NICE National Institute for Health and Care Excellence

Developing NICE guidelines: the manual

NICE process and methods
Published: 31 October 2014
Last updated: 29 May 2024

www.nice.org.uk/process/pmg20

Contents

1 Introduction

The National Institute for Health and Care Excellence (NICE) is a non-departmental public body that provides national guidance and advice to improve health and social care in England. Decisions on how NICE guidelines apply in other UK countries are made by ministers in the Welsh Government, Scottish Government and Northern Ireland Executive.

This manual explains the processes and methods used to develop and update NICE guidelines, the guidance that NICE develops covering topics across clinical care (in primary, secondary and community care settings), social care and public health. For more information on the other types of NICE guidance and advice (including technology appraisal guidance), see about NICE.

1.1 NICE guideline recommendations

NICE guidelines make evidence-based recommendations on a wide range of topics, from preventing and managing specific conditions, improving health, and managing medicines in different settings, to providing social care and support to adults and children, and planning broader services and interventions to improve the health of communities. Guidelines promote individualised care and integrated care (for example, by covering transitions between children's and adult's services and between health and social care).

Guideline recommendations set out:

- the care and services that are suitable for:
  - most people with a specific condition or need
  - particular populations, groups or people in particular circumstances or settings (for example, when being discharged from hospital)
- ways to promote and protect good health or prevent ill health
- the configuration and provision of health and social care services **and/or**
- how national and local public sector organisations and partnerships can improve the quality of care and services (for example, how the NHS and social care services work together).

Many guideline recommendations are for individual health and social care professionals, who should use them in their work together with their own judgement and discussion with people using services. Some recommendations are for local authorities, commissioners and managers, and cover planning, commissioning and improving services; others are for providers (organisations providing services), schools, and local and national organisations and partnerships in the public, private and voluntary sectors. Guideline recommendations are also useful for people who use health and social care services (including people who purchase their own social care), their families and carers, and organisations representing their interests.

In addition to the recommendations, guidelines also summarise the evidence behind the recommendations and explain how the recommendations were derived from the evidence.

1.2 Information about this manual

This manual explains the processes and methods NICE uses for developing, maintaining and updating NICE guidelines. It is primarily for:

- NICE staff involved in developing guidelines
- organisations contracted by NICE (such as those doing evidence reviews, economic analysis and other engagement work)
- members of the committees that develop the guidelines (see the section on who is involved).

It is also likely to be of interest to a broader audience, including developers of other guidance, stakeholders and users of NICE guideline recommendations.

The processes and methods described in this manual are based on internationally recognised standards, and the experience and expertise of the teams at NICE, the contractors that work with NICE, NICE committee members and stakeholders. They are based on internationally accepted criteria of quality, as detailed in the AGREE Enterprise's Appraisal of Guidelines for Research and Evaluation II (AGREE II) instrument, and primary methodological research and evaluation undertaken by the NICE teams.

As NICE changes its processes and methods in line with the NICE strategy for 2021 to 2026, we will update some recommendations rapidly in high priority areas to provide useful and useable content to users. This manual describes how we develop new

guidelines and update our existing guideline recommendations.

This manual covers methods and processes for clinical, public health and social care topics, and service guidance (see the appendix on service delivery – developing review questions, evidence reviews and synthesis).

NICE uses a different process for producing guidelines relating to health and social care emergencies. This process is covered in the appendix on process and methods for guidelines developed in response to health and social care emergencies.

NICE is piloting a different process for producing digital living guideline recommendations. The principles are covered in the appendix on interim principles for methods and processes for supporting digital living guideline recommendations and the appendix on surveillance decision framework and multi-criteria decision framework for deciding whether to develop or update recommendations and which methods to use.

1.3 Key principles for developing NICE guideline recommendations

We develop guideline recommendations according to our principles that guide the development of NICE guidance and standards.

Using the best available evidence to develop recommendations

NICE guideline recommendations are based on the best available evidence. We use a wide range of different types of evidence and other information – from scientific research using a variety of methods, to testimony from practitioners and people using services.

Review questions guide the search for evidence, and the type of evidence used depends on the type of question (see the chapter on developing review questions and planning the evidence review).

Whatever evidence is used, it is selected and quality assessed using clear, consistent and appropriate methods (see the chapters on identifying the evidence: literature searching and evidence submission, reviewing evidence and incorporating economic evaluation).

Involving people affected by the guideline

When developing guideline recommendations, we involve people who might be affected by the guideline recommendations in a collaborative and transparent way. This includes commissioners, practitioners and others involved in providing services. People using health and care services, carers and the public also contribute to ensure that guidelines address issues relevant to them, reflect their views, and meet their health and social care needs.

There are 2 main ways to get involved:

- organisations can register as a stakeholder
- individuals can apply to join a committee that works on developing guideline recommendations or individuals may be asked to advise a committee as co-opted experts on a particular issue.

There is more information about stakeholders and committee members in the section on who is involved and in our guide for stakeholders and the public on how to get involved in developing NICE guidelines.

The people and communities team at NICE provides advice and support to committees and NICE staff about involving the public in developing NICE guidelines.

The Guidelines International Network Public Toolkit includes our best practice examples on how to involve patients and the public in guideline development.

Practitioners and people who use health and care services, family members, carers and the public may also be involved as:

- expert witnesses invited to give testimony to the committee (see the section on other attendees at committee meetings in the chapter on decision-making committees)
- members of a reference group, focus group or other advisory group set up when standard involvement and consultation processes are likely to be insufficient (for example, when the topic covers a population group that is not part of the committee, such as children, or people with advanced dementia [see the section on what happens during consultation in the chapter on the validation process for draft guidelines, and dealing with stakeholder comments]).

Advancing equality and diversity, and upholding NICE principles

We are committed to ensuring that our guideline development process:

- fully meets duties under the Equality Act (2010) to have due regard to the need to eliminate discrimination, foster good relations and advance equality of opportunity in relation to people who share the protected characteristics of age, disability, gender reassignment, marriage and civil partnership, pregnancy and maternity, race, religion or belief, sex, and sexual orientation, including the public sector equality duty to tackle discrimination and provide equality of opportunity for all
- enables us to meet requirements under the Human Rights Act (1998)
- fully meets duties under the Health and Social Care Act 2012 to consider the degree of a person's need for health services or social care in England.

NICE's equality objectives and equality programme (2020 to 2024) summarises our legal and other equality obligations and describes our approach to meeting them. In addition, NICE's strategy for 2021 to 2026 reaffirms our commitment to identifying and reducing health inequalities through our work.

Reducing health inequalities

Health inequalities arise because of the conditions in which we are born, grow, live, work and age. Health inequalities can be considered across 4 dimensions:

- socio-economic status and deprivation (for example, unemployment, poor housing, poor education, low income or people living in deprived areas)
- protected characteristics defined in the Equality Act 2010
- vulnerable groups of society, or 'inclusion health' groups (for example, vulnerable migrants, people who are homeless, sex workers, and Gypsy, Roma and Travellers)
- geography (for example, urban or rural areas, coastal areas).

Health inequalities can be measured through examining differences in 5 domains:

- health status (for example, life expectancy and prevalence of health conditions)
- behavioural risks to health (for example, smoking, diet and physical activity)

- wider determinants of health (for example, income, education and access to green spaces)
- access to care (for example, availability of treatments)
- quality and experience of care (for example, levels of patient satisfaction).

Having due consideration for groups that may be affected by equality and health inequalities issues is an aspect of our compliance with both general public law requirements to act fairly and reasonably, and human rights obligations. It is also aligned to duties placed on the integrated care systems as outlined in the Health and Care Act 2022. We also have a moral, leadership and strategic duty to address health inequalities given our reputational role in delivering robust, independent and trusted advice to the UK health and care system.

Approaches to reducing health inequalities

We use evidence-based approaches to help identify and address equality and health inequalities issues throughout the guideline development process by:

- systematically identifying population groups that may experience health inequalities using an equality and health inequalities assessment form, which considers the 4 dimensions of health inequalities
- building on the key principles of co-design, co-production and community engagement to include diverse voices and perspectives that can help identify health inequalities and inform actions to reduce them
- proactively considering whether recommendations can be made to advance equality and reduce health inequalities.

1.4 Choice of guideline topics

We may develop guideline recommendations on new topic areas or update existing guideline recommendations. Decisions on which new topics to develop recommendations on will be based on the NICE-wide topic prioritisation process. This process ensures that NICE guidance reflects national priorities.

Further details on this process are available in the NICE-wide topic prioritisation: the

manual. Briefly this process involves:

- The identification of national priorities for the health and care system.
- Consideration of whether a new topic area or update of an existing area addresses and national priority and meets the eligibility criteria of being within NICE's remit and there is either a gap in the existing NICE portfolio or there is significant and unwarranted variation in practice.
- Where these criteria are met there is a further assessment via the prioritisation framework (alongside the NICE strategic principles for public health, social care and rare disease, if relevant).
- The prioritisation board will make decisions on if the topic should be prioritised for further development and what type(s) of guidance the topic is likely to be best addressed by. This decision alongside brief rationale/s is published on the NICE website.

There are updates that do not require assessment by the prioritisation board, these are usually:

- Related to a safety alert that NICE must respond to (for example an MHRA safety alert).
- An alignment of guidance of content already approved by the prioritisation board (for example an update to an incorporated technology appraisal recommendation/s).

NICE guidelines are a key source for the development of NICE quality standards. Therefore, any new guidelines we develop are usually agreed with the relevant body (NHS England or the Department of Health and Social Care).

NICE will routinely identify and contact relevant individuals and organisations at planned points in the topic prioritisation process. Additionally, there is a proforma for topic suggestions (topics@nice.org.uk).

1.5 Who is involved

The committee

The committee is the independent advisory group that considers the evidence and develops the recommendations, taking into account the views of stakeholders (see the chapter on decision-making committees). Committee members include practitioners (both specialists in the topic and generalists), service or care providers or commissioners, and others working in the area covered by the guideline for example researchers and academics. In addition, at least 2 members of every committee are people using services, their family members or carers, or members of the public and community or voluntary sector with relevant experience (lay members).

If needed for a topic, the committee can co-opt members with specific expertise to contribute to developing some of the recommendations.

Registered stakeholders

Registered stakeholders are organisations that have registered with us because they have an interest in the guideline topic, or they represent people whose practice or care may be directly affected by the guideline. They play an important role in developing and advocating for, or implementing, NICE guidelines. During guideline development, we keep registered stakeholders informed of progress by email. We also add information on progress to the guideline page on the NICE website. The plan for the development of guideline recommendations and details of the committee are also available on the NICE website.

Stakeholder organisations are encouraged to get involved in guideline development in a range of ways. The NICE website explains how to register as a stakeholder and how to contribute to the development of a guideline. Registered stakeholders comment on the draft scope (depending on the process followed) and draft guideline recommendations, and they may be invited to provide evidence during guideline development. We formally respond to comments from registered stakeholders, and these responses are published on the NICE website. We may work with key stakeholders to support implementation of published guideline recommendations.

Stakeholders include:

- national organisations for people who use health and social care services, their families and carers, and the public
- local Healthwatch organisations and local organisations that have no national group to participate on their behalf
- national organisations that represent health and social care practitioners and other relevant professionals whose practice may be affected by the guideline, or who can influence uptake of the guideline recommendations
- public sector providers and commissioners of care or services (for example, GP practices)
- private, voluntary sector and other independent providers of care or services
- groups who produce evidence reviews and guidelines
- companies that manufacture medicines, devices, equipment or adaptations, and commercial industries relevant to public health (excluding the tobacco industry)
- organisations that fund or carry out research
- government departments and national statutory agencies
- overseas agencies with a remit covering England.

Individuals cannot register as stakeholders but we encourage anyone with an interest in the topic to express their views to a registered stakeholder listed on the guideline page on the NICE website. Although we will consider comments on the draft scope and guideline from individuals, we do not have the resources to write responses to these comments.

Local or regional professional or practitioner groups, and local or regional groups for people who use health and social care services cannot register as stakeholders unless there is no national organisation that represents the group's specific interests.

Although we are established as an England-only body, our guideline recommendations are used in other countries in the UK. We want our recommendations to be useful in these countries, so encourage individuals and stakeholders from anywhere in the UK to take part in developing them.

Tobacco companies and those who speak for them or are funded by them (collectively referred to as 'tobacco organisations') cannot register as stakeholders and are simply

referred to as 'respondents'.

NICE staff and contractors who work with the committee

The committees are assisted by teams whose work covers evidence review and support, health economic evaluation, guideline recommendation development and quality assurance.

These teams are represented at committee meetings and contribute to discussions. They are not committee members, do not contribute to the quorum of the committee or the development of recommendations during meetings, and do not hold voting rights.

The **centre director** is responsible for ensuring that the guideline is produced in accordance with this manual. The centre director is also responsible for appointing the committee chair and committee members.

The **topic lead** is responsible for the development and quality of the guideline (including the scope), and has delegated responsibility for approving the consultation draft, the final guideline, and other documents, before final approval by NICE's guidance executive. The topic lead also advises the chair of the committee and the developer on matters of process.

The **clinical, medicines, nursing, public health and social care advisers** are responsible for providing advice during all stages of guideline development. Additional specialist adviser roles may be appointed by NICE as needed to support guideline development.

The **technical adviser** is responsible for the technical quality assurance of the evidence reviews and other work undertaken by the technical analyst. The technical adviser commissions, coordinates and quality assures any fieldwork.

The **economics lead** is responsible for ensuring the technical quality of the economic evidence and any economic analysis.

The **information specialist** identifies relevant literature to answer the review questions (see the chapter on identifying the evidence: literature searching and evidence submission), creates databases to manage the search results and keeps a log of search results and strategies.

The **technical analyst** critically appraises the evidence, performs additional analysis, distils it into evidence tables and writes brief summaries (including GRADE tables, GRADE-CERQual or evidence statements, if used) for presentation to the committee (see the chapter on reviewing evidence). The analyst also summarises the evidence and quality of the evidence for the committee and contributes to their discussions.

For most guidelines, an **economist** identifies potential economic issues in discussion with the committee, summarises the published economic evidence and performs additional economic analyses as needed (see the chapter on incorporating economic evaluation).

Administrators, **coordinators** and **project managers** provide administrative and management support to the committee, planning and scheduling the work, arranging meetings, liaising with stakeholders and all individuals and organisations contributing to the development of guidelines.

Staff not involved in developing the guideline carry out independent quality assurance of the guideline to ensure that processes have been followed appropriately, and that the methods are clear and transparent. This includes ensuring that the reviews of the evidence and any economic analysis are up-to-date, credible, robust and relevant. They also check that there is a valid link between the evidence and the recommendations.

Quality assurance takes place throughout development and during checks of the guideline recommendations after publication (surveillance). Independent quality assurance takes place in line with methods and processes set out in this manual, which are informed by the AGREE criteria. We do not routinely invite peer review but may occasionally use external expert peer review (see the section on external expert peer review in the chapter on the validation process for draft guidelines, and dealing with stakeholder comments).

Quality assurance of guidelines and recommendations includes the following actions:

- Ensuring that the scope and guideline are produced in accordance with this manual.
- Approving the consultation draft and the final guideline, before approval by NICE's guidance executive.
- Advising the committee chair and the development team on matters of method and process.
- Providing advice on the guideline topic during all stages of guideline development.

- Ensuring the technical quality of the non-economic and economic evidence reviews, and any related analyses.
- Providing advice on any medicines and prescribing aspects of the guideline.
- Ensuring that the guideline is clear and accessible to readers, and in line with NICE's style guide.

Support

Staff from other NICE teams may work on the guideline recommendations at different stages. They may attend committee meetings and comment on the guideline during consultation and at other times.

Media relations team

The media relations team supports committee members, the development team, and staff with responsibility for quality assurance, on all aspects of communications, including contacts with the media and managing any issues, throughout guideline development and after publication.

Resource impact assessment team

The resource impact assessment team works with the committee, and staff with responsibility for quality assurance, to provide information on the resource impact (costs and savings) of recommendations. Final cost estimates are available to support the implementation of the guideline.

National implementation team

The national implementation team provide support to enable the effective use of NICE guidance. They work in close partnership with key national partners and with health and care practitioners and commissioners to consider barriers to implementation and to develop implementation support resources and solutions to help overcome them.

Insight team

The insight team gathers and analyses data and intelligence to improve understanding of the uptake of our guidance, ensuring we have actionable insights to inform our work.

System implementation team

The system implementation team works with regional and local organisations to promote the guideline and help to put it into practice.

People and communities team

The people and communities team advises on ways to effectively involve people using health and care services, carers and the public, and supports their participation in our work. The team advertises for people with relevant lived experience to join committees and supports them in their roles as committee members. The team also works with organisations for people using services, carers and the public to support their engagement with us, including encouraging organisations to register as stakeholders and contribute to guideline activities.

Medicines optimisation team

The medicines optimisation team provides advice on medicines and prescribing issues in guidelines, including medicines safety, licensing and medicines optimisation.

Publishing team

Content designers from the publishing team work with the committee, the development team and staff with responsibility for quality assurance. They ensure that the guideline and related products are written and presented in a way that is clear to a range of audiences and meets accessibility standards (see NICE's webpage on accessibility).

1.6 Publication and implementation of the guideline recommendations

Guideline recommendations are published on the NICE website alongside the rationales for the recommendations, evidence review documents that include summaries of the committee discussions, and methods. This content may also be syndicated to other organisations to incorporate into their products, for example, decision support tools. Any resources to help users implement the guideline recommendations are also published.

Updating this manual

When significant changes are made, there will be a stakeholder consultation. The updated manual will then be published, along with a list of changes from the previous version of the manual. All registered stakeholders will be invited to comment on proposed changes.

We welcome comments on the content of this manual and suggested subjects for inclusion in the next update. These should be addressed to GuidelinesManualUpdate2022@nice.org.uk.

New methods may be piloted before formal consultation, to fully assess the implications before recommending changes. For small changes to be put in place without stakeholder consultation, they must fulfil all of the following criteria:

- no fundamental stage in the process is added or removed
- no fundamental method, technique or step is either added or removed
- no stakeholders will be obviously disadvantaged
- the efficiency, clarity or fairness of the process or methodology will be improved.

Changes that meet all of these criteria will be published on the NICE website. The manual will be updated, and changes from the previous version of the manual will be listed.

2 The scope

The scope sets out what a NICE guideline will and will not cover. Preparing the scope is the first step in developing a guideline. The scope is used to create a framework for the development work (see the chapters on developing review questions and planning the evidence review, identifying the evidence: literature searching and evidence submission, reviewing evidence, and incorporating economic evaluation).

2.1 Purpose of the scope

The scope of a guideline sets boundaries that ensure the work stays within the referral and the resulting guideline can support any relevant quality standard (see the section on choice of guideline topics in the introduction chapter). The scope for each guideline:

- defines the population(s) and setting(s) that will and will not be covered
- describes what the guideline will and will not consider
- identifies the activities, services or aspects of care that will be considered and lists the draft questions
- describes the economic perspective(s) to be used
- identifies potential equality and health inequality issues among groups sharing protected characteristics and how these will be considered.

Occasionally, it may be necessary when preparing the scope to seek clarification on the referral (for example, to clarify how the NICE guideline will add value in relation to existing non-NICE guidance or to specify the boundaries and the extent of the work) from the commissioning body (see the section on choice of guideline topics in the introduction chapter).

When we identify that recommendations need updating, the guideline scope is updated to reflect this. It lists:

- the sections that will be updated and
- any changes from the current guideline.

2.2 Who is involved in developing the scope

The draft scope is prepared by the development team.

Topic-specific expertise may be provided by members of the committee. Independent quality assurance is done throughout the scoping process by staff with responsibility for quality assurance.

When several related guidelines or guidelines within a topic suite are being developed simultaneously, cross-representation of expertise on each scoping group may also be considered.

The draft scope is signed off by a senior member of staff with responsibility for quality assurance.

2.3 Scope development process

The scope is developed in stages. Stages 1 to 6 apply to new guidelines and stage 7 is for updates to the guideline:

- stage 1: the scoping search
- stage 2: identifying the population, settings and activities, services or aspects of care
- stage 3: identifying and making decisions on overlaps with other NICE guidance
- stage 4: checking the population and selected activities, services or aspects of care with stakeholders
- stage 5: consulting on the draft scope
- stage 6: finalising the scope after consultation
- stage 7: updates to the guideline.

Stage 1: the scoping search

To support scope development a scoping search is done (see the chapter on identifying the evidence: literature searching and evidence submission). The search should not aim to

be exhaustive. It should be based on the need to inform the development of the draft scope and the issues to be discussed at a scoping workshop, if held (see stage 4: checking the population and selected activities, services or aspects of care with stakeholders).

