the guideline on our website regularly. You can also subscribe free of charge to our monthly e-newsletter 'NICE news', which lists forthcoming guidance, consultations on guidance that are in progress, and future events. The e-newsletter is also available on our website.

[9] www.nice.org.uk/ourguidance/niceguidancebytype/clinicalguidelines/shregistration/shregistration.jsp

# How stakeholders can get involved

Stakeholder organisations can contribute to and comment on the clinical guideline at various stages during its development. A summary of the clinical guideline development process showing the key points of stakeholder involvement is on page O-10.

Stakeholder involvement is managed by the Centre for Clinical Practice working with the PPIP at NICE.

### *Short clinical guidelines*

All tasks in this section described as being the responsibility of an NCC will usually be carried out by the Short Clinical Guidelines Team at NICE.

## The scope

### What is the scope?

The Department of Health gives NICE a short 'remit' for each clinical guideline. The next stage is to define exactly what the guideline will and will not cover. This process is called 'scoping', and the document containing this information is the scope.

The scope is drafted by the staff at the NCC, with input from the GDG Chair, the Clinical Adviser (if there is one), and the guidelines team and the PPIP team at NICE.

The scope gives an overview of what the clinical guideline will and will not include, and defines the aspects of care that it will cover. It may describe:

- groups of patients whose care is to be included or excluded – for example, particular age groups, or people with certain types of disease
- where treatment will be carried out – for example, by GPs (primary care), in hospital (secondary care) or in specialist units (tertiary care)
- treatments to be included and excluded – for example, diagnostic tests, surgical, medical and psychological treatments, rehabilitation.

The scope should also identify topics from other NICE guidance programmes (that is, technology appraisal, interventional procedures and public health guidance) that are relevant to the clinical guideline. For more information, see chapter 8 of 'The guidelines manual'.

## The stakeholder scoping workshop

We arrange a workshop for all registered stakeholder organisations before public consultation on the scope. Key staff from the Centre for Clinical Practice at NICE, the PPIP and the NCC attend, as well as the GDG Chair and (if applicable) the Clinical Adviser. People attending the meeting are sent a first draft of the scope, which is intended as a starting point for discussion. At the workshop we:

- provide an overview of the NICE clinical guideline development process
- describe how stakeholders can contribute to the guideline by:
  - commenting during the consultations on the draft scope and draft guideline
  - informing their members and associates about GDG vacancies
- discuss the first draft of the scope and hear stakeholders' views on the key clinical issues that the guideline will cover.

### What to do before the workshop

Each registered stakeholder can send one person to the workshop – please tell us who will be attending from your organisation. The person who attends should have a good understanding of the guideline topic. People attending from patient and carer organisations should have a good understanding of issues relating to the scope from a patient or carer perspective.

Note that each person is attending the workshop from their own perspective, and not to represent the views of their stakeholder organisation.

### Key point

The stakeholder scoping workshop takes place before the public consultation on the scope. Note that expressing views at the workshop does not replace the formal scope consultation process. You should still send comments on the scope to NICE during the consultation.

## Commenting on the draft scope

The NCC, GDG Chair, Clinical Adviser (if there is one) and NICE consider the issues raised at the scoping workshop and refine the draft scope for consultation. The draft scope is then posted on our website for a 4-week consultation period. We send a link to the document to registered stakeholders. Consultation dates are given on the website and in our monthly e-newsletter. Stakeholders should check the website regularly for any changes to timings.

We ask stakeholders to submit comments on the draft scope using the form provided. When commenting, it is important to take account of what NICE clinical guidelines can realistically be expected to cover (see box 3).

Some notes on how to comment during consultation are given in box 4 (these also apply to commenting on the draft guideline – see pages O-30 to O-31).

### Box 3 Considerations when commenting on the draft scope

- NICE clinical guidelines apply to the NHS only, so they will not address the independent sector specifically. However, whenever an independent hospital, clinic or care home, social services or the voluntary sector is commissioned to provide NHS-funded care, it will be expected to adhere to NICE guidelines.
- Guidelines are generally published within 2 years of the development process starting (1 year for short clinical guidelines), so that information is up to date at publication. If the scope is very wide it will not be possible to complete the work in this time, so the scope must be restricted to what can realistically be covered.
- Guidelines will, if appropriate, address what drugs to use. However, it is assumed that prescribers will use the summaries of product characteristics* of medicines they are considering prescribing for individual patients. Therefore guidelines do not usually contain detailed information on contraindications and side effects.
- The scope may specify or exclude certain groups of patients. It is helpful if stakeholders can comment on whether such inclusions or exclusions may discriminate on the grounds of race, disability, sex and gender, age, socioeconomic status, sexual orientation and gender reassignment, or religion or belief.
- Clinical guidelines can cover any aspect of healthcare, but do not generally address how services are organised, or the skills or staff required. The scope sometimes includes aspects of service delivery, but only if the Department of Health has requested this.

*The summary of product characteristics for a drug includes information on uses for which the drug is licensed, dosages and contraindications. Summaries of product characteristics can be found at www.emc.medicines.org.uk

## Box 4 A guide to commenting on drafts of the scope and the guideline

When the draft scope or guideline arrives, you should:

- circulate the draft within your organisation if appropriate, making it clear that it is for consultation and asking recipients to respond to you as the organisation's stakeholder contact (rather than responding directly to NICE)
- prepare your response and return it to NICE, remembering to:
  - collate the comments into one response from your organisation using the form provided (do not make changes to the draft document)
  - include the name of your organisation in the response
  - return the response by the closing date
- send comments electronically to the dedicated email address provided, adding your organisation's name in the subject box.

Please keep in mind the following:

- We will accept only one response from each registered stakeholder organisation. If several responses are received, it may be unclear which represents the view of the organisation. We do not have the resources to acknowledge or respond to comments from several individuals within a registered stakeholder organisation.
- All comments received from registered stakeholders will be made public on our website, so do not include confidential information (such as information about individual patients).
- Make sure that comments are constructive and clearly worded.
- We will not consider comments that are not prepared according to these instructions, or that arrive after the deadline.
- The Guidelines Coordinator (whose name is on the guideline page on our website) can answer questions on submitting comments.

Please see the document 'Protocol for managing guidance consultation comments'* for further details about how we deal with stakeholder comments received during consultation.

*www.nice.org.uk/media/307/97/Managingstakeholdercomments.pdf

**Key point**

Comments on the draft scope must be submitted by the end of the 4-week consultation period, using the form provided by NICE. We notify registered stakeholders of the deadline for submitting comments.

### The final scope

We collect together the stakeholder comments on the scope into a 'scope consultation table'. The NCC then finalises the scope, taking into account the comments received. We 'sign off' the final version of the scope, with the approval of the Guideline Review Panel Chair. The final scope is then posted on our website, along with the scope consultation table, which contains the NCC's responses to stakeholder comments.

## The clinical guideline

### Evidence from stakeholders

The NCC and GDG draft 'review questions' for the guideline from the key clinical issues defined in the final scope. Each review question takes account of issues that are important to patients, such as acceptability of treatment and patients' preferences for treatment options. There is more information about review questions, including examples, in chapter 4 of 'The guidelines manual' (see www.nice.org.uk/guidelinesmanual). A search of the scientific literature is carried out to answer the review questions.

For some of the review questions, the GDG and NCC may believe that their literature search has not found all the relevant information. For example:

- the NCC may be aware that further research is being carried out
- a drug or medical device may be relatively new
- studies may have been published only as abstracts
- the NCC may be looking for data on side effects, economic models or studies of the experiences of patients, carers or healthcare professionals.

In these situations, the NCC may call for evidence from stakeholders. They will specify the review question and the type of evidence they are looking for. These calls for evidence will be sent to all registered stakeholders, and may be made at any point during development of a clinical guideline. Stakeholders are usually given 4 weeks to respond.