Stage 2: identifying the population, settings and activities, services or aspects of care

Stage 2 includes identifying the population and settings, then considering the activities, services or aspects of care, including main outcomes, for inclusion in the scope. These may have emerged during preliminary work, or may be identified by the scoping search, by considering any health inequalities and impacts on equality, or by consulting experts.

Guidelines do not usually include activities, services or aspects of care that are covered by other arms-length or government bodies such as the Department of Health and Social Care, NHS England, UK Health Security Agency and the Office for Health Improvement and Disparities. For details on areas where we do not usually make recommendations see table 1 in the chapter on interpreting the evidence and writing the guideline.

Equality and equity issues at the scoping stage

During development of the scope, it is important to consider and assess any health inequalities and health inequities to establish:

- whether there is any risk of unlawful discrimination arising from the guideline
- whether the guideline offers any opportunities for advancing equality or reducing inequalities and health inequalities
- whether there might need to be reasonable adjustments to a recommendation to avoid putting any group of people covered by the scope at a substantial disadvantage
- whether, and to what extent, particular equality issues should be included in the scope.

Stage 3: identifying and making decisions on overlaps with other NICE guidance

Identifying related NICE guidance (both published and in development) is a key element of scoping. This helps to see where and how the guideline recommendations are likely to relate to existing recommendations in other guidance.

This process should aim to identify any gaps where new recommendations would be of value, and areas where recommendations already exist (see the chapter on linking to other guidance).

Stage 4: checking the population and selected activities, services or aspects of care with stakeholders

NICE values the views of stakeholders on whether the population group(s) and activities, services or aspects of care identified are relevant and appropriate. Stakeholders include organisations representing health and care professionals such as Royal Colleges and professional associations, organisations led by people using services, and organisations that represent the interests of people with the condition or people using services and their family members or carers, or the public.

For some guidelines, registered stakeholders (see the section on who is involved in the introduction chapter) may be invited to a scoping workshop to talk about the activities, services or aspects of care in the scope, and discuss any other issues as needed. A workshop may be held if the referral is in a new area, there is a new audience for NICE guidelines, or a guideline topic or an area of practice has unique complexities. Following discussions with the development team, staff with responsibility for quality assurance decide whether, and when, to hold a scoping workshop, and document the reasons for the decision.

If a scoping workshop has been held, the development team (with input from other teams) considers the issues raised and refines the scope after the workshop.

Equality and health inequalities assessment

Before the draft scope is signed off for consultation, an equality and health inequalities assessment is completed by the development team and the committee chair to show which equality issues have been identified and considered during scoping, and to provide

assurance that risks of adverse impacts on equality of any exclusions from the scope have been assessed and can be justified. The equality and health inequalities assessment is signed off by a member of staff with responsibly for quality assurance, and published on the NICE website with the draft scope. The assessment is updated by the development team and the committee chair after the scope consultation.

Stage 5: consulting on the draft scope

For new guidelines or for updates where consultation is required, the draft scope is signed off for consultation by a senior member of staff with responsibility for quality assurance. It is published on the NICE website for a 2- to 4-week consultation, and registered stakeholders are notified. Information and prompts to support stakeholder input are posted with the draft scope. The purpose of these prompts is to seek their views on activities, services or aspects of care (such as whether the identified outcome measures are in line with what matters to people with the condition or people using services) and to ask what should be included or excluded. We ask stakeholders to suggest areas where cost savings could be achieved.

Comments are invited from registered stakeholders. The NICE website explains how to register as a stakeholder and how to contribute to the development of a guideline. In particular circumstances, comments will also be requested from the relevant regulatory organisation; for example, the Medicines and Healthcare products Regulatory Agency (MHRA) when the off-label use of medicines is likely to be considered within the guideline or when advice is required on regulations related to medicines.

Registered stakeholders comment on the draft scope (and later on the draft guideline; see the section on what happens during consultation in the chapter on the validation process for draft guidelines, and dealing with stakeholder comments). Comments should be constructed as reasoned arguments and be submitted for the purpose of improving the draft scope. We reserve the right not to respond to comments that are hostile or inappropriate. Accepting late comments and responding to them is at the discretion of the development team.

Tobacco companies and those who speak for them or are funded by them (collectively referred to as 'tobacco organisations') cannot register as stakeholders. Tobacco organisations are simply referred to as 'respondents'. Any comments received during consultation from respondents are reviewed for factual inaccuracy claims and are made public along with any responses.

The development team, staff responsible for quality assurance and the people and communities team (see the section on who is involved in the introduction chapter) routinely review the list of registered stakeholders to check whether any key organisations are missing. Registered stakeholders are also encouraged to identify potential stakeholders who are not registered. When the guideline covers social care, staff with responsibility for social care should be asked about appropriate stakeholders.

Stage 6: finalising the scope after consultation

Dealing with stakeholder comments

After consultation, the development team finalises the scope based on the comments received ensuring that the scope stays in line with the referral for the guideline.

Sometimes registered stakeholders ask for the scope of a guideline to be broadened. If the development team considers that a request to expand the scope would mean the guideline could not be completed on schedule, this should be discussed with staff with responsibility for quality assurance. Sometimes lower-priority areas are removed from the scope to keep the development work manageable. This is done in collaboration with the lead for any related quality standards. Suggestions that are clearly outside the original referral should not be included.

All comments from registered stakeholders, and the actions taken by the development team in response to each comment, are clearly documented by the development team in a 'scope consultation table'. The process for responding to comments from registered stakeholders should follow the principles described in the section on principles of responding to stakeholder comments in the chapter on the validation process for draft guidelines, and dealing with stakeholder comments. Comments received from unregistered stakeholders and individuals are reviewed by the development team and staff with responsibly for quality assurance. We do not formally respond to these comments and do not publish them.

We reserve the right to summarise and edit comments received during consultations, or not to publish them at all, if we consider the comments are too long, or publication would be unlawful or otherwise inappropriate.

Signing off the final scope

The final scope is signed off by a senior member of staff with responsibility for quality assurance. Once the final scope has been published no changes should be made to it unless the guideline is being updated or there are exceptional circumstances (see the section on amending the final scope after publication on the NICE website).

The final scope, the scope consultation table with comments from registered stakeholders and responses to these comments, and the equality and health inequalities assessment are published on the NICE website.

Stage 7: updates to the guideline

When we identify that recommendations need updating, we add to the guideline scope to reflect this. It lists:

- the sections that will be updated and
- any changes from the current guideline.

Small updates to the guideline scope are not usually subject to consultation with stakeholders. For larger updates, or where areas are added to the guideline, consultation with stakeholders may take place. The decision to consult with stakeholders is made by staff with responsibility for quality assurance. For all updates, the scope is published on the NICE website and stakeholders are informed.

2.4 Amending the final scope after publication on the NICE website

There can be exceptional circumstances when the final scope may need amending after it has been signed off and published on the NICE website. For example, amendments may be needed in the light of policy changes, the withdrawal of a medicine, or to include a NICE technology appraisal in development (see the section on related NICE technology appraisal guidance in the chapter on linking to other guidance). The decision on whether to amend the scope is made by a senior member of staff with responsibility for quality assurance, based on advice from the committee or development team as appropriate.

If a final scope is amended after publication, registered stakeholders are informed and the

revised scope is published on the NICE website. Further consultation on the scope would not usually be done.

2.5 References and further reading

Kelly MP, Stewart E, Morgan A et al. (2009) A conceptual framework for public health: NICE's emerging approach. Public Health 123: e14–20

Kelly MP, Morgan A, Ellis S et al. (2010) Evidence-based public health: a review of the experience of the National Institute of Health and Clinical Excellence (NICE) of developing public health guidance in England. Social Science and Medicine 71: 1056–62

Kirkham JJ, Gorst S, Altman DG et al. (2016) Core Outcome Set–STAndards for Reporting: The COS-STAR Statement. PLoS Medicine 13: e1002148

Kirkham JJ, Davis K, Altman DG et al. (2017) Core Outcome Set-STAndards for Development: The COS-STAD Recommendations. PLoS Medicine 14: e1002447

Pawson R (2006) Evidence-based policy: a realist perspective. London: Sage

3 Decision-making committees

3.1 Introduction

This chapter describes the different types of decision-making committees, the training of committee members, the general principles of committee meetings and how committees make group decisions.

A decision-making committee (either a standing committee or a topic-specific committee) draws on its expertise to develop recommendations in the areas defined by the scope of the guideline. Convening an effective committee is therefore one of the most important stages in producing a NICE guideline. The committee:

- may refine and agree the review questions to be addressed by the evidence reviews (for example, when topic-specific input is needed to further define outcomes or specify appropriate comparators) as defined in the scope
- advises on developing the review protocol and alternative analyses
- considers the evidence
- develops the recommendations for practice and research
- considers the likely costs and savings associated with implementing the recommendations
- considers factors that may help or hinder implementation ('levers and barriers')
- advises on implementation support that may be needed.

Therefore the committee needs to be multidisciplinary and include:

- practitioners, professionals, providers, commissioners and researchers (specialists and generalists from the public, private or voluntary sectors, from other independent providers of care and support, or from services)
- lay members (people using services, family members and carers, and members of the public and community or voluntary sector with relevant experience).

Committee members are selected for their knowledge and experience, and do not represent their organisation(s). Lay members champion the perspectives of people who use services, carers or the public. They do not have a healthcare, public health or social care professional or practitioner background in the topic. Lay members should be willing to reflect the experiences of a wide range of people affected by the guideline rather than basing their views solely on personal experience. All committee members are independent of NICE.

In most cases, organisations with a direct commercial interest in interventions or services are not represented on the committee because of potential conflicts of interest, but they contribute to guideline development as registered stakeholders. However, there may be situations in which members of such organisations are included to ensure that this perspective is represented. For example, when guidelines are likely to cover systems and processes relevant to the pharmaceutical or medical devices industries, the committee may include members of the Association of the British Pharmaceutical Industry (ABPI). Any conflicts of interests should be managed in accordance with NICE's policy on declaring and managing interests for NICE advisory committees.

The exact composition of the committee is tailored to the guideline topic and is agreed by the developer and NICE staff with a role in guideline quality assurance. Developers should ensure that all committees can comprehensively consider mental health aspects of guideline topics. Members with expertise in service transformation, change management or implementation science are recruited if needed to support the committee in considering implementation issues throughout guideline development.

For some guideline topics, it may be important for the committee to include other types of expert (for example, an epidemiologist, researcher, statistician or economist with specialist knowledge, or a health inequalities champion). Members with experience of integrating delivery of services across service areas may also be recruited, particularly where the development of a guideline requires more flexibility than conventional organisational boundaries permit. If the guideline contains recommendations about services, committee members with a commissioning or provider background are needed, in addition to members from relevant professional or practitioner networks or local authorities.

When several related guidelines are being developed at the same time by different committees, some committee members may sit on more than one committee, depending on expertise.

As far as possible, the committee should aim for diversity in membership, an objective of NICE's equality scheme. Equality and anti-discrimination considerations are reflected at every stage of recruitment and anonymised data from recruitment is submitted to NICE's corporate office who use this to report to the NICE Board.

Ideally, committee members have experience of the UK healthcare setting and are drawn from different parts of England. But this depends on the expertise available and does not exclude people from other countries in the UK.

All committee members, including practitioner, provider, commissioner and lay members, have equal status, acknowledging the importance of the expertise and experience that each member brings to the committee.

The committee may also be supported by co-opted members who are invited to contribute to formulating recommendations in a specific part of the guideline only. They take part fully in discussions, but do not have voting rights or count towards quorum.

Expert witnesses may also be invited to some committee meetings to provide additional evidence (see the appendix on call for evidence and expert witnesses).

3.2 Forming the committee

The committee can be formed in 2 ways:

- from members of a NICE standing committee, with additional recruitment of topic expert members
- from multidisciplinary recruitment of all committee members (topic-specific committee).

The resulting committee should, as far as practically possible, reflect the range of stakeholders and groups whose activities, services or care will be covered by the guideline. The type of committee chosen is likely to reflect the nature of the work.

For all guidelines covering children and young people and those focusing on people with a learning disability or cognitive impairment, the developer should consider how to involve people affected by the guideline when they begin work on the topic, and should reflect this in its plans. NICE's people and communities team can advise on options and methods

for involving people affected by the guideline. They can target adverts for committee members to specific groups and signpost to sources of specialist advice (see the appendix on approaches to additional consultation and commissioned primary research and the page on the NICE website on involving children and young people in NICE's patient and public involvement policy). They will work with the developer on tailored support where this is needed.

For some topics it may be possible to recruit young people aged 16 to 18 years or people with a learning disability as members of the committee, or as co-opted members or topic experts. With specialist support and reasonable adjustments, a developer has successfully included people with a learning disability as committee members for NICE guidelines on improving care for this population. This was achieved through specialist support and adjustments to the process to enable their full and equal involvement.

When there are particular reasons why people affected by the guideline cannot be recruited as lay committee members (for example, when the guideline covers children), other approaches are needed to ensure that their views and experiences are incorporated in the recommendations. These might include working with an external agency to obtain user views or consulting with a reference group of people using services, at key stages of guideline development. For example, for the NICE guideline on child abuse and neglect, the developer commissioned the independent charity Against Violence and Abuse (AVA) to recruit and facilitate an expert reference group of young people as a subgroup of the committee (see the appendix on Children and Young People's Expert Reference Group summary report). Another approach could involve a targeted consultation with people affected by the guideline to fine-tune selected draft recommendations (see the section on what happens during consultation in the chapter on the validation process for draft guidelines, and dealing with stakeholder comments and the appendix on approaches to additional consultation and commissioned primary research).

If other approaches to engaging people affected by the guideline are needed, the developer should document the reasons, together with a proposal for the work, including possible methods to be used, and the anticipated costs and time. The proposal should be discussed and agreed with members of NICE staff with a quality assurance role and the people and communities team, and approved by the centre director. When the work is approved, the reasons for the approach and methods should be documented in the guideline.

All committee members are recruited in accordance with NICE's appointment to advisory

bodies policy and procedure. Positions are advertised on the NICE website and other appropriate places (for example, NICE Twitter, social media and websites of stakeholders, Royal Colleges and professional organisations), and relevant stakeholders are notified. Candidates are required to submit a declaration of interest, curriculum vitae (CV) and covering letter, or application form in the case of lay members. They are encouraged to complete an equality monitoring form.

The committee chair

The chair guides the committee in terms of task (developing the guideline, focusing on any referral, the scope and timescale) and process (how the group works). The chair helps the committee to work collaboratively, ensures a balanced contribution from all committee members, and is mindful that some members may need support to ensure full participation.

The chair is appointed for their expertise and skill in chairing groups, and although they may have some knowledge of the topic, this is not their primary role in the group. Specialist knowledge is provided by other committee members, including in many cases a topic adviser (see the section on topic-specific committees).

The chair ensures adherence to NICE's equality scheme and that the committee takes account of NICE's principles. The chair and a senior member of the developer's team consider any potential conflicts of interest of committee members. The chair also ensures that the guideline recommendations reflect the evidence and the committee's considerations. The chair should ideally be appointed before guideline scoping and may contribute to early development of the scope. NICE has some chairs who may be appointed to chair more than 1 topic-specific committee (see the section on topic-specific committees). More details on the role of the chair can be found in the appendix on guideline committee Terms of Reference and Standing Orders.

3.3 Standing committees

NICE has multiple standing committees in operation at any one time. Some guideline topics are allocated to a standing committee before scoping.

Standing committees usually include between 12 and 18 members (both practitioner and lay members). The size of the committee depends on the complexity and breadth of the guideline. Some members are generalists (core members) and some have specialist

expertise (topic expert members). Each standing committee consists of:

- a chair
- core members
- topic expert members.

Additional members with specialist knowledge may also be co-opted to the committee for 1 or more meetings to contribute to formulating recommendations in a specific part of the guideline.

More details on the role of committee members can be found in the appendix on guideline committee Terms of Reference and Standing Orders.

Core members of standing committees

The number of core members of a standing committee depends on the complexity and breadth of the guideline programme, but is usually between 6 and 12. This number allows members to contribute effectively to discussions while including a broad range of experience and knowledge. Core members should include at least 1 practitioner member.

Some core members will have experience of commissioning or implementing interventions, services or care at regional and local levels. Others will have specific expertise in assessing the quality of the evidence presented to the committee, and in its interpretation.

Core members should include at least 1 lay member. Lay members help ensure that the committee's recommendations are relevant to specific groups or to the general public. They also help to identify where the recommendations should acknowledge general or specific preferences and choice by people using services, family members and carers, or members of the wider public.

Core members of a standing committee are appointed to a single committee for a 3-year term in the first instance. All members are eligible for re-appointment after 3 years.

Topic expert members of standing committees

When a new guideline is allocated to a standing committee, the core members of the committee are complemented by topic expert members. They have specialist knowledge

of the topic and may include providers, commissioners and practitioners, and should include at least 1 lay member. The lay member either has direct experience of the topic or is a member of a relevant organisation or support group. The number of topic expert members varies but should be no more than half of the total number of committee members.

Topic expert members are usually recruited for a specific guideline, but may be appointed for up to 3 years, at the discretion of NICE, so that they can work on subsequent related guidelines. This might mean they move between standing committees during their term, depending on the guidelines being produced. All members are eligible for reappointment after 3 years.

The process of appointing topic expert members is completed at least 6 weeks before the first committee meeting for the guideline and takes into account the final scope of the guideline. Topic expert members are full members of the committee, with voting rights. They join in discussions, contribute to formulating recommendations and count towards the quorum.

3.4 Topic-specific committees

The chair and members of a topic-specific committee are appointed for the development of a particular guideline or for up to 3 years to work on multiple guidelines within a broad topic area, with membership subject to renewal for a total period of up to 10 years. Chairs may also be appointed to chair more than 1 topic-specific committee. The chair, the topic adviser (if there is one) and possibly 1 or 2 other committee members are likely to be appointed before guideline scoping and contribute to the development of the scope (see the chapter on the scope).

The final composition of a topic-specific committee is agreed by the developer and NICE staff with a role in guideline quality assurance and takes into account the final scope for the guideline. The committee usually comprises between 13 and 15 members. This number allows members to contribute effectively to discussions while including a broad range of experience and knowledge. Occasionally when the topic is very broad, a larger committee may be convened. A topic-specific committee is made up of:

- a chair
- a topic adviser (not all topic-specific committees have topic advisers)

- practitioner and professional members, providers and commissioners
- at least 2 lay members (people using services, family members and carers, and members of the public, community or voluntary sector with relevant experience); examples of topics where there may be more than 2 lay members include topics where personalisation and choice are particularly important, where a scope is large and covers different areas, and topics covering all age groups.

Additional members with specialist knowledge may also be co-opted to the committee for 1 or more meetings to contribute to formulating recommendations in a specific part of the guideline.

The topic adviser of a topic-specific committee

A topic adviser with specialist knowledge may be appointed to a topic-specific committee (for example, when the committee chair does not have topic expertise). The topic adviser is a member of the committee but also supports the developer. The topic adviser contributes to the development of the scope (see the chapter on the scope) and is therefore appointed before scoping work starts.

The topic adviser's exact responsibilities depend on the guideline and the expert input needed. Responsibilities may include working with the technical analyst on the evidence reviews (if topic-specific knowledge is needed), or checking the guideline to ensure that the terminology and language are correct.

Practitioner and professional members of topic-specific committees

Practitioner and professional members of a topic-specific committee may be recruited before the scope is finalised (see the section on stages of scope development in the chapter on the scope). They should reflect the views and experiences of practitioners, professionals, providers and commissioners working in the area covered by the guideline.

Practitioner and professional committee members have appropriate knowledge and skills; detailed research expertise is not necessary, although an understanding of evidence-based practice is essential and some experience of service transformation or delivering integrated services across boundaries is desirable. Practitioner and professional members contribute their own views to the committee and do not represent the views of their

professional organisations.

A topic-specific committee usually includes between 9 and 11 practitioner or professional members (occasionally when the topic is very broad, more members may be recruited). The spread of interest and experience of practitioner and professional members is agreed between the developer and members of NICE staff with responsibility for guideline quality assurance.

Lay members of topic-specific committees

All committees have at least 2 lay members with experience or knowledge of issues that are important to people using services, family members and carers, and the community affected by the guideline. This helps to ensure that the guideline is relevant to people affected by the recommendations and acknowledges general or specific preferences and choice.

Lay members usually have direct experience of the condition, services or topic being covered by the guideline – as a patient, person using services, carer or family member, or as a member or an officer of a lay stakeholder organisation or support group. However, they do not represent the views of any particular organisation.

When appropriate, lay members may be recruited before the scope is finalised (see the section on stages of scope development in the chapter on the scope).