As well as published studies, stakeholders may submit relevant unpublished data or studies. Any confidential information should be clearly marked (for example, by using a highlighter pen, or the highlighter function in an electronic version). The NCC also asks stakeholders to complete a checklist that lists and identifies the location of all of the confidential information contained in their submission.

Box 5 summarises what may, and may not, be considered confidential by NICE.

### Box 5 A guide to submitting confidential information

- Data that may influence share price values ('commercial in confidence') or are 'intellectual property' (that is, awaiting publication) may be considered as confidential.
- Information marked as confidential should be kept to an absolute minimum, for example just the relevant part of a sentence or a particular result from a table.
- NICE will not agree to a whole study being designated as confidential. As a minimum, a structured abstract of the study or economic model will have to be made available for public disclosure during consultation on the clinical guideline.
- Results derived from calculations using confidential data will not be considered confidential unless releasing those results would enable back-calculation to the original confidential data.

It is important that the amount of confidential information in a submission is kept to a minimum. At the least, a summary should be publicly available by the time the draft guideline is consulted on. We need to be able to justify the recommendations in our clinical guidelines on the basis of the evidence considered by the GDG, so the guidelines team and the NCC will work with the data owners to find an agreed solution to the balance between confidentiality and transparency[10].

[10] For further details see the document 'Agreement between the Association of the British Pharmaceutical Industry (ABPI) and the National Institute for Clinical Excellence (NICE) on guidelines for the release of company data into the public domain during a health technology appraisal' (www.nice.org.uk/page.aspx?o=229411).

The types of information listed in box 6 will not be considered by the GDG.

### Box 6 Stakeholder material not eligible for consideration

- Studies with weak designs if better-designed studies are available.
- Promotional literature.
- Papers, commentaries and editorials that interpret the results of a published paper.
- Representations and experiences of individuals (unless assessed as part of a well-designed study or a survey).

## Consultation on the draft clinical guideline

The GDG takes 12–18 months to develop a draft of the clinical guideline once the scope has been finalised. There is then an 8-week consultation period when registered stakeholders can comment on the draft guideline.

### *Short clinical guidelines*

- Development of the draft guideline takes 4–6 months.
- The consultation period for the draft guideline is 4 weeks.

We notify registered stakeholders by email when the consultation draft of the guideline is posted on our website. Comments should be submitted using the form provided via the dedicated email address for the guideline. When commenting on the guideline, stakeholders should consult the final scope (on our website) to check what the guideline will and will not cover.

Stakeholders can comment on the full guideline (which includes the draft recommendations as well as explanations of how the GDG has interpreted the evidence to make the recommendations) and/or the 'NICE guideline' (which contains just the draft recommendations and only brief supporting information).

Issues that stakeholders may wish to comment on during consultation include:

- a general view (either positive or negative) of the quality and content of the draft guideline
- points or areas that appear to fall within the scope but are not covered in the draft guideline
- any gaps in the evidence that the recommendations are based on
- potential inconsistencies in the interpretation of the evidence
- disagreements with the interpretation of the evidence
- the practical value of the guideline
- wording (for example, could the recommendations be clearer, or the language more patient-centred; could the wording be perceived as excluding patients or groups of patients?)
- whether the recommendations discriminate against some groups on the grounds of race, disability, sex and gender, age, socioeconomic status, sexual orientation and gender reassignment, or religion or belief
- how easy the recommendations will be to implement
- the potential cost of implementing the recommendations.

Some notes on how to comment on the draft guideline are given in box 4 (see page O-27).

**Key point**

There is a single consultation period when registered stakeholders can comment on the draft clinical guideline (8 weeks for standard guidelines and 4 weeks for short guidelines). The GDG will not consider comments that are submitted late.

## Finalising the clinical guideline

We collect together all the comments from registered stakeholders in a 'guideline consultation table', and pass them to the NCC to consider. The NCC adds its responses to the consultation table.

In very rare cases, we may decide to hold a second consultation on all or part of the guideline (see section 11.4 of 'The guidelines manual' for more details).