3.5 Other attendees at committee meetings

Expert witnesses

If the committee does not have sufficient evidence to make recommendations in a particular area (for example, if there are gaps in the evidence base or subgroups are under-represented), it may call on external experts (expert witnesses) who can provide additional evidence from their experience and specific expertise, to help the committee to consider and interpret the evidence.

Once the committee has established that it needs evidence in a particular area from an expert witness, committee members and NICE's people and communities team are asked by the developer to nominate experts who might fulfil this role. Expert witnesses may also

be identified by the developer or NICE staff with a quality assurance role, or if required, by active recruitment through stakeholder organisations and the NICE website. Experts may be drawn from a wide range of areas as appropriate, including government and policy, research, practice, people using services and carers or the community and voluntary sector.

Before seeking expert testimony from children or other vulnerable groups, expert advice should be sought about the ethics and implications of the involvement. Adjustments and additional support may be needed, such as giving testimony via video recording, or in private session. There is no minimum age for young people providing expert testimony. If a child or young person attends a committee meeting, they must be accompanied by their parent, carer or other appropriate adult with responsibility for their welfare (see appendix 1 on children, young people and safeguarding in NICE's patient and public involvement policy).

Expert witnesses attending a committee meeting are invited to answer questions from members of the committee. They may be invited to present their evidence at a committee meeting in the form of expert testimony based on a written paper. The written expert testimony paper may be shared with the committee before the meeting or the paper may be submitted by the expert after the meeting. Sometimes the developer writes up the expert testimony and agrees this with the witness after the meeting. Expert testimony papers are posted on the NICE website with other sources of evidence when the guideline is published.

Expert witnesses are not members of the committee; they do not have voting rights, and they should not be involved in the final decisions or influence the wording of recommendations.

Committee support roles

The committee is assisted by a range of people, who have a role in:

- quality assurance
- development
- support.

These are technical and project management staff from the developer or NICE staff with a

quality assurance role. Other NICE staff (such as content designers and members of the NICE resource impact assessment team) also attend some meetings. They are not committee members and do not have voting rights at committee meetings.

Public access to meetings

Enabling public access to NICE advisory body meetings is part of NICE's commitment to openness and transparency. It enables stakeholders and the public to better understand how evidence is assessed and interpreted, how consultation comments are taken into account and how recommendations are formulated. Public access to all guideline committee meetings is not possible because of financial pressures. In addition, all or part of a meeting may need to be held in closed session because expert evidence involves the disclosure of a person's health problems, or the consideration of national policy that has not been agreed by ministers, or the drafting of recommendations might affect commercial interests.

Standing committee meetings are usually held in public; topic-specific committee meetings are held in private. Public access to standing committee meetings is arranged according to the NICE policy (see the appendix on guideline committee Terms of Reference and Standing Orders).

To promote public access to standing committee meetings, NICE publishes a notice with a draft agenda, alongside details of how the meeting can be accessed, on its website in advance of the meeting.

Standing committee meetings may be held entirely in public or split into 2 parts: part 1 with the public having access, and part 2 a closed session with no public access. On rare occasions a standing committee meeting may be entirely closed. The decision to hold a closed session is made by the committee chair and the NICE centre director responsible for the guideline.

3.6 Code of conduct and declaration of interests

Declaring interests

All committee members, including the chair, and anyone who has direct input into the guideline (including the developer and expert witnesses) must declare any potential

conflicts of interest in line with NICE's policy on declaring and managing interests for NICE advisory committees. For committee members, including the chair, this happens on application for committee membership. Any relevant interests, or changes to interests, should also be declared publicly at the start of each committee meeting. Before each meeting, any potential conflicts of interest are considered by the committee chair and a senior member of the developer's team. Any decisions to exclude a person from all or part of a meeting should be documented. Any changes to a member's declaration of interests should be recorded in the minutes of the meeting. Declarations of interests are recorded in a register for each guideline and are published on NICE's website.

Code of conduct and confidentiality

NICE has developed a code of conduct for committee members (see the appendix on code of conduct for committee members), which sets out the responsibilities of NICE and the committee, and the principles of transparency and confidentiality.

Everyone who sees documents containing confidential information or who is party to part 2 (closed session) discussions about a guideline before public consultation must sign a confidentiality agreement before becoming involved.

If committee members are asked by external parties – including stakeholders, their professional organisation or the media – to provide information about the work of the committee, they should contact the developer for advice.

Terms of Reference for committees

Details of the role of committee members, and the procedural rules for managing the work of committees, can be found in the appendix on guideline committee Terms of Reference and Standing Orders.

All committee members should be committed to developing NICE guidelines according to NICE's methods and processes, and to working within NICE's equality scheme.

New members, with the exception of co-opted members, should not usually be added to the committee after the first meeting, because this may disturb the group dynamic. However, when a resignation leaves a gap in experience and expertise, recruitment of new members is considered.

3.7 Identifying and meeting training needs of committee members

Induction

All committee members, including topic expert members and co-opted members, receive an induction from NICE or the developer covering:

- key principles for developing NICE guidelines
- the process of developing NICE guidelines, including the importance of being familiar with relevant chapters of this manual
- how the elements of the guideline development process fit together, and the relationship to quality standards and products supporting implementation
- the role of the committee, including Terms of Reference and Standing Orders, and how lay members contribute
- the role of the developer and NICE teams
- formulating review questions
- reviewing evidence
- the basics of how economics methods are used in decision-making
- developing and wording recommendations
- how guidelines are presented on the NICE website
- information about resource impact and how this is considered alongside the economic evidence
- information about implementation
- NICE's principles and equality scheme
- declaration of interests.

The induction may be scheduled on appointment of the member, or during an early committee meeting. To work effectively, committee members may need training and

support in some technical areas of guideline development, such as systematic reviewing and economics. Such sessions are arranged by NICE or the developer, as required. Before beginning their work in a committee, members may also be invited to observe a meeting of another committee.

All committee members are encouraged to provide information to NICE staff about any needs they have for additional support to enable them to participate fully in the work of the committee.

Training for the committee chair

The person selected to perform the role of committee chair may need support and training so that they can carry out their role effectively. The chair needs in-depth knowledge of the NICE guideline development process and an understanding of group processes. Anyone appointed as a committee chair is required to attend the chairs' induction session provided by NICE, which in addition to the above covers the key tasks that the chair is expected to perform, including:

- facilitating meetings so that all committee members are involved
- ensuring that lay members of the committee can contribute to discussions
- ensuring that people requiring adjustments who are members (including topic expert, co-opted or lay members) can contribute to discussions
- declaring interests and dealing with conflicts of interest
- planning and organising the work of the committee, including how the evidence is considered, consensus approaches and developing recommendations.

In addition to the specific induction session, the developer should identify and meet any additional training needs of a committee chair. The developer may consider a 'buddying' approach in which a new committee chair learns from someone with previous experience.

Training for lay members

Lay members of the committee are offered support and training by NICE's people and communities team. This is in addition to the induction and any training they receive alongside other members of the committee, and allows specific questions and needs to be addressed. The training advises lay members about effective participation in guideline

development and gives them the opportunity to learn from people who have had a lay role on previous committees. Public involvement advisers also offer ongoing support tailored to a lay member's needs.

3.8 Committee meetings

General principles

The committee is multidisciplinary and its members bring with them different beliefs, values and experience. All these perspectives are valued by NICE and should be considered. Each member should have an equal opportunity to contribute to the development of the guideline, and should receive any additional support they need to do this. For this reason, it is important for the chair to check that the terminology used is understood by all committee members and is clarified if needed. The chair should ensure that there is sufficient discussion to allow a range of possible approaches to be considered, while keeping the group focused on the guideline scope, the evidence being reviewed, and the timescale of the project.

Meeting documentation

Meeting documentation is usually sent to committee members to arrive at least 5 working days before a committee meeting.

The developer takes formal minutes during committee meetings and these are reviewed and approved at the next meeting. The approved minutes of each meeting are posted on the NICE website during guideline development. The information includes:

- where the meeting took place
- who attended
- apologies for absence
- declarations of interests of those attending, including actions and decisions made about any conflicts of interest
- a list of the subjects discussed
- date, time and venue of the next meeting.

Meeting schedule

The number of committee meetings depends on the size and scope of the topic. There are usually between 2 and 15 meetings for each topic; most are 1-day meetings, but some may take place over 2 days.

Initial meetings

During the initial meeting(s), it may be helpful to establish a framework that clarifies the objectives of the committee, the specific tasks that need to be carried out and the timetable. This allows the group to focus on the task and to develop a working relationship that is structured and well defined.

Initial meetings may be used to consider the background to the guideline, the scope, and review protocols for the evidence reviews and any economic analysis that is needed. Drafts or completed evidence reviews may be included in initial meetings if they are available. At initial meetings of standing committees, topic expert members may be invited to give presentations on their area of work, practice or experience, to familiarise core committee members with key topic issues.

If review questions and protocols are still in development, the developer evidence review team will draft the review protocols and present them to the committee for comment (the chapter on developing review questions and planning the evidence review describes the process of developing review questions). The committee is asked whether the draft review protocols and economic analysis are likely to answer the review questions. Committee members are asked to suggest any amendments or improvements (for example, to further define outcomes or specify appropriate comparators).

During initial meetings, the committee may also be asked to discuss the development plans and to suggest areas that might benefit from expert testimony. The committee may be asked to suggest people who can provide that testimony and discuss and consider evidence. They will also be asked if they are aware of any health inequalities or equity issues.

For some topics, the committee may also be asked to discuss options or plans for involving groups who may not be part of the decision-making process (for example, children and young people or people with a cognitive impairment; see the section on forming the committee).

Development meetings

Evidence reviews and economic analyses are presented to the committee over the course of a defined number of meetings. The committee considers the evidence review for each review question, any economic analyses and any additional evidence (for example, expert testimony, views of people using services from a reference group, information from focus groups or other exceptional consultation activity). It discusses how these answer the review questions and summarises each area of evidence. To facilitate guideline development, the committee may work in smaller subgroups whose proposals are then considered by the whole committee.

The committee also discusses the wording of any draft recommendations (see the chapter on interpreting the evidence and writing the guideline). The discussions and rationales for the recommendations are recorded.

NICE staff (for example, the lead content designer, public involvement adviser and media relations lead for the guideline and members of the resource impact assessment team) may give presentations and/or provide information to explain their roles to the committee. Committee members may be asked to volunteer to work with NICE on the following:

- activities and tools that support implementation of the guideline (see the chapter on support for putting the guideline recommendations into practice)
- promoting the guideline (see the chapter on finalising and publishing the guideline recommendations).
- developing content for the information for the public section on the topic webpage, including the guideline's key messages.

Final meetings

Towards the end of guideline development, the committee discusses and agrees the final wording of the draft guideline for consultation, including the draft recommendations (see the chapter on interpreting the evidence and writing the guideline).

After consultation the committee discusses the comments received during consultation, any changes needed to the guideline, and agrees the final wording of the recommendations (see the chapter on finalising and publishing the guideline recommendations).

Record-keeping

The developer should maintain records throughout guideline development and ensure that record-keeping standards are appropriate for audit. The following information should be readily available if requested by NICE staff with a quality assurance role:

- details of the committee members, including declarations of interest
- details of the search strategies, including when the most recent searches were conducted
- details of the draft recommendations and the rationales for the recommendations
- details of the included and excluded studies and associated review protocols
- data-extraction forms
- draft evidence reviews with evidence tables
- details of the economic analysis, including any working models
- minutes of committee meetings
- any additional information presented to the committee (for example, expert testimony papers, presentations, examples of practice).

3.9 Making group decisions and reaching consensus

Reaching agreement

Committee members need to make collective decisions throughout guideline development. This can include agreeing the review questions (see the chapter on developing review questions and planning the evidence review) and protocols (see the section on planning the evidence review in the chapter on developing review questions and planning the evidence review), interpreting the evidence to answer these questions (see the chapters on reviewing evidence and incorporating economic evaluation) and developing recommendations (see the chapter on interpreting the evidence and writing the guideline).

There are many different approaches to making group decisions, and there are no rules

that set out which approach should be used in which circumstances. Also, because committees work in different ways to reflect their individual membership, it is difficult to be prescriptive about the approach that should be used. In most cases, the committee reaches decisions through a process of informal consensus. In all cases the approach used should be documented.

The role of the committee chair in reaching consensus is to ensure that:

- everyone on the committee, including lay members, can present their views
- assumptions can be debated
- discussions are open and constructive.

The chair needs to allow sufficient time for all committee members to express their views without feeling intimidated or threatened, and should check that all of them agree to endorse any recommendations. If the committee cannot come to consensus in a particular area, the reasons for this should be documented, and the wording of the recommendation reflect any underlying uncertainty (see the chapter on interpreting the evidence and writing the guideline).

Formal consensus within the group

In exceptional circumstances, some committees may choose to use formal voting procedures or formal consensus methods for certain decisions (for example, when members disagree or when there is no evidence; see the appendix on guideline committee Terms of Reference and Standing Orders). NICE does not offer advice on which of the many methods might be used. However, the methods for achieving consensus should be recorded in the minutes of the meeting and a clear statement made about the factors that have been considered. This should also be documented in the guideline methods, ensuring the process is as transparent as possible.

The views of all committee members should be considered, irrespective of the method used to reach consensus.

Formal consensus outside the group

In certain cases, for example, if the literature search has found no evidence that addresses the review question, the committee may wish to identify wider views on best practice by

using formal consensus methods (such as the Delphi technique or the nominal-group technique) outside of the group. The use of these methods and the constituency of the wider group should be discussed on a case-by-case basis with members of NICE staff with responsibility for guideline quality assurance, and the NICE people and communities team lead. The final decision on whether these methods are warranted is made by NICE staff with responsibility for quality assurance. If it is decided that such methods may be used, the planning and methods will be clearly documented and the methods described in the guideline. In all cases the approach used should be documented.

3.10 Further contributions of committee members

Some committee members may be invited to apply to join the Quality Standards Advisory Committee that is developing a quality standard related to the guideline. Some members may also be invited to contribute to a future review of the guideline and a check of the need for an update.

3.11 References and further reading

Choudhry NK, Stelfox HT, Desky AS (2002) Relationships between authors of clinical practice guidelines and the pharmaceutical industry. Journal of the American Medical Association 287: 612–7

Eccles M, Grimshaw J, editors (2000) Clinical guidelines from conception to use. Abingdon: Radcliffe Medical Press

Elwyn G, Greenhalgh T, Macfarlane F (2001) Groups: a guide to small groups. In: Healthcare, management, education and research. Abingdon: Radcliffe Medical Press

Hutchinson A, Baker R (1999) Making use of guidelines in clinical practice. Abingdon: Radcliffe Medical Press

Kelly MP, Moore TA (2012) The judgement process in evidence-based medicine and health technology assessment. Social Theory and Health 10: 1–19

4 Developing review questions and planning the evidence review

At the start of guideline development, the key issues and draft questions listed in the scope should be translated into review questions and review protocols.

Review questions and review protocols must be clear and focused and build on the draft questions listed in the scope. They provide the framework for the design of the literature searches, inform the planning, methods and process of the evidence review, and act as a guide for the development of recommendations by the committee. Review protocols may also be used to inform surveillance of guidelines, and future updates (see the chapters on ensuring that published guidelines are current and accurate and updating guideline recommendations).

This chapter describes how review questions and review protocols are developed and agreed. It describes the different types of review question and provides examples. It also provides information on the different types of evidence and how to plan the evidence review. The best approach may vary depending on the topic. Options should be considered by the development team, and the chosen approach discussed and agreed with staff with responsibility for quality assurance. The approach should be documented in the review protocol (see the appendix on review protocol templates) and the guideline, together with any reasons for the choice, if the approach is non-standard.

4.1 Number of review questions

The number of review questions for each guideline depends on the topic and the breadth of the scope. However, it is important that the total number of review questions:

- provides sufficient focus for the guideline, and covers all key areas outlined in the scope
- focuses on key questions that are likely to generate useful or needed recommendations
- can be completed in the time and with the resources available.

Review questions can vary considerably in terms of both the number of included studies and the complexity of the question and analyses. For example, a single review question might involve a complex comparison of several interventions, many studies that are using different study designs (including systematic reviews), multiple population subgroups, or the intervention itself may be considered complex (see the section on review questions about complex interventions). At the other extreme, a review question might investigate the effects of a single intervention compared with a single comparator and there may be few primary studies or no study meeting the inclusion criteria. The number of review questions for each guideline, and how much time and resources are needed for them, will therefore vary depending on the topic and its complexity, and the nature of the evidence.

4.2 Developing review questions from the scope

The review questions should cover all key areas specified in the scope but should not introduce new areas. They should build on the draft questions in the scope and usually contain more detail. For details on developing the scope, see the chapter on the scope.

Review questions are usually drafted by the development team. They are then refined and agreed with the committee members. This enables the literature search to be planned efficiently. Sometimes the draft questions from the scope need refining before the development of review protocols, or very occasionally after the evidence has been searched. Any such changes to review questions (with reasons) should be agreed with a member of staff with a quality assurance role. All changes should be clearly recorded in the review protocol and evidence review document, so they are auditable.

NICE guidelines should not reiterate or rephrase recommendations from the NICE guidelines on patient experience in adult NHS services, service user experience in adult mental health, people's experience in adult social care services, babies, children and young people's experience of healthcare, shared decision making, or other NICE guidelines on the experience of people using services. However, whether there are specific aspects of views or experiences that need addressing for a topic, should be considered during the scoping of every guideline. Specific aspects identified during scoping should be included in the scope if they are not covered by existing guidelines and are supported as a priority area. These are likely to be topic-specific and should be well-defined and focused.

4.3 Formulating and structuring different review questions

Review questions should be clear and focused. The exact structure of each question depends on what is being asked. The aims of questions will differ, but are likely to cover at least one of the following:

- effectiveness of an intervention or interventions
- accuracy or effectiveness of diagnostic tests or test strategies
- prognosis of outcomes over time, based on the characteristics of the person using services
- predicting an individual prognosis or identifying an individual diagnosis
- views and experiences of people using services, family members or carers, or of those commissioning and providing services
- service delivery
- epidemiology or aetiology of a disease or condition
- equality and health inequalities issues.

The nature and type of review questions determines the type of evidence that is most suitable (Petticrew and Roberts 2003). There are examples of different types of review questions and the type of evidence that might best address them throughout this chapter. When developing review questions, it is important to consider if any additional information is needed for any planned economic modelling. This might include information about quality of life, rates of adverse events in particular populations, and use of health and social care services.

Conceptual frameworks and logic models

When review questions are being developed, it can sometimes be useful to use a conceptual framework or a logic model (or reuse ones that were developed for the scope). These display pathways through which actions and interventions are expected to lead to differences in outcomes. They can be used to potentially identify which interventions are most likely to be effective when targeted at particular places in a pathway. Conceptual

frameworks and logic models can be produced by the development team or taken from previously published literature in the topic area.

Conceptual frameworks or logic models are helpful in many topics and can be particularly useful when:

- the evidence reviews for the guideline will be considering complex interventions
- the context around an area is complex (for example, interactions with policy, multiple bodies involved in delivering an intervention, or where the commissioning or delivery approaches are unclear or complex).

An understanding of the expected mechanisms by which outcomes occur can help guide which interventions are most worthwhile to study in particular settings or questions.

Conceptual frameworks and logic models are used to aid the committee in developing review questions and protocols that are most likely to produce useful results. However, they are not intended to constrain the recommendations the committee may make in the guideline (for example, the committee is not limited to making recommendations that fit within the causal pathways in the framework or model). Examples of conceptual frameworks and logic models are given in figure 1 and figure 2.

Figure 1 Example of a conceptual framework/logic model for promoting mental wellbeing at work

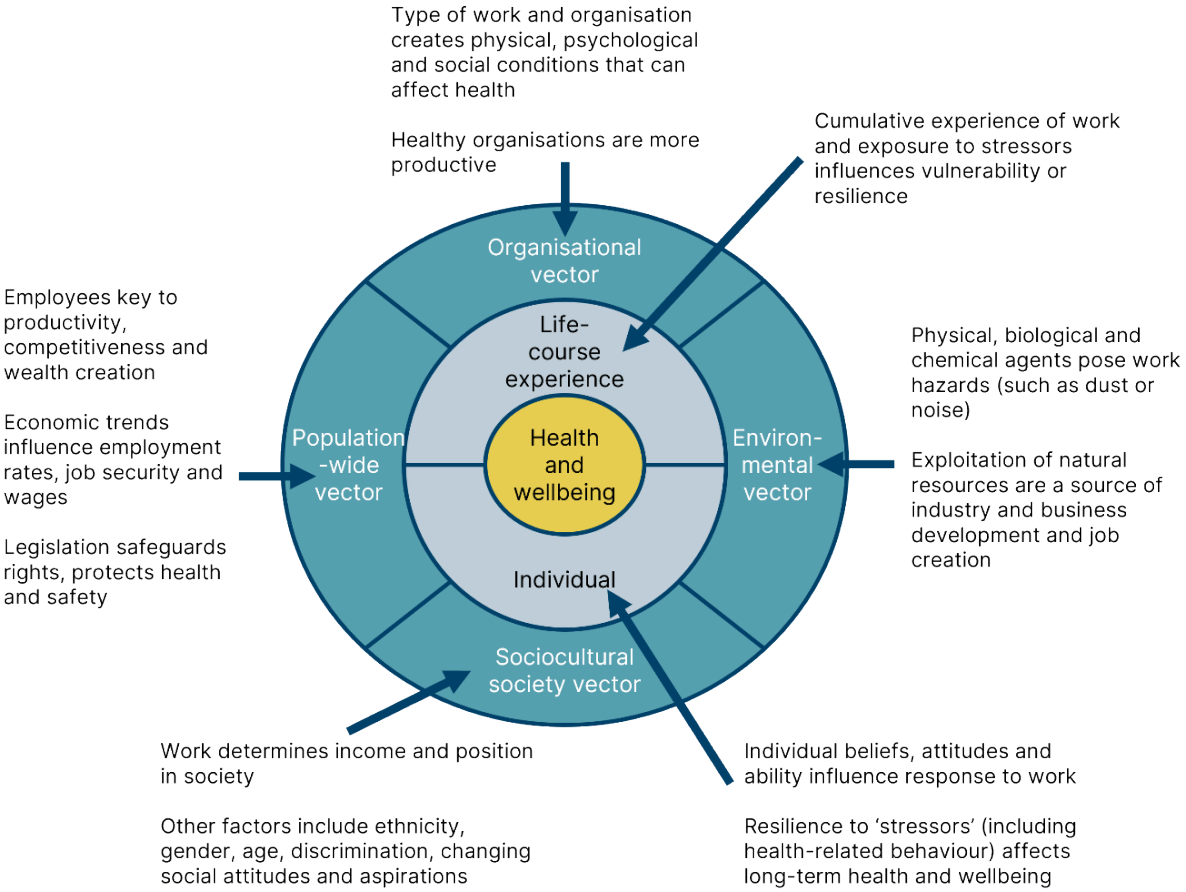
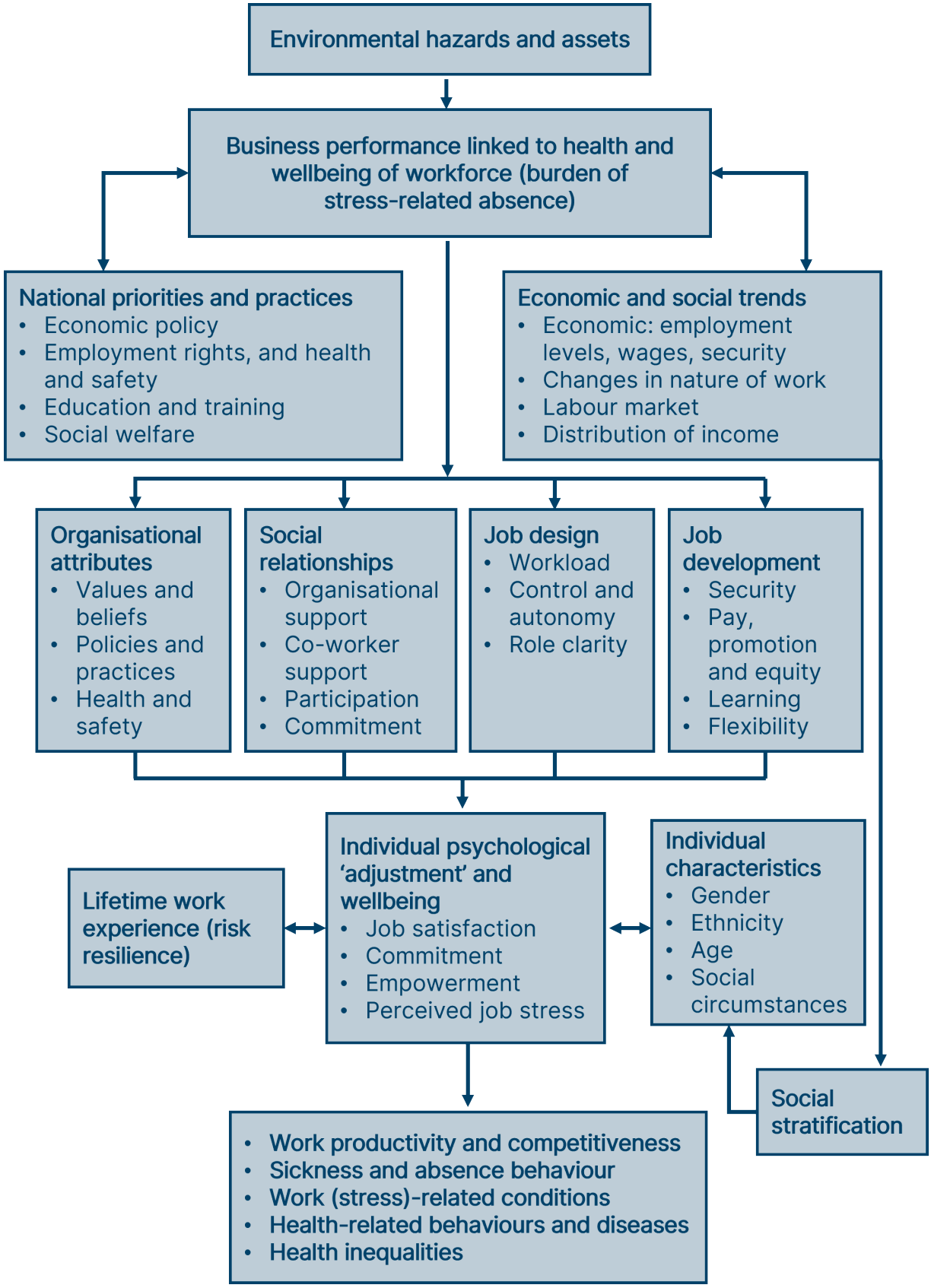


Figure 2 Example of a conceptual framework/logic model for promoting mental wellbeing at work



Review questions about the effectiveness of an intervention

A helpful structured approach for developing questions about interventions is the PICO (population, intervention, comparator and outcome) framework (see box 4.1). The setting for the question should also be specified if relevant.

Box 4.1 Formulating a review question on the effectiveness of an intervention using the PICO framework

Population: Which population are we interested in? How best can it be described? Are there subgroups that need to be considered?

Intervention: Which intervention, treatment or approach should be examined?

Comparators: What are the alternatives to the intervention being examined (for example, other interventions, usual care, placebo)?

Outcome: Which outcomes should be considered to assess how well the intervention is working (including outcomes on both benefits and harms)? What is important for people using services? Core outcome sets should be used if suitable based on quality and validity; one source is the COMET database. The Core Outcome Set Standards for Development (COS STAD) and Core Outcome Set Standards for Reporting (COS STAR) should be used to assess the suitability of identified core outcome sets.

For each review question, factors that may affect the outcomes and effectiveness of an intervention, including any wider social factors that may affect health and any health inequalities, should be considered (see the section on considering health inequalities when preparing review protocols). Outcomes (on both benefits and harms) and other factors that are important should be specified in the review protocol. In general, a range of 5 to 9 outcomes should be defined. Guidance on prioritising outcomes is provided by the GRADE working group.

When designing review questions and protocols, it is important to consider possible intercurrent events (events that occur after starting treatment and either preclude the observation of the variable, or affect its interpretation, such as death, non-adherence to treatment or stopping treatment) and how these will be dealt with in any analysis in the guideline, if different identified studies analyse the data in different ways. Clinical trials are increasingly following the Estimand framework (European Medicines Agency), which attempts to increase the clarity of the precise treatment effect that an individual study is estimating.

Examples of review questions on the effectiveness of interventions are shown in box 4.2.

Box 4.2 Examples of review questions on the effectiveness of interventions

- What are the most effective blood pressure targets for reducing the risk of future cardiovascular events in adults with diagnosed primary hypertension and established cardiovascular disease?
- What approaches are effective in improving access to and/or engagement with health and social care for people experiencing homelessness?
- What pharmacological (antimicrobial and non-antimicrobial) and non-pharmacological interventions are effective in managing acute uncomplicated otitis media?
- Which of the harm-reduction services offered by needle and syringe programmes (including advice and information on safer injecting, onsite vaccination services, and testing for hepatitis B and C and HIV) are effective in reducing blood-borne viruses and other infections among people who inject drugs?
- When escalating from oxygen therapy, which non-invasive modality is most effective in adults in hospital with suspected or confirmed COVID-19?
- What are the most effective combined approaches to identifying, assessing and monitoring the health, social care and education needs (including changing needs) of disabled children and young people with severe complex needs?
- What is the effectiveness of SGLT2 inhibitors for children, young people and adults with chronic kidney disease and type 2 diabetes?
- What is the effectiveness of selective laser trabeculoplasty as a first-line treatment compared with intraocular pressure-lowering eyedrops in adults with ocular hypertension or chronic open-angle glaucoma?
- For children and young people with complex rehabilitation needs after traumatic injury that involves spinal cord injury, what specific rehabilitation programmes and packages are effective?
- What interventions are effective in improving access to diagnosis and treatment services for people with suspected or diagnosed depression, and improve referral from primary to secondary and tertiary levels of care in populations or groups with low uptake in the UK?

Review questions about pharmacological interventions will usually only include medicines with a UK marketing authorisation, based on regulatory assessment of safety and efficacy. Use of a licensed medicine outside the terms of its marketing authorisation (off-label use) may be considered in some circumstances; for example, if this use is common practice in the UK, if there is good evidence for this use, or if there is no other medicine licensed for the indication. Off-label use is particularly common in pregnant women and in children and young people because these groups have often been excluded from clinical trials during medicine development.

Medicines with no UK marketing authorisation for any indication (unlicensed medicines) will not usually be considered in a guideline because there is no UK assessment of safety and efficacy to support their use. Unlicensed medicines may be included in exceptional circumstances, for example in complex conditions when there are no other treatment options. This should be agreed with the medicines adviser responsible for quality assurance (see also the section on recommendations on medicines, including off-label use of licensed medicines).

The Medicines and Healthcare products Regulatory Agency (MHRA) is responsible for ensuring that medicines meet the required standards of safety, that is, expected benefits outweigh risks of harmful effects. Therefore, we take account of national medicines safety advice and do not assess the safety of licensed medicines unless prior agreement has been reached with the MHRA (these discussions can be initiated via the medicines adviser responsible for quality assurance). In particular, we rarely consider review questions that just cover the clinical risk-benefit balance of a drug versus placebo or no treatment for its licensed indication. National medicines safety advice does not require formal quality assessment but should be reported in the evidence review document and used to inform committee discussions. See the chapter on interpreting the evidence and writing the guideline for more information.

When the effectiveness of a medicine is being considered, it is expected that the evidence review will include medicines safety outcomes, such as adverse events. These outcomes are also likely to be included as part of assessing the overall impact on quality of life in any cost-effectiveness analysis, or as part of supporting shared decision making.

Sources of national medicines safety advice include:

- MHRA drug safety updates
- NHS England's national patient safety alerts.

If there is a need for additional information on drug safety for a particular topic (for example, pregnancy or breastfeeding) advice on how to obtain this should be sought from the medicines adviser responsible for quality assurance.

When including an off-label use of a licensed medicine, it is usually possible to extrapolate national medicines safety advice, for example, if the population is similar and the recommended dosage is likely to be within the licensed dose range. The approach should be discussed with the committee as early as possible and agreed with the team responsible for quality assurance.

In some circumstances, primary evidence on medicines safety may be needed, for example, for unlicensed medicines, or off-label use of licensed medicines when it is not expected that extrapolating safety information from the licensed uses is appropriate. This should be identified as early as possible, stated in the review protocol, and agreed with the team responsible for quality assurance. It is important to note that different evidence types will often be needed for assessing safety, because both rare serious events and long-term events are important, and are often not captured in research studies.

Study designs for review questions about the effectiveness of an intervention

Randomised controlled trials (RCTs) are the preferred study design for estimating the effects of interventions. This is because randomisation ensures that any differences in known and unknown baseline characteristics between groups are due to chance; blinding (where applied) prevents knowledge of treatment allocation from influencing behaviours; and standardised protocols ensure consistent data collection.

However, RCTs are not always available or may not be sufficient to address the review question of interest.

In such cases or other circumstances where the randomised evidence is insufficient for decision making, non-randomised studies may sometimes be appropriate for estimating the effects of interventions. There are also some published studies that include analysis of large, high-quality primary data sources (such as patient registries). The Medical Research Council (MRC) has produced guidance on evaluating complex interventions (Skivington et al. 2021b) and using natural experiments to evaluate population health interventions (UK research and Innovation [UKRI], 2022). More information on appropriate use for non-randomised evidence to estimate the effects of interventions is given in the NICE real-world evidence framework.

Review questions that consider antimicrobial interventions or interventions that could reduce the use of antimicrobials

Review questions on antimicrobial interventions should take account of the principles of good antimicrobial stewardship – see the NICE guideline on antimicrobial stewardship. Antimicrobial-sparing interventions (such as self-care treatments) should be considered for self-limiting conditions (such as sore throat and conjunctivitis) or other conditions where they are relevant.

In line with these principles, the review protocol should include how the following will be considered in the evidence review:

- antimicrobial resistance
- antibiotic choice, duration, dosage and route of administration
- reviewing and stepping down treatment, if appropriate (for example, if intravenous or prophylactic antibiotics are included).

Information on antimicrobial resistance can be identified from various sources, for example from:

- UKHSA's English surveillance programme for antimicrobial utilisation and resistance (ESPAUR) report
- UKHSA's AMR local indicators
- summaries of product characteristics.

Any relevant information should be summarised in the evidence review document and does not need quality assessment, see the chapter on interpreting the evidence and writing the guideline for more information.

Studies identified from the literature search may report outcomes that measure resistance or usage (for antimicrobial-sparing interventions). These outcomes should be routinely included in the review protocol.

In most situations, evidence reviews should aim to identify the most appropriate choice(s) of antibiotics and the optimal duration of treatment, to minimise the risk of antimicrobial resistance (also see the section on wording the recommendations in the chapter on interpreting the evidence and writing the guideline).

Review questions about complex interventions

Skivington et al. 2021b states: 'An intervention might be considered complex because of properties of the intervention itself, such as the number of components involved; the range of behaviours targeted; expertise and skills required by those delivering and receiving the intervention; the number of groups, settings, or levels targeted; or the permitted level of flexibility of the intervention or its components'.

Review questions about the effectiveness of complex interventions may need additional analyses to understand the complexity of the interventions, or additional types of evidence to answer different aspects of the question such as qualitative research or real-world evidence. For example, additional evidence might address the views of people using the service or intervention, or barriers to use, as reported by practitioners or providers, or the impact on health inequalities. In this case, 2 related review questions (quantitative and qualitative) may be used to address the issues. These different forms of data can be considered separately, or together (integrated) if needed to understand the underlying problem. A review of effectiveness may also include evidence of the intervention's mechanism of action, that is, evidence of how the intervention works. All these examples of questions about the effectiveness of interventions may be addressed by 2 related review questions or by a mixed methods review.

When formulating questions to assess complex interventions, routinely consider whether to factor in the sociocultural acceptability and accessibility of an intervention, as well as contextual factors that impact on intervention feasibility. Qualitative evidence synthesis is one method of exploring these factors (Booth et al. 2019). An extended question framework (PerSPEcTiF) is proposed to recognise these wider issues, while also being particularly suited to qualitative evidence synthesis and complex intervention reviews (see table 21.5a in chapter 21 of the Cochrane Handbook for Systematic Reviews of Interventions). The PerSPEcTiF model of question formulation includes the following elements: Per = Perspective, S = Setting, P = Phenomenon of interest or problem, E = Environment, (c) = Comparison (optional), Ti = Time or timing, F = Finding.

Examples of review questions that include analysis targeting complex interventions are shown in box 4.3.

Box 4.3 Examples of review questions that include analysis targeting complex interventions

What interventions and services are most effective for supporting the wellbeing of informal carers of people living with dementia?

What intervention components (alone or in combination) and approaches are most effective and acceptable in helping children and young people living with overweight or obesity?

Are psychological interventions with a particular component (or combination of components) effective for people with coronary heart disease in relation to reducing all-cause mortality, cardiac mortality, non-fatal myocardial infarction, total cholesterol, blood pressure, anxiety and depression?

School-based self-management interventions for asthma in children and adolescents: a mixed methods systematic review:

- (a) What intervention components and processes are aligned with successful school-based asthma self-management intervention implementation?
- (b) What is the effectiveness of school-based interventions for improvement of asthma self-management on children's outcomes?

Community engagement for health via coalitions, collaborations and partnerships (online social media and social networks):

- (a) What is the extent of community engagement across design, delivery and evaluation in online social media and online social networking interventions?
- (b) What health issues and populations have been studied using online social media and social networking?
- (c) How effective are online social networks in improving health and wellbeing and reducing health inequalities?
- (d) Do particular programme features (for example, health topic, extent of engagement, population type) account for heterogeneity in effect size estimates across studies?
- (e) What processes are aligned with effective interventions?

Review questions about the accuracy of diagnostic tests

Review questions about diagnosis are concerned with the performance of a diagnostic test or test strategy to identify the presence of a current condition in people. They begin at the point in the diagnostic process when a professional has already made an initial diagnosis, or diagnoses, based on their clinical judgement. Diagnostic tests to confirm or rule out the initial diagnosis can include history-taking, symptoms, signs, identification tools, laboratory or pathological examination, and imaging tests.

Broadly, there are 2 types of review questions about diagnostic tests:

- questions about the diagnostic accuracy (or diagnostic yield) of a test or a number of tests compared individually against a comparator (the reference standard)
- questions about the diagnostic accuracy (or diagnostic yield) of a test strategy (such as serial testing) against a comparator (the reference standard).

Questions looking at the accuracy of multivariable diagnostic prediction models are covered in the section on review questions about predicting an individual prognosis or identifying an individual diagnosis.

In studies of the accuracy of a diagnostic test, the results of the test under study (the index test) are compared with those of the best available test (the reference standard). It is important to be clear when deciding on the question what the exact proposed use of the test is (for example, as an identification tool, an initial 'triage' test or after other tests).

The PICTO (population, index test, comparator, target condition and outcome) framework can be useful when formulating review questions about diagnostic test accuracy (see box 4.4). However other frameworks (such as PPIRT; population, prior tests, index test, reference standard, target condition) can be used if helpful.

Box 4.4 Features of a well-formulated review question on diagnostic test accuracy using the PICTO framework

Population: To which populations would the test be applicable? How can they be best described? Are there subgroups that need to be considered?

Index test: The test or test strategy being evaluated for accuracy.

Comparator or reference standard: The test with which the index test is being compared, usually the reference standard (the test that is considered to be the best available method for identifying the presence or absence of the condition of interest – this may not be the one that is routinely used in practice).

Target condition: The disease, disease stage or subtype of disease that the index test(s) and the reference standard are being used to identify.

Outcome measure: The diagnostic accuracy of the test or test strategy for detecting the target condition. This is usually reported as test parameters, such as sensitivity, specificity, predictive values, likelihood ratios, or – when multiple thresholds are used – a receiver operating characteristic (ROC) curve.

A review question about diagnostic test accuracy is usually best answered by a cross-sectional study in which both the index test and the reference standard are performed on the same sample of people. Cohort and case–control studies are also used to assess the accuracy of diagnostic tests, but these types of study design are more prone to bias (and often result in inflated estimates of diagnostic test accuracy). Further advice on the types of study to include in evidence reviews of diagnostic test accuracy can be found in the Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy.

Examples of review questions on diagnostic test accuracy are shown in box 4.5.

Box 4.5 Examples of review questions on diagnostic test accuracy

In people with suspected hypertension, which test is most accurate in identifying whether hypertension is present, as indicated by the reference standard of ambulatory blood pressure measurement?

In adults with diabetes, what are the best clinical predictors or biomarker tests (alone or in combination) to distinguish between diagnosis of type 1 diabetes, type 2 diabetes, and other forms of diabetes?

Which of the following, alone or in combination, constitutes the most accurate pathway for diagnosing prostate cancer: multiparametric MRI; transrectal ultrasonography (TRUS) biopsy; transperineal template biopsy?

In adults, children, and young people from black, Asian and other minority ethnic groups with chronic kidney disease, what is the diagnostic accuracy of estimated glomerular filtration rate (eGFR) calculations?

In people with suspected prostate cancer (with any of the following symptoms – any lower urinary tract symptoms, such as nocturia, urinary frequency, hesitancy, urgency or retention or erectile dysfunction or visible haematuria), what is the diagnostic accuracy of fixed prostate-specific antigen (PSA) test threshold compared to age-adjusted PSA thresholds?