The NCC makes changes to the guideline in the light of comments made during the consultation by:

- registered stakeholders
- the Guideline Review Panel
- external reviewers (see section 11.2.2 of 'The guidelines manual')
- other teams at NICE (such as the PPIP, the editors and the implementation team).

Comments from the Guideline Review Panel and from NICE staff are entered into the guideline consultation table and are responded to in the same way as comments from registered stakeholders, but they are not posted on our website.

In response to advice from the Guideline Review Panel, and in consultation with the GDG, the guideline is revised.

## The pre-publication check

The pre-publication check enables registered stakeholders to point out any factual errors and inaccuracies that exist in the revised full guideline after consultation. More details are given in section 12.2 of 'The guidelines manual'.

A pre-publication check is not a second consultation or an opportunity to reopen arguments and issues on which the GDG has made recommendations. Nor is it an opportunity for stakeholders to ask why the guideline has not been amended in response to their comments. New evidence will not be accepted.

Factual errors are instances where there is an error of fact in the proposed full final guideline that should be corrected before publication. Factual errors do not include disagreements surrounding scientific or clinical interpretation or judgement. Box 7 gives examples of what we may consider to be a factual error.

### Box 7 Examples of what may be considered as a factual error

- Incorrect referencing of studies, for example wrong year or wrong journal.
- Errors in the transcription of data, for example '4.9 months' instead of '4.9 years', '£100' instead of '£1000'.
- Incorrect reference to the licensed indications of a drug.
- Errors of fact in appraising a study, such as describing it as randomised when it wasn't.

The pre-publication check happens after the NCC and the GDG have responded to stakeholder comments from consultation on the draft guideline, after the Guideline Review Panel has reviewed the stakeholder comments and responses, and before our Guidance Executive approves the final version of the guideline. However, final editing of recommendation wording may take place after the pre-publication check.

The full guideline is posted on our website for a period of 15 working days, along with the guideline consultation table that lists comments received during consultation from stakeholders and the responses from the NCC and GDG. Registered stakeholders are alerted by email. Stakeholders are invited to report factual errors using a standard form. Reports of errors received after the 15-working-day period, from non-registered stakeholders, or in a format other than using the standard form are not considered.

### *Short clinical guidelines*

The full guideline is posted on our website for the pre-publication check for a period of 10 working days.

NICE, the NCC and the GDG Chair consider the reports of errors received from registered stakeholders and respond only to those related to factual errors as defined above. A decision is made on whether corrections to the guideline are needed.

If corrections are not needed, the guideline is considered and submitted to NICE's Guidance Executive for approval ('sign-off'). If corrections are needed, these are carried out and the full guideline is revised by the NCC and resubmitted to NICE, together with a table of comments about the factual errors and the NCC's responses. The revised guideline is submitted to Guidance Executive for approval.

After sign-off, the different versions of the guideline are published as described below.

## Publication

Once Guidance Executive has given final approval of ('signed off') the clinical guideline, the different versions are published (see box 2). Registered stakeholders are notified when the guideline is published. If applicable, the comments and responses from the pre-publication check are published on our website along with the final guideline.

Any stakeholder comments on the published guideline (other than those about errors that require correction) are addressed when the guideline is updated (see page O-35).

# After publication

## Implementation support

Stakeholders are encouraged to use their networks and influence to encourage implementation of the clinical guideline at both national and local level.

We develop tools to help the NHS implement our clinical guidelines, and these are available on our website. These routinely include the following:

- costing tools:
  - a costing report that estimates the national savings and costs associated with implementation
  - a costing template that can be used to estimate the local costs and savings involved
- a slide set (in the form of a PowerPoint presentation) that highlights the key priorities and provides a framework for local discussion
- clinical audit support to help monitor and review local practice.

Depending on the topic, we may also produce other tools. These can include implementation advice to aid with action planning at an organisational level, referral letter templates, flow charts, fact sheets and checklists. Tools may be produced jointly with other organisations such as professional or patient groups.