In people with suspected cow's milk allergy, should skin prick tests rather than an oral food challenge with cow's milk be used for diagnosis and management?

What are the symptoms and signs of urinary tract infection (UTI) in babies, children and young people under 16 years old?

Although assessing test accuracy is important for establishing the usefulness of a diagnostic test, the value of a test usually lies in how useful it is in guiding treatment decisions or the provision of services, or supporting shared decision making, and ultimately in improving outcomes. Review questions aimed at establishing the value of a diagnostic test in practice can be structured in a similar way to questions about interventions. The best study design is a test-and-treat RCT. These compare outcomes for people who have a new diagnostic test (in combination with a management strategy) with outcomes of people who have the usual diagnostic test and management strategy. These types of study are not very common so evidence about diagnostic test accuracy is usually

also needed.

Review questions about prognosis

Information about prognosis can be used within guidelines to:

- classify people into risk categories (for example, cardiovascular risk or level of need) so that different interventions or preventative strategies can be applied
- define subgroups of populations that may respond differently to interventions
- identify factors that can be used to adjust for case mix (for example, in investigations of heterogeneity)
- help determine longer-term outcomes not captured within the timeframe of a trial (for example, for use in an economic model).

Review questions about prognosis address the likelihood of an outcome for an individual person from a population at risk for that outcome, based on the presence of a proposed prognostic factor or factors (see box 4.6). A helpful structured approach for developing questions about prognosis is the PICOTS (population, index prognostic factor, comparator prognostic factors, outcome, timing, setting) framework.

Review questions about prognosis may be closely related to questions about aetiology (cause of a disease or need) if the outcome is viewed as the development of the disease or need based on a number of risk factors.

Questions looking at the accuracy of multivariable prognostic prediction models are covered in the section on review questions about predicting an individual prognosis or identifying an individual diagnosis.

Box 4.6 Examples of review questions on prognosis

What is the best combination of measures of kidney function and markers of kidney damage to identify increased risk of progression in adults, children and young people with chronic kidney disease?

What are the factors (for example, mental health problems, substance misuse, medication that may cause impulse control disorders) that may increase the chance of a person participating in harmful gambling?

A review question about prognosis is best answered using a prospective cohort study with multivariable analysis. Case-control studies and cross-sectional studies are not usually suitable for answering questions about prognosis because they do not estimate baseline risk; they give only an estimate of the likelihood of the outcome for people with and without the prognostic factor. When developing a review question on prognosis, it is also important to consider possible confounding factors and whether some prognostic factors are modifiable and others non-modifiable.

Review questions about predicting an individual prognosis or identifying an individual diagnosis

Statistical analyses can be used to develop prediction models for a specific diagnosis or prognosis. These models are usually developed using multivariable modelling methods. Multivariable prediction models are developed to help healthcare professionals estimate the probability or risk that a specific disease or condition is present (diagnostic prediction models) or that a specific event will occur in the future (prognostic prediction models). They are usually developed using a multivariable model – a mathematical equation that relates multiple predictors for a particular person to the probability of or risk for the presence (diagnosis) or future occurrence (prognosis) of a particular outcome. Other names for a prediction model include risk prediction model, predictive model, prognostic (or prediction) index or rule, and risk score.

Diagnostic prediction models can be used to inform who should be referred for further testing, whether treatment should be started directly, or to reassure patients that a serious cause for their symptoms is unlikely. Prognostic prediction models can be used for planning lifestyle or treatment decisions based on the risk for developing a particular outcome or state of health in a given period.

Prediction model studies can be broadly categorised into those that develop models,

those that validate models (with or without updating the model) and those that do both. Studies that report model development aim to derive a prediction model by selecting the relevant predictors and combining them statistically into a multivariable model. Logistic and Cox regression are most frequently used for short-term (for example, disease absent versus present, 30-day mortality) outcomes and long-term (for example, 10-year risk) categorical outcomes, respectively. Studies may also focus on quantifying how much value a specific predictor (for example, a new predictor) adds to the model. Outcomes that are important should be agreed by the committee and specified in the review protocol.

Quantifying the predictive ability of a model using the same data from which the model was developed (often referred to as apparent performance) tends to overestimate performance. Studies reporting the development of new prediction models should always include some form of validation to quantify any optimism in the predictive performance (for example, calibration and discrimination). There are 2 types of validation: internal validation and external validation. Internal validation uses only the original study sample with methods such as bootstrapping. External validation evaluates the performance of the model with data not used for model development. The data may be collected by the same investigators or other independent investigators, typically using the same predictor and outcome definitions and measurements, but using a different sample (for example, from a later time period). If validation indicates poor performance, the model can be updated or adjusted on the basis of the validation data set. For more information on validating prediction models, see Steyerberg et al. 2001, 2003, 2009; Moons et al. 2012; Altman et al. 2009; and Justice et al. 1999.

For more information on developing review questions about prediction models, see the TRIPOD statement in the Annals of Internal Medicine and the TRIPOD statement: explanation and elaboration in the Annals of Internal Medicine.

Although assessing predictive accuracy is important for establishing the usefulness of a prediction model, the value of a prediction model lies in how useful it is in guiding treatment or management decisions, or the provision of services, and ultimately in improving outcomes. Review questions aimed at establishing the value of a prediction model in practice, for example, to compare outcomes of people who were identified from a prediction model (in combination with a management strategy) with outcomes of people who were identified opportunistically (in combination with a management strategy) can be structured in the same way as questions about interventions.

Box 4.7 Examples of review questions on prediction models

Diagnostic prediction models

Which scoring tools for signs and symptoms (including Centor and FeverPAIN) are most accurate in predicting sore throat caused by group A beta-haemolytic streptococcus infection in primary care?

What is the accuracy of clinical prediction models and tools (clinical history, cardiovascular risk factors, physical examination) in evaluating stable chest pain of suspected cardiac origin?

Prognostic prediction models

What is the effectiveness of prediction tools for identifying women at risk of pelvic floor dysfunction?

Are kidney failure prediction equations good predictors of progression, kidney failure or end-stage renal disease?

What risk assessment or prediction tool best identifies people with multiple conditions who are at risk of unplanned hospital admission?

What risk tool best identifies people with type 2 diabetes who are at risk of reduced life expectancy?

Which risk assessment tools are the most accurate in predicting the risk of fragility fracture in adults with osteoporosis or previous fragility fracture?

What factors and baseline characteristics are accurate in predicting positive treatment outcomes in people with pancreatic cancer?

In people with localised or locally advanced prostate cancer, which risk stratification models, tools and categorising systems perform better in indicating risk of poor outcomes?

In people with stable chronic obstructive pulmonary disease, does routine assessment using a multidimensional severity assessment index (such as BODE [BMI, airflow obstruction, dyspnoea/ breathlessness and exercise capacity]) better predict

outcomes than forced expiratory volume in 1 second (FEV1) alone?

Qualitative review questions

In some circumstances, specific questions should be formulated about the views and experience of people using services, family members and carers, and the public. The views and experiences of those commissioning and providing services may also be relevant. Qualitative questions do not have to be linked to an effectiveness question; they can stand alone in a single evidence review. Qualitative questions can cover a range of dimensions, including:

- views and experiences of people using or providing services, family members or carers or the public on the effectiveness and acceptability of interventions
- preferences of people using services, family members or carers or the public for different treatment or service options, including the option of foregoing treatment or care
- views and experiences of people using or providing services, family members or carers or the public on what constitutes a desired, appropriate or acceptable outcome
- elements of care or a service that are of particular importance to people using or providing services
- factors that encourage or discourage people from using interventions or services
- the specific needs of certain groups of people using services, including those sharing the characteristics protected by the Equality Act (2010)
- information and support needs specific to the topic
- which outcomes reported in studies of interventions are most important to people using services, family members or carers or the public
- health and care inequalities.

As for other types of review question, questions that are broad and lack focus (for example, 'What is the experience of living with condition X?') should not be asked. The question should be more focused by including the phenomenon of interest such as

acceptability, accessibility, preferences, information and support needs, feasibility or implementation.

The PICO (Population, phenomena of Interest, Context) framework and the SPIDER framework in McMaster University's National Collaborating Centre for Methods and Tools registry are examples of frameworks that can be used to structure qualitative evidence synthesis (QES) review questions. Examples of QES review questions are shown in box 4.8.

Box 4.8 Examples of QES review questions

What works well, and what could be improved, about access to, engagement with and delivery of health and social care for people experiencing homelessness?

What is the experience of disabled children and young people with severe complex needs and their families and carers of joint delivery of health, social care and education services?

What are the barriers and facilitators to, and key aspects of (including systems and processes), the successful implementation or delivery of mental wellbeing interventions, programmes, policies or strategies at work?

What factors influence the acceptability of, access to, and uptake of cardiac rehabilitation services?

What are the views and experiences of health, social care and other practitioners about the practicality and implementation of home-based intermediate care?

What are the barriers and facilitators to identifying children and young people at risk of poor social, emotional and mental wellbeing?

Other evidence used for views and experiences

While qualitative studies often answer a review question about the views or experiences of people using or providing services, family members or carers or the public, other options include quantitative patient preference studies. Information on views and experiences is also becoming increasingly available as part of some intervention studies, for example, collecting qualitative data from trial participants. When qualitative and quantitative data are generated from the same trial in this way, they are referred to as 'trial siblings'. For

more information see chapter 21 of the Cochrane Handbook for Systematic Reviews of Interventions.

When there is a lack of evidence on issues important to people affected by the guideline (including families and carers, where appropriate), the development team should consider seeking additional information outside of that from formal literature searching. This could be through a call for evidence (see the section on calls for evidence from stakeholders in the chapter on identifying the evidence: literature searching and evidence submission). It could also be through additional consultation, commissioned primary research, or by approaching experts who may have access to additional data sources (such as surveys of people's views and experiences) or who may be able draw on their experience of working in the field. For more information, see:

- the section on other attendees at committee meetings in the chapter on decision-making committees
- the appendix on approaches to additional consultation and commissioned primary research
- the appendix on call for evidence and expert witnesses.

Mixed methods approaches

For further information on mixed methods reviews, see chapter 8 of the JBI Manual for Evidence Synthesis.

Lizarondo et al. 2020 states: 'The core intention is to combine quantitative and qualitative data (from primary studies) or integrate transformed quantitative evidence and qualitative evidence to create a breadth and depth of understanding that can confirm or dispute evidence and ultimately answer the review question/s posed'.

'Dependent on the nature of the review question mixed methods systematic reviews may allow for:

- an examination of the degree of agreement between quantitative and qualitative data to validate or triangulate results/findings,
- identification of discrepancies within the available evidence,

- determination of whether the quantitative and qualitative data address different aspects of a phenomenon of interest, and
- one type of data that can explore, contextualize or explain the findings of the other type of data.'

A mixed method approach is only needed when a question cannot be answered by separate quantitative or qualitative evidence synthesis and when multiple perspectives are needed to understand the underlying problem. For example, if the method of intervention delivery affects whether people engage with it and that has an impact on its effectiveness. Another example is when integrating the 2 types of evidence provides more evidence or explanation than the separate quantitative and qualitative evidence reviews. Reasons to use a mixed methods review may include:

- to use qualitative data (such as barriers and facilitators or explanatory factors from peoples' experiences of having or giving an intervention) to:
  - explain quantitative results (for example if the mechanisms of action behind an intervention are unclear, or very different results are seen in different studies or populations), or
  - supplement quantitative evidence when it is limited (for example in certain populations).
- when results are inconclusive, for example if a systematic review of quantitative data finds no effects, but there is a plausible mechanism of action and it is unclear why the intervention shows no effect. It should be noted that:
  - inconclusive results may arise because other outcomes may be more important than those assessed,
  - plausible mechanisms do not always translate into real-world effects or
  - people with experience of the intervention may be able to offer other explanations for an intervention's lack of impact.
- to better contextualise results, for example to understand how to reach certain populations (this may include removing structural barriers, or other things potentially relevant to health inequalities such as lower uptake in some groups than others, including why and how can it be addressed)
- when it is unclear who the intervention is effective for and when it is most effective,

and not enough data can be gathered from subgroup analysis or meta-regression (for example, the more sensitive the intervention is to context, the more important it may be to look at the context for where, when and in whom it is most effective)

- when the implementation of the intervention needs to be considered, either because:
  - of previous evidence that implementation is complicated and inconsistent leading to mixed use or differential uptake across the country, or
  - implementation is likely to be complicated and a recommendation may be needed to provide guidance on what could be useful to achieve this.

Using a mixed methods approach requires integration of evidence. That is:

- quantitative and qualitative data are synthesised separately and juxtaposed so that different dimensions of a phenomenon (qualitative) may explain the outcomes of the quantitative synthesis (convergent segregated), or
- when both quantitative and qualitative data can answer a single question and so data are transformed, and quantitative and qualitative studies synthesised simultaneously (convergent integrated).

When developing a mixed methods review question, it is helpful to consider a number of things to help guide the method to take, such as sequence of synthesis, approach to integration and the nature of the question (for more information, see sections 8.2 and 8.3 in chapter 8 of the JBI Manual for Evidence Synthesis).

Examples of mixed methods review questions are shown in box 4.9.

Box 4.9 Examples of mixed methods review questions

Convergent segregated:

What is the effectiveness of integrated working among registered social workers and other practitioners to support adults with complex needs, and based on their views and experiences what are the barriers to integrated working?

- (a) What is the effectiveness of integrated working among registered social workers and other practitioners to support adults with complex needs?
- (b) Based on the views and experiences of everyone involved, what are the facilitators and barriers to integrated working between registered social workers and other practitioners to support adults with complex needs?

What is the efficacy of telehealth and mobile health interventions and what are the benefits and challenges of these interventions in patients with inflammatory bowel disease?

- (a) Are telehealth and mobile health interventions effective in improving the health-related outcomes of adults with IBD?
- (b) What are the perceived benefits and challenges of telehealth and mobile health interventions by adults with IBD?

What are the effects of clinical supervision of healthcare professionals on organisational outcomes?

- (a) What are healthcare professionals' experiences, views, and opinions regarding clinical supervision as it relates to organisational processes and outcomes?
- (b) What can be inferred from the qualitative synthesis of healthcare professionals' experiences or views that can explain the effects of clinical supervision or inform its appropriateness and acceptability for health professionals?

Convergent integrated

End-of-life care preferences of older patients with multimorbidity: willingness to receive life-sustaining treatments, place of care, and shared decision-making

processes (González-González 2021)

Lifestyle interventions through participatory research: a mixed methods systematic review of alcohol and other breast cancer behavioural risk factors: what works and how? (Thomas 2022)

What is the clinical effectiveness of self-management in adolescents with asthma, and what factors are perceived by them as important to maintain adherence to their self-management plan? (Lizarondo 2021)

Review questions about service delivery

Guidelines sometimes cover areas of service delivery. These might include the relative effectiveness of different models of service delivery, how delivery of services could improve, how delivery of services impact on health inequalities, or what the core components of services are and how different components could be re-configured.

Box 4.10 Examples of review questions on service delivery

In people with hip fracture what is the effectiveness of hospital-based multidisciplinary rehabilitation on the following outcomes: functional status, length of stay in secondary care, mortality, place of residence or discharge, hospital readmission and quality of life?

What is the effectiveness of surgeon seniority (consultant or equivalent) in reducing the incidence of mortality, the number of people requiring reoperation, and poor outcome in terms of mobility, length of stay, wound infection and dislocation?

What are the best service models to support the identification of people who may be entering their last year of life?

What are the most effective approaches and activities to normalise shared decision making in the healthcare system?

What are the most effective service models for weight management services that would improve uptake in population groups with low uptake?

What types of needle and syringe programmes (including their location and opening times) are effective?

What regional or city-level commissioning models, service models, systems and service structures are effective in:

- reducing diagnostic delay for tuberculosis (TB)
- improving TB contact tracing
- improving TB treatment completion?

A review question about the effectiveness of service delivery models is usually best answered by a pragmatic RCT, if it is feasible to do one. However, a wide variety of methodological approaches and study designs can be used, including non-randomised studies that report observational data (including routine healthcare and audit data), experimental and qualitative evidence. Other types of questions on service delivery are also likely to be answered using evidence from study types other than RCTs. For example, to determine whether an intervention will work for a particular subgroup or setting that

does not have specific evidence from an RCT, we might want to know how the intervention works, which will require evidence of the relevant underlying mechanisms.

Depending on the type of review questions, the PICO framework may be appropriate but other frameworks can be used.

When a topic includes review questions on service delivery, approaches described in the chapter on incorporating economic evaluation and the appendix on service delivery – developing review questions, evidence reviews and synthesis may be used. Such methods should be agreed with staff with responsibility for quality assurance and should be clearly documented in the guideline.

Review questions about descriptive epidemiology

Some epidemiological reviews describe the problem under investigation and can be used to inform other review questions. For example, an epidemiological review of incidence or prevalence of a condition would provide baseline data for further evidence synthesis, an epidemiological review of accidents would provide information on the most common accidents, as well as morbidity and mortality statistics, and data on inequalities in the impact of accidents. These review questions may also be necessary to provide input data for economic modelling.

Box 4.11 Examples of review questions that might benefit from a descriptive epidemiological review

What are the patterns of physical activity among children from different populations and of different ages in England?

Which populations of children are least physically active and at which developmental stage are all children least physically active?

Which population groups are disproportionately affected by type 2 diabetes mellitus?

What is the incidence of Lyme disease in the UK?

The structure of the question and the type of evidence will depend on the aim of the review.

Other epidemiological reviews describe relationships between epidemiological factors and

outcomes – a review on associations. If an epidemiological review has been carried out, information will have been gathered from observational studies on the nature of the problem. However, further analysis of this information – in the form of a review on associations – may be needed to establish the epidemiological factors associated with any positive or negative behaviours or outcomes.

Box 4.12 Examples of review questions that might benefit from a review on epidemiological associations

What factors are associated with children's or young people's physical activity and how strong are the associations?

What physiological and aetiological factors are associated with coeliac disease?

What physical, environmental and sociological factors are associated with the higher prevalence of multiple sclerosis in European countries?

What factors are associated with a higher mortality rate from breast cancer in people from the most deprived quintile?

Review questions that consider cost effectiveness

In most NICE guidelines, it is expected that considerations around cost effectiveness will be included for all review questions. Therefore, it is not necessary to explicitly mention cost effectiveness in the review question itself. If some review questions for a guideline will not consider cost effectiveness, it should be specified which questions will and will not be including these considerations. For more information on review questions that consider cost effectiveness, see the chapter on incorporating economic evaluation.

4.4 Review protocols

For each guideline evidence review, a review protocol is prepared that outlines the background, the objectives and the planned methods. In addition, the review protocol should make it possible for the review to be repeated by others. A protocol should also make it clear how equality and health inequalities issues have been considered in planning the review work (see the section on considering health inequalities when preparing review protocols).

Structure of the review protocol

The protocol should describe any differences from the methods described in this manual (see the chapters on identifying the evidence: literature searching and evidence submission, reviewing evidence, and incorporating economic evaluation), rather than duplicating the manual.

Templates for the 4 common review types (intervention, diagnostic accuracy, qualitative and prognosis) are available in the appendix on the review protocol templates. These templates can be amended to match the specifics of an individual review question. For reviews not covered by these templates, a similar level of detail should be provided in the review protocol.

When a guideline is updating an evidence review from a published NICE guideline, the protocol from the published guideline, if available, should be used as a starting point for developing the new review protocol. It should be updated based on any changes since the original protocol was developed (such as new interventions and comparators, and extensions of the population). The level of changes needed is likely to depend on the length of time since the original protocol was developed. No more changes than necessary should be made, as the closer the protocol remains to the original, the easier it will be to reuse data extracted in that original review.

Process for developing the review protocol

The review protocol should be drafted by the development team, with input from the guideline committee, and then reviewed and approved by staff with responsibility for quality assurance. This should take place after agreeing the review question and before starting the evidence review.

Although an original systematic review of primary studies is the most common method to answer a review question, there are several other possible alternatives that should be considered, such as:

- Making use of a previously published systematic review or qualitative evidence synthesis. This review could either be used without further modification, or as a starting point for additional work (for example, to include studies published after the review search date or additional outcomes that may be relevant to the guideline but were not included in the original review). See the section on existing systematic

reviews for more details.

- Doing a review of reviews. This involves doing a systematic search for published systematic reviews or qualitative evidence syntheses and using these reviews as the included evidence. No further analysis is done on the primary studies in those reviews.
- Using formal consensus methods, such as Delphi panels or nominal group technique. These techniques can be used instead of doing formal evidence searches, or as a way to interpret the evidence found from these searches (for example, if a large volume of lower quality evidence is available).
- Using informal committee consensus to make recommendations, without searching for evidence first. This is only likely to be suitable in situations where the development team is confident that no evidence is likely to exist that would help inform recommendations.
- Adapting recommendations from previously published guidelines, either other NICE guidelines or guidelines from other organisations that are assessed to be sufficiently high quality using the AGREE II instrument.
- Doing primary analysis of real-world data (such as routinely collected NHS or registry data). The NICE real-world evidence framework provides advice on situations where this type of analysis may be appropriate and outlines best practices for identifying and assessing data sources and doing the analysis. For questions on effectiveness of interventions such analyses are likely to be undertaken when randomised evidence is not available or sufficient to address the research question of interest, while for other question types (such as prognostic or epidemiological) this may represent the optimal type of evidence.
- Using calls for evidence and expert witnesses to obtain evidence that may not be available from standard literature searches (see the section on other attendees at committee meetings in the chapter on decision-making committees, and the appendix on call for evidence and expert witnesses).

More than one of these methods may need to be used for some review questions, if different parts of the question need different approaches. It will often not be possible to describe evidence reviews using any of these methods in the format of a standard review protocol. When this is not possible a narrative description of the review plan can be produced instead. This should clearly describe the planned approach to the review, and the reason why this approach was preferred over doing an original review of published primary data.

When considering which of these potential approaches to use, it is important to consider the trade-off between the optimal evidence to address a question, and the additional time and resources needed to gather that evidence. In particular, it should be considered whether any additional work is likely to lead to different recommendations being made.

All review protocols should consider registering on the PROSPERO database before the data extraction commences, if possible and appropriate. The review questions are published on the NICE website at least 6 weeks before consultation on the draft recommendations. Any changes made to a protocol during guideline development should be agreed with staff who have responsibility for quality assurance and the version on the website updated. Any deviations from the signed-off and published protocol should be clearly stated and justified in the evidence review document. If protocols are published anywhere else (for example, on the PROSPERO database) development teams should ensure the versions of the protocol are consistent.

Existing systematic reviews

Often reviews of quantitative or qualitative studies (secondary evidence) already exist (for example, those developed by internationally recognised producers of systematic reviews such as Cochrane, the Campbell Collaboration and the Joanna Briggs Institute among others). Existing reviews may include systematic reviews (with or without a meta-analysis or individual participant data analysis) and non-systematic literature reviews and meta-analyses. Well-conducted systematic reviews may be of particular value as sources of evidence (see the appendix on appraisal checklists, evidence tables, GRADE and economic profiles for checklists to assess risk of bias or quality of studies when developing guidelines). Some reviews may be more useful as background information or as additional sources of potentially relevant primary studies. This is because they may:

- not cover inclusion and exclusion criteria relevant to the guideline topic's referral and parameters (for example, comparable research questions, relevant outcomes, settings, population groups or time periods)
- group together different outcome or study types
- include data that are difficult or impossible to separate appropriately
- not provide enough data to develop recommendations (for example, some reviews do not provide sufficient detail on specific interventions making it necessary to refer to the primary studies).

Conversely, some high-quality systematic reviews (as assessed using the checklists recommended in the appendix on appraisal checklists, evidence tables, GRADE and economic profiles) may provide enhanced data not available in the primary studies. For example, authors of the review may have contacted the authors of the primary studies or other related bodies in order to include additional relevant data in their review, or may have undertaken additional analyses (such as individual participant data analyses). In addition, if high-quality reviews are in progress (protocol published) at the time of development of the guideline, the development team may choose to contact the authors for permission to access pre-publication data for inclusion in the guideline (see the appendix on call for evidence and expert witnesses).

Systematic reviews can also be useful when developing the scope and when defining review questions, outcomes and outcome measures for the guideline evidence reviews. The discussion section of a systematic review can also help to identify some of the limitations or difficulties associated with a topic, for example, through a critical appraisal of the limitations of the evidence. The information specialists may also wish to consider the search strategies of high-quality systematic reviews. These can provide useful search approaches for capturing different key concepts. They can also provide potentially useful search terms and combinations of terms, which have been carefully tailored for a range of databases.

High-quality systematic reviews that are directly applicable to the guideline review question can be used as a source of data instead of doing an original review. In such circumstances it can sometimes be beneficial to contact and collaborate with the authors of the original review, because it can be more efficient to share data rather than extract it from the published study.

When considering using results from an existing high-quality review, an assessment should be made of whether the parameters (for example, research question, PICO, inclusion and exclusion criteria) of the review are sufficiently similar to the review protocol of the guideline review question. If they are, the development teams and the committee should make a judgement on whether it is necessary to do an additional search for primary studies published after the search date covered by the existing review.

Considering health inequalities when preparing review protocols

When developing review protocols it is important to identify any health inequalities that may be relevant to the review question. This should involve identifying any issues from the

equality and health inequalities assessment that are particularly relevant to the particular question, as well as documenting any new issues identified by the committee.

The committee will need to consider these issues and any gaps in the evidence when interpreting the result of the review. However, in some circumstances it may be advisable to make modifications to the review protocol to ensure health inequalities are appropriately addressed, for example by:

- including relevant subgroups
- including outcomes that may be correlated to or explain inequalities (for example, including adherence as an outcome if this is a possible mechanism by which health inequalities are generated or exacerbated)
- including a wider range of study types if there are reasons to believe some groups are systematically excluded from a particular study design but will be included in others.

Other evidence

Depending on the nature of the guideline topic and the review question, other sources of relevant evidence such as reports, audits or service evaluations from the published or grey literature may be included. Often these will not need identifying from a systematic literature search (for example, if there is a national organisation responsible for producing reports on a particular subject). This should be agreed with staff who have responsibility for quality assurance and documented in the review protocol. When it is necessary to assess the quality, reliability and applicability of this evidence, it should be assessed according to standard processes (see the appendix on appraisal checklists, evidence tables, GRADE and economic profiles).

See also the chapter on linking to other guidance (which also covers using evidence from non-NICE guidance).

4.5 References and further reading

Altman DG, Vergouwe Y, Royston P et al. (2009) Prognosis and prognostic research: validating a prognostic model. BMJ 338: b605

Booth A, Noyes J, Flemming K, et al. Formulating questions to explore complex

interventions within qualitative evidence synthesis. BMJ Glob Health 2019;4:e001107

Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy | Cochrane Screening and Diagnostic Tests The Cochrane Collaboration

Collins G, Reitsma J, Altman D et al. (2015) Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. Annals of Internal Medicine 162: 55–63

Craig P, Dieppe P, McIntyre S et al. on behalf of the MRC (2008) Developing and evaluating complex interventions: the new Medical Research Council guidance. London: Medical Research Council

Craig P, Cooper C, Gunnell D et al. on behalf of the MRC (2011) Using natural experiments to evaluate population health interventions: new Medical Research Council guidance. London: Medical Research Council

Davis S, Ross H (2021) Telehealth and mobile health interventions in adults with inflammatory bowel disease: A mixed-methods systematic review. Research in nursing & health. Issue: 1. p.155-172

González-González A, Schmucker C, Nothacker J et al. (2021) End-of-Life Care Preferences of Older Patients with Multimorbidity: A Mixed Methods Systematic Review. J. Clin. Med. 10:91

Higgins JPT, Thomas J, Chandler J et al., editors (2022) Cochrane Handbook for Systematic Reviews of Interventions, version 6.2. The Cochrane Collaboration

Justice AC, Covinsky KE, Berlin JA (1999) Assessing the generalizability of prognostic information. Annals of Internal Medicine 130: 515–24

Kirkham JJ, Gorst S, Altman DG et al. (2016) Core Outcome Set–STAndards for Reporting: The COS-STAR Statement. PLoS: 21

Kirkham JJ, Davis K, Altman DG et al. (2017) Core Outcome Set-STAndards for Development: The COS-STAD Recommendations. PLoS: 23

Lizarondo L, Stern C, Carrier J et al. (2020) Chapter 8: Mixed methods systematic reviews.

In: Aromataris E, Munn Z (Editors), JBI Manual for Evidence Synthesis. JBI Downloaded 21-04-2022.

Martin P, Lizarondo L, Kumar S, Snowdon D (2021) Impact of clinical supervision on healthcare organisational outcomes: A mixed methods systematic review. PLoS ONE 16(11): e0260156

Moons KG, Kengne AP, Grobbee DE et al. (2012) Risk prediction models: II. External validation, model updating, and impact assessment. Heart 98: 691–8

Moons KGM, Altman DG, Reitsma JB et al. (2015) Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): explanation and elaboration. Annals of Internal Medicine 162: W1–W73

Noyes J, Booth A, Cargo M et al. (2022) Chapter 21: Qualitative evidence. In: Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA (editors). Cochrane Handbook for Systematic Reviews of Interventions version 6.3. Cochrane.

Petticrew M, Roberts H (2003) Evidence, hierarchies, and typologies: horses for courses. Journal of Epidemiology and Community Health 57: 527–9

Riley R, Moons K, Snell K et al. (2019) A guide to systematic review and meta-analysis of prognostic factor studies | The BMJ. BMJ 364:k4597

Skivington K, Matthews L, Simpson S et al. (2021a) Framework for the development and evaluation of complex interventions: gap analysis, workshop and consultation-informed update. Health Technol Assess 25:57

Skivington K, Matthews L, Simpson S et al. (2021b). A new framework for developing and evaluating complex interventions: update of Medical Research Council guidance. BMJ 374:n2061

Steyerberg E (2009) Clinical prediction models: a practical approach to development, validation, and updating. Springer

Steyerberg E, Harrell F, Borsboom G et al. (2001) Internal validation of predictive models - efficiency of some procedures for logistic regression analysis. Journal of Clinical Epidemiology 54: 774–81

Steyerberg EW, Bleeker SE, Moll HA et al. (2003) Internal and external validation of predictive models: A simulation study of bias and precision in small samples. Journal of Clinical Epidemiology 56: 441–7

Thomas J, Miller E, Ward P. (2022) Lifestyle Interventions through Participatory Research: A Mixed-Methods Systematic Review of Alcohol and Other Breast Cancer Behavioural Risk Factors. Int. J. Environ. Res. Public Health 19:980

Thomas J, Petticrew M, Noyes J et al. (2022) Chapter 17: Intervention complexity. In: Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA (editors). Cochrane Handbook for Systematic Reviews of Interventions version 6.3. Cochrane.

Viswanathan M, Melissa L, McPheeters Met et al. (2017) AHRQ series on complex intervention systematic reviews-paper 4: selecting analytic approaches. Journal of Clinical Epidemiology, 90:28

5 Identifying the evidence: literature searching and evidence submission

5.1 Introduction

The systematic identification of evidence is an essential step in developing NICE guideline recommendations.

This chapter sets out how evidence is identified at each stage of the guideline development cycle. It provides details of the systematic literature searching methods used to identify the best available evidence for NICE guidelines. It also provides details of associated information management processes including quality assurance (peer review), re-running searches, and documenting the search process.

Our searching methods are informed by the chapter on searching & selecting studies in the Cochrane Handbook for Systematic Reviews of Interventions and the Campbell Collaboration's searching for studies guide. The Summarized Research in Information Retrieval for HTA (SuRe Info) resource also provides research-based advice on information retrieval for systematic reviews.

Our literature searches are designed to be systematic, transparent, and reproducible, and minimise dissemination bias. Dissemination bias may affect the results of reviews and includes publication bias and database bias.

We use search methods that balance recall and precision. When the need to reduce the number of records for sifting requires pragmatic search approaches that may increase the risk of missing relevant studies, the context and trade-offs are discussed and agreed within the development team and made explicit in the reported search methods.

A flexible approach to identifying evidence is adopted, guided by the subject of the review question (see the chapter on developing review questions and planning the evidence review), type of evidence sought, and the resource constraints of the evidence review. Often an evidence review will be an update of our earlier work, therefore the approach can be informed by previous searches and surveillance reviews (see the chapter on ensuring that published guidelines are current and accurate).

5.2 Searches during guideline recommendation scoping and surveillance

Scoping searches

Scoping searches are top-level searches to support scope development. The purpose of the searches is to investigate the current evidence around the topic, and to identify any areas where an evidence review may be beneficial and any research gaps. The results of the searches are used to draft the scope of the upcoming guideline or update and to inform the discussions at scoping workshops (if held). Scoping searches do not aim to be exhaustive.

In some cases, scoping searches are not required when it is more efficient to use the surveillance review (see the chapter on the scope).

The sources searched at scoping stage will vary according to the topic, type of review questions the guideline or update will seek to address, and type of evidence sought. Each scoping search is tailored using combinations of the following types of information:

- NICE guidance and guidance from other organisations
- policy and legislation guides
- key systematic reviews and epidemiological reviews
- economic evaluations
- current practice data, including costs and resource use and any safety concerns
- views and experiences of people using services, their family members or carers, or the public
- other real-world health and social care data (for example audits, surveys, registries, electronic health records, patient-generated health data), if appropriate
- summaries of interventions that may be appropriate, including any national safety advice
- statistics (for example on epidemiology, natural history of the condition, service configuration or national prevalence data).

All scoping searches are fully documented and if new issues are identified at a scoping workshop, the search is updated. A range of possible sources considered for scoping searches is provided in the appendix on suggested sources for scoping.

Health inequalities searches

The purpose of these searches is to identify evidence to help inform the scope, health inequalities briefing, or the equality and health inequalities assessment (EHIA). They help identify key issues relevant to health inequalities on the topic, for example covering protected characteristics, groups experiencing or at risk of inequalities, or wider determinants of health.

The searches involve finding key data sources, such as routinely available national databases, audits or published reports by charities, non-governmental bodies, or government organisations.

Surveillance searches

Surveillance determines whether published recommendations remain current. The searches are tailored to the evidence required. This may include searches for new or updated policies, legislation, guidance from other organisations, or ongoing studies in the area covered by the evidence review.

If required, published evidence is identified by searching a range of bibliographic databases relevant to the topic. Surveillance searches generally use the same core set of databases used during the development of the original evidence review. A list of sources is given in the appendix on sources for evidence reviews.

The search approach and sources will vary between topics and may include:

- population and intervention searches
- focused searches for specific question areas
- forward and backward citation searching.

Searches usually focus on randomised controlled trials and systematic reviews, although other study types will be considered where appropriate, for example for diagnostic questions.

The search period starts at either the end of the search for the last update of a guideline evidence review, or at the last search date for any previous surveillance check. Where appropriate, living evidence surveillance could be set up to continuously monitor the publication of new evidence over a period of time until impact reaches the threshold for actions. For more information on NICE guideline recommendation surveillance, see the chapter on ensuring that guideline recommendations are current and accurate and appendix on surveillance - interim principles for monitoring approaches of guideline recommendations.

5.3 Searches during guideline recommendation development

Search protocols

Search protocols form part of the wider guideline review protocol (see the appendix on the review protocol template). They pre-define how the evidence is identified and provide a basis for developing the search strategies.

Once the final scope is agreed, the information specialist develops the search protocols and agrees them with the development team before the evidence search begins.

A search protocol includes the following elements:

- approach to the search strategy, tailored to the review question and eligibility criteria
- sources to be searched
- plans to use any additional or alternative search techniques, when known at the protocol development stage, and the reasons for their use
- details of any limits to be applied to the search
- references to any key papers used to inform the search approach.

Sources

Searches are done on a mix of bibliographic databases, websites and other sources, depending on the subject of the review question and the type of evidence sought.

For most searches there are key sources that are prioritised, and other potentially relevant sources that can be considered. It is important to ensure adequate coverage of the relevant literature and to search a range of sources. However, there are practical limits to the number of sources that can be searched in the standard time available for an evidence review.

The selection of sources varies according to the requirements of the review question.

Clinical intervention sources

For reviews of the effectiveness of clinical interventions the following sources are prioritised for searching:

- the Cochrane Central Register of Controlled Trials (CENTRAL)
- the Cochrane Database of Systematic Reviews (CDSR)
- Embase
- MEDLINE.

Clinical safety sources

In addition to the sources searched for clinical interventions, the following should be prioritised for clinical safety review questions:

- MHRA drug safety updates
- National patient safety alerts.

Antimicrobial resistance sources

For reviews of antimicrobial resistance, the following sources should be prioritised:

- UK Health Security Agency's English surveillance programme for antimicrobial utilisation and resistance (ESPAUR) report
- UK Health Security Agency's antimicrobial resistance local indicators.

Cost-effectiveness sources

For reviews of cost effectiveness, economic databases are used in combination with general bibliographic databases, such as MEDLINE and Embase (see appendix G on sources for economic reviews).

Economic evaluations of social care interventions may be published in journals that are not identified through standard searches. Targeted searches based on references of key articles and contacting authors can be considered to identify relevant papers.

Topic-specific sources

Some topics we cover may require the use of topic-specific sources. Examples include:

- PsycINFO (psychology and psychiatry)
- CINAHL (nursing and allied health professions)
- ASSIA (Applied Social Sciences Index and Abstracts)
- HealthTalk, and other sources to identify the views and experiences of people using services, carers and the public
- social policy and practice
- sociological abstracts
- transport database
- Greenfile (environmental literature)
- HMIC (Health Management Information Consortium).

Searching for model inputs

Evidence searches may be needed to inform design-oriented conceptual models. Examples include precise searches to find representative NHS costs for an intervention or finding out the proportion of people offered an intervention who take up the offer.

Some model inputs, such as costs, use national sources such as national list prices or national audit data. In some cases, it may be more appropriate to identify costs from the

academic literature. Further advice on methods to identify model inputs are also informed by Paisley (2016) and Kaltenhaler et al. (2011). See also the chapter on incorporating economic evaluation.

Real-world data

Information specialists can identify sources of real-world data (such as electronic health records, registries, and audits) for data analysts to explore further. The Health Data Research Innovation Gateway can be used to identify datasets. The NICE real-world evidence framework (2022) has additional guidance on searching for and selecting real-world data sources.

Grey literature

For some review questions, for example, where significant evidence is likely to be published in non-journal sources and there is a paucity of evidence in published journal sources, it may be appropriate to search for grey literature. Useful sources of grey literature include:

- HMIC (Health Management Information Consortium)
- GOV.uk
- TRIP database
- social policy and practice
- Canadian Agency for Drugs and Technology in Health (CADTH) Grey Matters resource.

Committee members may also be able to suggest additional appropriate sources for grey literature.

A list containing potential relevant sources is provided in the appendix on sources for evidence reviews.

Developing search strategies

The approach to devising and structuring search strategies is informed by the review protocol. The PICO (population, intervention, comparator and outcome) or SPICE (setting,

perspective, intervention, comparison, evaluation) frameworks may be used to structure a search strategy for intervention review questions. For other types of review questions, alternative frameworks may be more suitable.

It is sometimes more efficient to conduct a single search for multiple review questions, rather than conducting a separate search for each question.

Some topics may not easily lend themselves to PICO- or SPICE-type frameworks. In these cases, it may be better to combine multiple, shorter searches rather than attempting to capture the entire topic using a single search. This is often referred to as multi-stranded searching.

In some instances, for example where the terminology around a topic is diffuse or ill defined, it may be difficult to specify the most appropriate search terms in advance. In these cases, an iterative approach to searching can be used.

In an iterative approach, searching is done in several stages, with each search considering the evidence that has already been retrieved (for example, see Booth et al. 2020). Searching in stages allows the reviewers to review the most relevant, high-quality information first and then make decisions for identifying additional evidence if needed.

Decisions to use iterative approaches are agreed by the development team and staff with responsibility for quality assurance because it can affect timelines.

Updating previous work

Where high-quality review-level evidence is available on a topic, the review team may choose to update or expand this previous work rather than duplicating the existing findings. In these cases, the original review searches are re-run and expanded to account for any differences in scope and inclusion criteria between the original review and the update.

Cost-effectiveness searches

There are several methods that can be used to identify economic evaluations:

- All relevant review questions can be covered by a single search using the population search terms, combined with a search filter, to identify economic evidence.

- The search strategies for individual review questions can be combined with search filters to identify economic evidence. If using this approach, it may be necessary to adapt strategies for some databases to ensure adequate sensitivity.
- Economic evidence can be manually sifted while screening evidence from a general literature search (so no separate searches are required).

The rationale for the selected approach is recorded in the search protocol.

Where searches are needed to populate an economic model, these are usually done separately.

Identifying search terms

Search terms usually consist of a combination of subject headings and free-text terms from the titles and abstracts of relevant references.

When identifying subject headings, variations in thesaurus and indexing terms for each database should be considered, for example MeSH (Medical Subject Headings) in MEDLINE and Emtree in Embase. Not all databases have indexing terms and some contain records that have not yet been indexed.

Free-text terms may include synonyms, acronyms and abbreviations, spelling variants, old and new terminology, brand and generic medicine names, and lay and medical terminology.