## Comments and correcting errors

Comments on published clinical guidelines should be sent to us at nice@nice.org.uk

Sometimes a comment after publication may highlight a potential error in a clinical guideline. This might be in either the interpretation or the presentation of the evidence considered by the GDG. In these cases the Director of the Centre for Clinical Practice and the NCC will consider whether the potential error:

- may result in harm to patients
- undermines the conclusions on which the recommendations were based
- indicates serious problems with our quality-assurance procedures.

If one of these criteria is met, the comment will be referred to our Guidance Executive, which decides what action to take. If the Guidance Executive does not accept that an error has been

made, the individual or organisation that made the comment will be notified. If the Guidance Executive accepts that an error has been made, a note will be put on our website, and the versions of the document on the website will be amended. Depending on the nature and significance of the error and the time since publication, registered stakeholders may also be notified in writing.

## Reviewing and updating clinical guidelines

There is a formal process for reviewing and updating clinical guidelines, which is managed by NICE and the NCC. Chapter 14 of 'The guidelines manual' (www.nice.org.uk/guidelinesmanual) gives details of this process.

Usually a guideline is considered for updating 3 years after publication. In order to be brought up to date, a guideline may require:

- a full update (in exceptional circumstances)
- a partial update
- no update.

Other possible options are:

- transferring the guideline to a 'static list'
- withdrawing the guideline.

A partial update may also be carried out before the usual 3 years if significant new evidence emerges.

In cases where there is to be a full update, or a partial update where new key areas are to be included in the guideline, the usual process for producing and consulting on the scope is followed (see pages O-24 to O-28). The time needed to conduct a partial update is agreed between NICE and the NCC, but will be no more than 18 months. Stakeholders are informed.

A partial update of a guideline may also be carried out when some recommendations need updating but no new areas need to be included. In these cases the original scope is used and is not consulted on. NICE informs the stakeholders that it is conducting a partial update.

A guideline will be transferred to a 'static list' if the recommendations are unlikely to change in the foreseeable future, and so no further update is planned.

A guideline may be withdrawn if its recommendations no longer apply, but it is not a sufficient priority for updating. This decision will be consulted on with stakeholders.

## General information about clinical guidelines on the NICE website (www.nice.org.uk)

Our website contains the following general information about NICE and clinical guidelines:

- contact details for NICE
- lists of clinical guidelines that are published and in development
- stakeholder registration form
- information on NICE staff involved in producing clinical guidelines
- information on the NCCs
- information on the Guideline Review Panels
- information on topic selection
- general information about how clinical guidelines are developed
- 'The guidelines manual', which gives more detailed information about the methods used for developing NICE clinical guidelines
- advertisements for the positions of GDG Chair and GDG members for each clinical guideline
- general information on the implementation of clinical guidelines:
  - implementation tools
  - examples of how organisations have successfully met the challenges of putting NICE guidance into practice (the shared learning database)
- details of NICE commissioning guides, which provide support for the local implementation of clinical guidelines through commissioning
- information on NICE's Patient and Public Involvement Programme (PPIP)
- information on other NICE guidance.

## Information about individual clinical guidelines

The following details for each clinical guideline will be made available on our website (www.nice.org.uk), and updated regularly:

- the remit from the Department of Health
- a list of registered stakeholders
- contact details of the NCC coordinating development of the guideline
- a schedule for development of the guideline
- the consultation draft of the scope
- the final scope
- a table of stakeholder comments on the consultation draft of the scope and responses
- project history, and information on the progress of the guideline
- members of the GDG
- minutes of GDG meetings
- the consultation draft of the guideline
- a table of stakeholder comments on the consultation draft of the guideline and responses
- the 'pre-publication' version of the guideline
- a list of factual errors in the pre-publication version of the guideline reported by stakeholders (if applicable) and responses
- details of related NICE technology appraisal, interventional procedure and public health guidance
- all versions of the published guideline – full guideline, 'NICE guideline', quick reference guide and 'Understanding NICE guidance'
- tools to support implementation of the guideline.