For updates, previous search terms, including those from surveillance searches, are reviewed and used to inform new search terms. New or changed terms are identified, as well as any changes to indexing terms. This also applies when an existing review, for example a Cochrane review, is being updated to answer a review question.

Key studies can be a useful source of search terms, as can reports, guidelines, topic-specific websites, committee members and topic experts.

Some websites and databases have limited search functionality. It may be necessary to use fewer search terms or do multiple searches of the same resource with different search term combinations.

It may be helpful to use frequency analysis or text mining to develop the search-term

strategy. Tools such as PubReMiner and Medline Ranker can help, either by highlighting search terms that might not otherwise be apparent, or by flagging terms of high value when exhaustive synonym searching is unfeasible or inadvisable.

Search limits

The application of limits to search strategies will reflect the eligibility criteria in the review protocol. Typically, English language limits, date limits, and the exclusion of conference abstracts and animal studies are usually done as a matter of routine.

Search filters

A search filter is a string of search terms with known (validated) performance. When a particular study design is required for a review question, relevant search filters are usually applied to literature search strategies.

Other search filters relating to age, setting, geography, and health inequalities are also applied as relevant. The most comprehensive list of available search filters is the search filter resource of the InterTASC Information Specialists' SubGroup. This resource also includes critical appraisal tools, which are used for filter selection.

Economics-related filters

A variety of search filters of relevance to cost effectiveness are available. These include filters for economic evaluations, quality of life data, and cost-utilities data. It may be necessary to use more than 1 filter to identify relevant data. In addition, it may be appropriate to add geographic search filters, such as those for the UK or Organisation for Economic Co-operation and Development (OECD) countries, to retrieve economic studies relevant to the UK or OECD (Ayiku et al. 2017, 2019, 2021).

Use of machine learning-based classifiers

Machine learning-based classification software has been developed for some study types (for example the Cochrane RCT classifier, Thomas et al. 2020). These classifiers apply a probability weighting to each bibliographical reference within a set of search results. The weighting relates to the reference's likelihood to be a particular study type, based on a model created from analysis of known, relevant papers. The weightings can then be used to either order references for screening or be used with a fixed cut-off value to divide a list

of references into those more likely to be included, and those that can be excluded without manual screening.

We support the use of machine classifiers if their performance characteristics are known, and if they improve efficiency in the search and screening process. However, caution is needed when using classifiers, because they may not be as effective if used on data that is different to the type of data for which they were originally developed. For example, the Cochrane RCT classifier is reported to have over 99% recall for health studies but showed "unacceptably low" recall for educational research (Stansfield et al. 2022).

Priority screening, a type of machine classifier that orders references for manual sifting based on previous sifting decisions, is considered in the chapter on reviewing evidence.

Additional search techniques

Additional search techniques are used alongside database searching when it is known, or reasonably likely, that relevant evidence is not indexed in bibliographic databases, or when it will be difficult to retrieve relevant evidence from databases in a way that adequately balances recall and precision. Additional search techniques include forward and backward citation searching, journal hand-searches and contacting experts and stakeholders.

Existing reviews may provide an additional source of primary studies, with reference lists being used as an indirect method of identifying primary research.

Various tools, including Citationchaser and Web of Science, are available to speed up the process of citation searching. These may not be as comprehensive as manual reference list checking (due to limitations of the underlying data sources), but the trade-off in terms of speed is generally acceptable.

All search techniques should follow the same principles of transparency, rigour and reproducibility as other search methods.

If possible, additional search techniques should be considered at the outset and documented in the search protocol. They should also be documented in the supporting appendices for the final evidence review.

5.4 Health inequalities and equality and diversity

All searches aim to be inclusive. This may mean not specifying any population groups.

Searches should avoid inadvertently excluding relevant groups. For example, if the population group is older people, a search for older people should pick up subpopulations such as disabled older people.

Additional search strategies may be needed to target evidence about people with protected characteristics or people experiencing or at risk from other inequalities.

Searches may need to be developed iteratively to ensure coverage of the health inequalities issues or evidence on the impacts of an intervention on equality.

Appropriate terminology for the search should be used, considering how language has evolved.

5.5 Quality assurance

Quality assuring the literature search is an important step in developing guideline recommendations. Studies have shown that errors do occur.

For each search (including economic searches), the initial MEDLINE search strategy is quality assured by a second information specialist. A standardised checklist, based on the PRESS peer review of electronic search strategies: 2015 guideline statement, is used to ensure clarity and consistency when quality assuring search strategies.

The information specialist carrying out the quality assurance process also considers how appropriate the overall search approach is to the parameters of the evidence review (for example, the time available to carry out the review). The quality assurance comments are recorded and the information specialist who conducted the search should respond to the comments and revise the search strategy as needed.

Search strategy translations across the remaining databases are also checked by a second information specialist to ensure that the strategies have been adapted appropriately, in accordance with the interfaces and search functionality of the sources used.

5.6 Documenting the search

Details of the evidence search are included as appendices to the individual evidence reviews. They are published for consultation alongside the draft evidence review and included in the final version.

Records are kept of the searches undertaken during guideline recommendation development for all review questions to ensure that the process for identifying the evidence is transparent and reproducible.

We use the PRISMA-S: an extension to the PRISMA statement for reporting literature searches in systematic reviews to inform search reporting. The search documentation is an audit trail that allows the reader to understand both the technical aspect of what was done (such as which sources were searched; what platform was used and on what date; any deviations from the original search protocol) and the underlying rationale for the search approach where this may not be immediately apparent.

Documenting the search begins with creating the search protocol (see the section on search protocols). If using an iterative or emergent stepped approach, initial search strategies, key decision points and the reasons for subsequent search steps are clearly documented in the search protocol and final evidence review. When using a proprietary search engine such as Google, whose underlying algorithm adapts to different users, the search is reported in a way that should allow the reader to understand what was done.

5.7 Re-running searches

Searches undertaken to identify evidence for each review question (including economics searches) may be re-run before consultation or before publication. For example, searches are re-run if the evidence changes quickly, there is reason to believe that substantial new evidence exists, or the development time is longer than usual.

A decision to re-run searches is taken by the development team and staff with responsibility for quality assurance.

If undertaken, searches are re-run at least 6 to 8 weeks before the final committee meeting before consultation.

If evidence is identified after the last cut-off date for searching but before publication, a

judgement on its impact is made by the development team and staff with responsibility for quality assurance. In exceptional circumstances, this evidence can be considered if its impact is judged as potentially substantial.

5.8 Calls for evidence from stakeholders

In some topic areas or for some review questions, staff with responsibility for quality assurance, the development team or the committee may believe that there is relevant evidence in addition to that identified by the searches. In these situations, the development team may invite stakeholders, and possibly also other relevant organisations or individuals with a significant role or interest (see expert witnesses in the section on other attendees at committee meetings in the chapter on decision-making committees), to submit evidence. A call for evidence is issued directly to registered stakeholders on the NICE website. Examples and details of process are included in the appendix on call for evidence and expert witnesses. Confidential information should be kept to an absolute minimum.

5.9 References and further reading

Ayiku L, Levay P, Hudson T et al. (2017) The medline UK filter: development and validation of a geographic search filter to retrieve research about the UK from OVID medline. Health Information and Libraries Journal 34(3): 200–216

Ayiku L, Levay P, Hudson T et al. (2019) The Embase UK filter: validation of a geographic search filter to retrieve research about the UK from OVID Embase. Health Information and Libraries Journal 36(2): 121–133

Ayiku L, Hudson T, Williams C et al. (2021) The NICE OECD countries' geographic search filters: Part 2-validation of the MEDLINE and Embase (Ovid) filters. Journal of the Medical Library Association 109(4): 583–9

Booth A, Briscoe S, Wright JM (2020) The "realist search": a systematic review of current practice and reporting. Research Synthesis Methods 11: 14–35

Canadian Agency for Drugs and Technologies in Health (2019) Grey Matters: a practical tool for searching health-related grey literature [online; accessed 24 July 2023]

Glanville J, Lefebvre C, Wright K (editors) (2008, updated 2017) The InterTASC Information Specialists' Subgroup Search Filters Resource [online; accessed 24 July 2023]

Kaltenthaler E, Tappenden P, Paisley S (2011) NICE DSU Technical support document 13: identifying and reviewing evidence to inform the conceptualisation and population of cost-effectiveness models [online; accessed 24 July 2023]

Kugley S, Wade A, Thomas J et al. (2017) Searching for studies: a guide to information retrieval for Campbell systematic reviews. Oslo: The Campbell Collaboration

Lefebvre C, Glanville J, Briscoe S et al. Chapter 4: Searching for and selecting studies. In: Higgins JPT, Thomas J, Cumpston M et al. (editors). Cochrane Handbook for Systematic Reviews of Interventions version 6.2 (updated February 2021). Cochrane, 2021

McGowan J, Sampson M, Salzwedel DM et al. (2016) PRESS Peer Review of Electronic Search Strategies: 2015 guideline statement. Journal of Clinical Epidemiology 75: 40–6

National Institute for Health and Care Excellence (2022) NICE real-world evidence framework [online; accessed 24 July 2023]

Paisley S (2016) Identification of key parameters in decision-analytic models of cost-effectiveness: a description of sources and a recommended minimum search requirement. Pharmacoeconomics 34: 597–8

Rethlefsen M, Kirtley S, Waffenschmidt S et al. (2021) PRISMA-S: an extension to the PRISMA statement for reporting literature searches in systematic reviews. Systematic Reviews 10: 39

Stansfield C, Stokes G, Thoman J (2022) Applying machine classifiers to update searches: analysis from two case studies. Research Synthesis Methods 13: 121–33

Summarized research for Information Retrieval in HTA (SuRe Info) [online; accessed 24 July 2023]

6 Reviewing evidence

Reviewing evidence is an explicit, systematic and transparent process that can be applied to both quantitative (experimental and observational) and qualitative evidence (see the chapter on developing review questions and planning the evidence review). The key aim of any review is to provide a summary of the relevant evidence to ensure that the committee can make fully informed decisions about its recommendations. This chapter describes how evidence is reviewed in the development of guidelines.

Evidence reviews for NICE guidelines summarise the evidence and its limitations so that the committee can interpret the evidence and make appropriate recommendations, even where there is uncertainty.

Most of the evidence reviews for NICE guidelines will be presenting syntheses of evidence from systematic literature searches for primary research studies. Evidence identified during these literature searches and from other sources (see the chapter on identifying the evidence: literature searching and evidence submission) should be reviewed against the review protocol to identify the most appropriate information to answer the review questions. The evidence review process used to inform guidelines must be explicit and transparent, and involves 8 main steps:

- writing the review protocol (see the section on planning the evidence review in the chapter on developing review questions and planning the evidence review)
- identifying and selecting relevant evidence (including a list of excluded studies with reasons for exclusion)
- critical appraisal (assessing the study design and its methods)
- extracting relevant data
- synthesising the results (including statistical analyses such as meta-analysis)
- assessing quality and certainty in the evidence
- interpreting the results
- considering health inequalities.

Any substantial deviations from these steps need to be agreed, in advance, with staff with responsibility for quality assurance. Additional considerations for reviews using alternative methods not based primarily on literature reviews of primary studies (such as formal consensus methods, adapting recommendations from other guidelines or primary analyses of real-world data) are discussed in the section on presenting evidence for reviews other than reviews of primary studies.

For all evidence reviews and data synthesis, it is important that the method used to report and evaluate the evidence is easy to follow. It should be written up in clear English and any analytical decisions should be clearly justified.

Updating previous NICE reviews

In many cases, the evidence reviews will be an update of a previous review we've done on the same or a similar topic, to include more recently published evidence. In these cases, a judgement should be made on what elements of the previous review can be reused, and which need to be redone, based on the level of similarity between the original and new review questions, protocols and methods. Examples of elements that can be considered for reuse include:

- literature searches and literature search results
- evidence tables for included studies
- critical appraisal of included studies
- data extraction and meta-analysis
- previously identified information on equalities and health inequalities.

6.1 Identifying and selecting relevant evidence

The process of selecting relevant evidence is common to all evidence reviews based on systematic literature searches; the other steps are discussed in relation to the main types of review question. The same rigour should be applied to reviewing all data, whether fully or partially published studies or unpublished data supplied by stakeholders. Care should be taken to identify and remove multiple reports of the same study to prevent double-counting.

Published studies

Titles and abstracts of the retrieved citations should be screened against the inclusion criteria defined in the review protocol, and those that do not meet these should be excluded. A percentage should be screened independently by 2 reviewers (that is, titles and abstracts should be double-screened). The percentage of records to be double-screened for each review should be specified in the review protocol.

If reviewers disagree about a study's relevance, this should be resolved by discussion or by recourse to a third reviewer. If, after discussion, there is still doubt about whether or not the study meets the inclusion criteria, it should be retained. If there are concerns about the level of disagreement between reviewers, the reasons should be explored, and a course of action agreed to ensure a rigorous selection process. A further proportion of studies should then be double-screened to validate this new process until appropriate agreement is achieved.

Once the screening of titles and abstracts is complete, full versions of the selected studies should be obtained for assessment. As with title and abstract screening, a percentage of full studies should be checked independently by 2 reviewers, with any differences being resolved and additional studies being assessed by multiple reviewers if sufficient agreement is not achieved. Studies that fail to meet the inclusion criteria once the full version has been checked should be excluded at this stage.

The study selection process should be clearly documented and include full details of the inclusion and exclusion criteria. A flow chart should be used to summarise the number of papers included and excluded at each stage and this should be presented in the evidence review document (see the PRISMA statement). Each study excluded after checking the full version should be listed, along with the reason for its exclusion. Reasons for study exclusion need to be sufficiently detailed for people to be able to understand the reason without needing to read the original paper (for example, avoid stating only that 'the study population did not meet that specified in the review protocol', but also include why it did not match the protocol population).

Priority screening

Priority screening refers to any technique that uses a machine learning algorithm to enhance the efficiency of screening. Usually, this involves taking information on previously included or excluded papers, and using this to order the unscreened papers from those

most likely to be included to those least likely. This can be used to identify a higher proportion of relevant papers earlier in the screening process, or to set a cut-off for manual screening, beyond which it is unlikely that additional relevant studies will be identified.

There is currently no published guidance on setting thresholds for stopping screening where priority screening has been used. Any methods used should be documented in the review protocol and agreed in advance with the team with responsibility for quality assurance. Any thresholds set should, at minimum, consider the following:

- the number of references identified so far through the search, and how this identification rate has changed over the review (for example, how many candidate papers were found in each 1,000 screened)
- the overall number of studies expected, which may be based on a previous version of the guideline (if it is an update), published systematic reviews, or the experience of the guideline committee
- the ratio of relevant/irrelevant records found at the random sampling stage (if undertaken) before priority screening.

The actual thresholds used for each review question should be clearly documented, either in the guideline methods chapter or in the evidence review documents. Examples of how this has been implemented can be found in NICE's guidelines on autism spectrum disorders in under 19s and prostate cancer.

Ensuring relevant records are not missed

Regardless of the level of double-screening, and whether or not priority screening was used, additional checks should always be made to reduce the risk that relevant studies are not identified. These should include, at minimum:

- checking reference lists of identified systematic reviews, even if these reviews are not used as a source of primary data
- checking with the guideline committee that they are not aware of any relevant studies that have been missed
- looking for published papers associated with any key trial registry entries or published protocols that have been identified.

It may be useful to test the sensitivity of the search by checking that it picks up known studies of relevance.

Conference abstracts

Conference abstracts seldom contain enough information to allow confident judgements about the quality and results of a study. It can be difficult to trace the original studies or additional data, and the information found may not always be useful. Also, good-quality studies will often publish full text papers after the conference abstract, and these will be identified by routine searches. Conference abstracts should therefore not routinely be included in the search strategy and review, unless there are good reasons for doing so. If a decision is made to include conference abstracts for a particular review, the justification for doing so should be clearly documented in the review protocol. If conference abstracts are searched for, the investigators may be contacted if additional information is needed to complete the assessment for inclusion.

National policy, legislation and medicines safety advice

Relevant national policy, legislation or medicines safety advice may be identified in the literature search and used to inform guidelines (such as drug safety updates from the Medicines and Healthcare products Regulatory Agency [MHRA]). This evidence does not need critical appraisal in the same way as other evidence, given the nature of the source. National policy, legislation or medicines safety advice can be quoted verbatim as evidence (for example, the Health and Social Care Act [2012]), where needed, and a summary of any relevant medicines safety advice identified should be included in the evidence review document.

Unpublished data and studies in progress

Any unpublished data should be quality assessed in the same way as published studies (see the section on assessing evidence: critical appraisal, analysis, and certainty in the findings). If additional information is needed to complete the quality assessment, the investigators may be contacted. Similarly, if data from in-progress studies are included, they should be quality assessed in the same way as published studies. Confidential information should be kept to a minimum, and a structured abstract of the study must be made available for public disclosure during consultation on the guideline. Additional considerations for reviews using primary analyses of real-world data are discussed in the section on presenting evidence for reviews other than reviews of primary studies.

Grey literature

Grey literature may be quality assessed in the same way as published literature, although because of its nature, such an assessment may be more difficult. Consideration should therefore be given to the elements of quality that are most likely to be important (for example, elements of the study methodology that are less clearly described than in a published article, because of the lack of need to go through the peer-review process, or conflicts of interest in the study).

6.2 Assessing evidence: critical appraisal, analysis, and certainty in the findings

Introduction

Assessing the quality of the evidence for a review question is critical. It requires a systematic process of assessing both the appropriateness of the study design and the methods of the study (critical appraisal) as well as the certainty of the findings (using an approach, such as GRADE).

Options for assessing the quality of the evidence should be considered by the development team. The chosen approach should be discussed and agreed with staff with responsibility for quality assurance, where the approach deviates from the standard (described in critical appraisal of individual studies). The agreed approach should be documented in the review protocol (see the appendix on review protocol templates) together with the reasons for the choice. If additional information is needed to complete the data extraction or quality assessment, study investigators may be contacted, although this is not something that is done routinely.

Critical appraisal of individual studies

Every study should be appraised using a checklist appropriate for the study design (see the appendix on appraisal checklists, evidence tables, GRADE and economic profiles). If a checklist other than those listed is needed, or the one recommended as the preferred option is not used, the planned approach should be discussed and agreed with staff with responsibility for quality assurance and documented in the review protocol.

The ROBINS-I checklist is currently only validated and recommended for use with non-

randomised controlled trials and cohort studies. However, there may be situations where a mix of non-randomised study types is included within a review. It can then be helpful to use this checklist across all included study types to maintain consistency of assessment. If this is done, additional care should be taken to ensure all relevant risks of bias for study designs for which ROBINS-I is not currently validated (such as case–control studies) are assessed.

In some evidence reviews, it may be possible to identify particular risk of bias criteria that are likely to be the most important indicators of biases for the review question (for example, conflicts of interest or study funding, if it is an area where there is known to be concern about the sponsorship of studies). If any such criteria are identified, these should then be used to guide decisions about the overall risk of bias of each individual study.

Sometimes, a decision might be made to exclude certain studies at particularly high risk of bias, or to explore any impact of bias through sensitivity analysis. If so, the approach should be specified in the review protocol and agreed with staff with responsibility for quality assurance.

Criteria relating to key areas of bias may also be useful when summarising and presenting the evidence (see the section on summarising evidence). Topic-specific input (for example, from committee members) may be needed to identify the most appropriate criteria to define subgroup analyses, or to define inclusion in a review, for example, the minimum biopsy protocol for identifying the relevant population in cancer studies.

For each criterion that might be explored in sensitivity analysis, the decision on whether it has been met or not (for example, which population subgroup the study has been categorised as), and the information used to arrive at the decision (for example, the study inclusion criteria, or the actual population recruited into the study), should be recorded in a standard template for inclusion in an evidence table (see the appendix on appraisal checklists, evidence tables, GRADE and economic profiles).

Each study included in an evidence review should be critically appraised by 1 reviewer and a proportion of these checked by another reviewer. Any differences in critical appraisal should be resolved by discussion or involving a third reviewer.

Data extraction

Study characteristics should be extracted to a standard template for inclusion in an

evidence table (see the appendix on appraisal checklists, evidence tables, GRADE and economic profiles). Care should be taken to ensure that newly identified studies are cross-checked against existing studies to avoid double-counting. This is particularly important where there may be multiple reports of the same study.

If complex data extraction is done for a review question (for example, situations where a large number of transformations or adjustments are made to the raw data from the included studies), data extraction should be checked by a second reviewer to avoid data errors, which are time-consuming to fix. This may be more common in reviews using more complex analysis methods (for example, network meta-analyses or meta-regressions) but decisions around dual data extraction should be based on the complexity of the extraction, not the complexity of the analysis.

Analysing and presenting results for studies on the effectiveness of interventions

Meta-analysis may be appropriate if treatment estimates of the same outcome from more than 1 study are available. Recognised approaches to meta-analysis should be used, as described in the handbook from Cochrane, in Higgins et al. (2021) and documents developed by the NICE Guidelines Technical Support Unit.

There are several ways of summarising and illustrating the strength and direction of quantitative evidence about the effectiveness of an intervention, even if a meta-analysis is not done. Forest plots can be used to show effect estimates and confidence intervals for each study (when available, or when it is possible to calculate them). They can also be used to provide a graphical representation when it is not appropriate to do a meta-analysis and present a pooled estimate. However, the homogeneity of the outcomes and measures in the studies needs to be carefully considered: a forest plot needs data derived from the same (or justifiably similar) population, interventions, outcomes and measures.

Head-to-head data that compares the effectiveness of interventions is useful for a comparison between 2 active management options. A network meta-analysis (NMA) is a method that can include trials that compare the interventions of interest head-to-head and also trials that allow indirect comparisons via other interventions.

The same principles of good practice for evidence reviews and meta-analyses should be applied when conducting network meta-analyses. The reasons for identifying and selecting the randomised controlled trials (RCTs) should be explained. This includes the

reasons for selecting the treatment comparisons, and whether any interventions that are not being considered as options for recommendations will be included within the network to allow for indirect comparisons between interventions of interest. The methods of synthesis should be described clearly either in the methods section of the evidence review document or the guideline methods chapter.

When multiple competing options are being appraised, network meta-analysis is the preferred approach to use, and should be considered in such cases. The data from individual trials should also be documented (usually as an appendix). If there is doubt about the inclusion of particular trials (for example, because of concerns about limitations or applicability), a sensitivity analysis in which these trials are excluded may also be presented. The level of consistency between the direct and indirect evidence on the interventions should be reported, including consideration of model fit and comparison statistics such as the total residual deviance, and the deviance information criterion (DIC). Results of any further inconsistency tests done, such as deviance plots or those based on node-splitting, should also be reported.

In addition to the inconsistency checks described above, which compare the direct and indirect evidence within a network meta-analysis model, results from direct comparisons may also be presented for comparison with the results from a network meta-analysis (thus comparing the direct and overall network meta-analysis results to aid validity checks and interpretation, rather than direct and indirect to check consistency). These may be the results from the direct evidence within the network meta-analysis, or from direct pairwise comparisons done outside the network meta-analysis, depending on which is considered more informative.

When evidence is combined using network meta-analyses, trial randomisation should typically be preserved. If this is not appropriate, the planned approach should be discussed and agreed with staff with responsibility for quality assurance. A comparison of the results from single treatment arms from different RCTs is not acceptable unless the data are treated as observational and analysed as such.

Further information on complex methods for evidence synthesis is provided by the documents developed by the NICE Guidelines Technical Support Unit. The methods described in these documents should be used as the basis for analysis, and any deviations from these methods clearly described and justified, and agreed with staff who have responsibility for quality assurance.

To promote transparency of health research reporting (as endorsed by the EQUATOR network), evidence from a network meta-analysis should usually be reported according to the criteria in the modified PRISMA-NMA checklist in the appendix on network meta-analysis reporting standards.

Evidence from a network meta-analysis can be presented in a variety of ways. The network should be presented diagrammatically with the available treatment comparisons clearly identified, and show the number of trials in each comparison. Further information on how to present the results of network meta-analyses is provided by the documents developed by the NICE Guidelines Technical Support Unit.

There is no NICE-endorsed approach for assessing the quality or certainty of outputs derived from network meta-analysis. At a minimum, a narrative description of the confidence in the results of the network meta-analysis should be presented, considering all the areas in a standard GRADE profile (risk of bias, indirectness, inconsistency and imprecision). Several other approaches have been suggested in the literature that may be relevant in particular circumstances (Phillippo et al. 2019, Phillippo et al. 2017, Caldwell et al. 2016, Purhan et al. 2014, Salanti et al. 2014). The approach to assessing confidence in results should take into account the particular questions the network meta-analysis is trying to address. For example, the approach to imprecision may be different if a network meta-analysis is trying to identify the single most effective treatment, compared to creating a ranking of all possible treatments.

Dealing with complex interventions

Analysing quantitative evidence on complex interventions may involve considering factors other than effectiveness. This includes:

- whether there are particular circumstances when the interventions work
- is there interaction, synergy or mediation between intervention components
- which factors impact on implementation
- is the intervention feasible and acceptable in different contexts
- how might this enhance or reduce the interventions' effect in different circumstances (see sections 17.2 and 17.5 in the Cochrane Handbook for Systematic Reviews of Interventions).

Different analytical approaches are relevant to different types of complexity and question (see table 1 in Higgins et al. 2019). The appropriate choice of technique will depend on the review question, available evidence, time needed to do the approach and likely impact on guideline recommendations. The approach should be discussed and agreed with staff who have responsibility for quality assurance.

Further information on complex methods for evidence synthesis is provided by the documents developed by the NICE Guidelines Technical Support Unit and NICE's Decision Support Unit.

Additional information is available from:

- Agency for Healthcare Research and Quality (AHRQ) series on complex intervention systematic reviews (2017)
- Viswanathan et al. (2017) AHRQ series: paper 4
- BMJ series on complex health interventions in complex systems: concepts and methods for evidence-informed health decisions (Higgins, 2019)
- chapter 17 of the Cochrane Handbook for Systematic Reviews of Interventions.

Analysing and presenting results of studies of diagnostic test accuracy

Information on methods of presenting and synthesising results from studies of diagnostic test accuracy is available in the Cochrane Handbook for Systematic Reviews of Interventions. When meta-analyses of paired accuracy measures (such as sensitivity and specificity) are done, bivariate analysis should be used where possible, to preserve correlations between outcomes. Univariate analyses can still be used if there are insufficient studies for a bivariate analysis.

Meta-analyses should not normally be done on positive and negative predictive values, unless the analysis takes account of differences in prevalence. Instead, analyses can be done on sensitivity and specificity and these results applied to separate prevalence estimates to obtain positive and negative predictive values, if these are outcomes specified in the review protocol.

If meta-analysis is not possible or appropriate (for example, if the differences between populations, references standard or index test thresholds are too large), there should be a

narrative summary of the results that were considered most important for the review question.

Analysing and presenting results of studies of prognosis, or prediction models for a diagnosis or prognosis

There is currently no consensus on approaches for synthesising evidence from studies on prognosis, or prediction models for diagnosis or prognosis. The approach chosen should be based on the types of data included (for example, prognostic accuracy data, prediction models, or associative studies presenting odds ratios or hazard ratios). For prognostic accuracy data, the same approach for synthesis can be taken as with diagnostic accuracy data, with the addition of the need to consider length of follow-up as part of the analysis. When considering meta-analysis, reviewers should consider how similar the prognostic factors or predictors and confounding factors are across all studies reporting the same outcome measure. It is important to explore whether all likely confounding factors have been accounted for, and whether the metrics used to measure exposure (or outcome) are universal. When studies cannot be pooled, results should be presented consistently across studies. For more information on prognostic reviews, see Collins 2015 and Moons 2015.

Analysing, synthesising and presenting results of qualitative evidence

Qualitative evidence occurs in many forms and formats and so different methods may be used for synthesis and presentation (such as those described by the Cochrane Qualitative & Implementation Methods Group).

Qualitative evidence should be synthesised and then summarised using GRADE-CERQual (see GRADE-CERQual Implementation series). If synthesis of the evidence is not appropriate, a narrative summary may be adequate; this should be agreed with staff with responsibility for quality assurance. The approach used may depend on the volume of the evidence. If the qualitative evidence is extensive, then a recognised method of synthesis is preferable (normally aggregative, thematic or framework synthesis type approaches). If the evidence is disparate and sparse, a narrative summary may be appropriate.

The simplest approach to synthesise qualitative data in a meaningful way is to group the findings in the evidence tables (comprising of 'first order' participant quotes and participant observations as well as 'second order' interpretations by study authors). Then, to write third-order interpretations based on the reviewers' interpretations of the first and second-order constructs synthesised across studies. These third-order interpretations will

become themes and subthemes or 'review findings'. This synthesis can be carried out if enough data are found, and the papers and research reports cover the same (or similar) context or use similar methods. These should be relevant to the review questions and could, for example, include intervention, age, population or setting.

Synthesis can be carried out in several ways (as noted above), and each may be appropriate depending on the question type, and the evidence identified. Papers reporting on the same findings can be grouped together to compare and contrast themes, focusing not just on consistency but also on any differences. The narrative should be based on these themes.

A more complex but useful approach is 'conceptual mapping' (see Johnson et al. 2000). This involves identifying the key themes and concepts across all the evidence tables and grouping them into first level (major), second level (associated) and third level (subthemes) themes. Results are presented in schematic form as a conceptual diagram and the narrative is based on the structure of the diagram.

Integrating and presenting results of mixed methods reviews

If a mixed methods approach has been identified as needed (see the chapter on developing review questions and planning the evidence review), then the approach to integration needs consideration. Integration refers to A) how quantitative and qualitative evidence are combined following separate synthesis (convergent-segregated) or, B) how quantitative and qualitative data that have been transformed are merged (convergent-integrated).

- A) The convergent-segregated approach consists of doing separate quantitative and qualitative syntheses (as usual), followed by integration of the results derived from each of the syntheses. Integrating the quantitative and qualitative synthesised findings gives a greater depth of understanding of the phenomena of interest compared to doing 2 separate component syntheses without formally linking the 2 sets of evidence.
- B) All qualitative evidence from a convergent-segregated mixed methods review should be synthesised and then summarised using GRADE-CERQual. If appropriate, all quantitative data (for example, for intervention studies) should be presented using GRADE. An overall summary of how the quantitative and qualitative evidence are linked should ideally be presented in either matrices or thematic diagrams. It should also be summarised in the review using the approach questions in the section on

integration of quantitative and qualitative evidence to frame the integration evidence summary (JBI manual for evidence synthesis).

Integration of quantitative and qualitative evidence

The integration section should provide a summary that represents the configured analysis of the quantitative and qualitative evidence. This can include matrices, look-up tables or thematic maps, but as a minimum should include statements that address all of the following questions:

- Are the results and findings from individual syntheses supportive or contradictory?
- Does the qualitative evidence explain why the intervention is or is not effective?
- Does the qualitative evidence explain differences in the direction and size of effect across the included quantitative studies?
- Which aspects of the quantitative evidence were or were not explored in the qualitative studies?
- Which aspects of the qualitative evidence were or were not tested in the quantitative studies?

'All of the questions above should be answered, but dependent on the evidence included in the review it is acknowledged that some responses will be more detailed than others' (JBI manual for evidence synthesis).

This should be reported as a summary of the mixed findings after reporting on the effectiveness and qualitative evidence synthesis.

A) The convergent-integrated approach refers to a process of combining extracted data from quantitative studies (including data from the quantitative component of mixed methods studies) and qualitative studies (including data from the qualitative component of mixed methods studies) and involves data transformation.

B) The convergent-segregated approach is the standard approach to adopt in most of our mixed methods reviews. If convergent-segregated is not the planned approach, data transformation methods and outcome reporting should be discussed and agreed with staff who have responsibility for quality assurance and documented in the review protocol.

Certainty or confidence in the findings of analysis

Once critical appraisal of the studies and data analysis are complete, the certainty or confidence in the findings should be presented (for individual or synthesised studies) at outcome level using GRADE or GRADE-CERQual. Although GRADE has not been formally validated for all quantitative review types (such as prognostic reviews), GRADE principles can be applied and adapted to other types of questions. Any substantial changes made by the development team to GRADE should be agreed with staff with responsibility for quality assurance before use.

If using GRADE or GRADE-CERQual is not appropriate, the planned approach should be discussed and agreed with staff with responsibility for quality assurance. It should be documented in the review protocol (see the appendix on review protocol templates) together with the reasons for the choice.

Certainty or confidence in the findings by outcome

Before starting an evidence review, the outcomes of interest which are important to people using services and the public for the purpose of decision making should be identified. The reasons for prioritising outcomes should be stated in the evidence review document. This should be done before starting the evidence review and clearly separated from discussion of the evidence, because there is potential to introduce bias if outcomes are selected when the results are known. An example of this would be choosing only outcomes for which there were statistically significant results.

The committee discussion section should also explain how the importance of outcomes was considered when discussing the evidence. For example, the committee may want to define prioritised outcomes into 'critical' and 'important'. Alternatively, they may think that all prioritised outcomes are crucial for decision making. In this case, there will be no distinction between 'critical' or 'important' for all prioritised outcomes. The impact of this on the final recommendations should be clear.

GRADE and GRADE-CERQual assess the certainty or confidence in the review findings by looking at features of the evidence found for each outcome or theme. GRADE is summarised in box 6.1, and GRADE-CERQual in box 6.2.

Box 6.1 GRADE approach to assessing the certainty of evidence for intervention studies

GRADE assesses the following features for the evidence found for each outcome:

- study limitations (risk of bias) – the internal validity of the evidence
- inconsistency – the heterogeneity or variability in the estimates of treatment effect across studies
- indirectness – the extent of differences between the population, intervention, comparator for the intervention and outcome of interest in the studies from that in the review protocol
- imprecision – the level of certainly in the effect estimate
- other considerations – publication bias, the degree of selective publication of studies.

In a standard GRADE approach, the certainty or confidence of evidence is classified as high, moderate, low or very low. In the context of NICE guidelines, it can be interpreted as follows:

- **High** – further research is very unlikely to change our recommendation.
- **Moderate** – further research may have an important impact on our confidence in the estimate of effect and may change the strength of our recommendation.
- **Low** – further research is likely to have an important impact on our confidence in the estimate of effect and is likely to change the recommendation.
- **Very low** – any estimate of effect is very uncertain and further research will probably change the recommendation.

Box 6.2 GRADE-CERQual approach to assessing the confidence of evidence for qualitative studies

GRADE-CERQual assesses the following features for the evidence found for each finding:

- methodological limitations – the internal validity of the evidence
- relevance – the extent to which the evidence is applicable to the context in the review question
- coherence – the extent of the similarities and differences within the evidence
- adequacy of data – the extent of richness and quantity of the evidence.

In a standard GRADE-CERQual approach, the certainty or confidence of evidence is classified as high, moderate, low or very low. In the context of NICE guidelines, it can be interpreted as follows:

- **High** – it is highly likely that the review finding is a reasonable representation of the phenomenon of interest.
- **Moderate** – it is likely that the review finding is a reasonable representation of the phenomenon of interest.
- **Low** – it is possible that the review finding is a reasonable representation of the phenomenon of interest.
- **Very low** – it is unclear whether the review finding is a reasonable representation of the phenomenon of interest.

The approach we take differs from the standard GRADE and GRADE-CERQual system in 2 ways:

- it also integrates a review of the quality of cost-effectiveness studies (see the chapter on incorporating economic evaluation)
- it does not use 'overall summary' labels for the quality of the evidence across all outcomes, or for the strength of a recommendation, but uses the wording of recommendations to reflect the strength of the evidence (see the chapter on interpreting the evidence and writing the guideline).

GRADE or GRADE-CERQual tables summarise the certainty in the evidence and data for each critical and each important outcome or theme and include a limited description of the certainty in the evidence. GRADE or GRADE-CERQual tables should be available (in an appendix) for each review question.

For mixed methods findings there is no recognised approach to combining the certainty of evidence from GRADE and GRADE-CERQual. The certainty and confidence ratings should be reported for both evidence types within the evidence summary of integrated findings and their impact on decision making described in the relevant section of the review.

Alternative approaches to assessing imprecision in GRADE

For information on assessing imprecision the standard GRADE approach can be used. If this approach is not used, the approach should be agreed with staff who have responsibility for quality assurance.

6.3 Equality and diversity considerations

Our equality and diversity duties are expressed in a single public sector equality duty ('the equality duty', see the section on key principles for developing NICE guideline recommendations in the introduction chapter). The equality duty supports good decision making by encouraging public bodies to understand how different people will be affected by their activities. As much of our work involves developing advice for others on what to do, this includes thinking about how people will be affected by our recommendations when they are implemented (for example, by health and social care practitioners).

6.4 Health inequalities

In addition to meeting our legal obligations, we are committed to going beyond compliance, particularly in terms of tackling health inequalities. Specifically, we consider that we should also take account of the 4 dimensions of health inequalities – socioeconomic status and deprivation, protected characteristics (defined in the Equality Act 2010), inclusion health groups (such as people experiencing homelessness and young people leaving care), and geography. Wherever possible, our guidance aims to reduce and not increase identified health inequalities.

Ensuring inclusivity of the evidence review criteria

Any equality criteria specified in the review protocol should be included in the evidence tables. At the data extraction stage, reviewers should refer to the health inequalities framework criteria (including age, gender/sex, sexual orientation, gender reassignment, disability, ethnicity, religion, place of residence, occupation, education, socioeconomic position and social capital; Gough et al. 2012) and any other relevant protected characteristics, and record these where reported, if specified in the review protocol. See the section on reducing health inequalities in the introduction chapter. Review inclusion and exclusion criteria should also take the relevant groups into account, as specified in the review protocol.

Equalities and health inequalities should be considered during the drafting of the evidence reviews, including any issues documented in the equality and health inequalities assessment. Equality and health inequality considerations should be included in the data extraction process and should be recorded in the committee discussion section. Equalities and health inequalities are also considered during surveillance and updating. See chapters on ensuring that published guidelines are current and accurate and updating guideline recommendations for more information.

6.5 Summarising evidence

Presenting evidence

The following sections should be included in the evidence review document:

- an introduction to the evidence review
- a description of the studies or other evidence identified, in either table or narrative format
- evidence tables (usually presented in an appendix)
- full GRADE or GRADE-CERQual profiles (in an appendix)
- evidence summaries (of the results or conclusions of the evidence)
- an overall summary of merged quantitative and qualitative evidence (either using matrices or thematic diagrams) and the integration questions for mixed methods

reviews

- results from other analysis of evidence, such as forest plots, area under the curve graphs, network meta-analysis (usually presented in an appendix; see the appendix on network meta-analysis reporting standards).

The evidence should usually be presented separately for each review question; however, alternative methods of presentation may be needed for some evidence reviews (for example, where review questions are closely linked and need to be interpreted together).

Any substantial deviations in presentation need to be agreed, in advance, with staff with responsibility for quality assurance.

Describing the included evidence

A description of the evidence identified should be produced. The content of this will depend on the type of question and the type of evidence. It should also identify and describe any gaps in the evidence, and cover at minimum:

- the volume of information for the review question(s), that is, the number of studies identified, included, and excluded (with a link to a PRISMA selection flowchart, in an appendix)
- the study types, populations, interventions, settings or outcomes for each study related to a particular review question.

Evidence tables

Evidence tables help to identify the similarities and differences between studies, including the key characteristics of the study population and interventions or outcome measures.

Data from identified studies are extracted to standard templates for inclusion in evidence tables. The type of data and study information that should be included depends on the type of study and review question, and should be concise and consistently reported.

The types of information that could be included for quantitative studies are:

- bibliography (authors, date)

- study aim, study design (for example, RCT, case–control study) and setting (for example, country)
- funding details (if known)
- population (for example, source and eligibility, and which population subgroup of the protocol the study has been mapped to, if relevant)
- intervention, if applicable (for example, content, who delivers the intervention, duration, method, dose, mode or timing of delivery, and which intervention subgroup of the protocol the study has been mapped to, if relevant)
- comparator, if applicable (for example, content, who delivers the intervention, duration, method, dose, mode or timing of delivery)
- method of allocation to study groups (if applicable)
- outcomes (for example, primary and secondary and whether measures were objective, subjective or otherwise validated, and the timepoint at which these outcomes were measured)
- key findings (for example, effect sizes, confidence intervals, for all relevant outcomes, and where appropriate, other information such as numbers needed to treat and considerations of heterogeneity if summarising a systematic review or meta-analysis)
- inadequately reported data, missing data or if data have been imputed (include method of imputation or if transformation is used)
- overall comments on quality, based on the critical appraisal and what checklist was used to make this assessment. When study details are inadequately reported, or absent, this should be clearly stated.

If data are not being used in any further statistical analysis, or are not reported in GRADE tables, effect sizes (point estimate) with confidence intervals should be reported, or back calculated from the published evidence where possible. If confidence intervals are not reported, exact p values (whether or not significant), with the test from which they were obtained, should be described. When confidence intervals or p values are inadequately reported or not given, this should be stated. Any descriptive statistics (including any mean values and degree of spread such as ranges) indicating the direction of the difference between intervention and comparator should be presented. If no further statistical information is available, this should be clearly stated.

The type of data that could be reported in evidence tables for qualitative studies includes:

- bibliography (authors, date)
- study aim, study design and setting (for example, country)
- funding details (if known)
- population or participants
- theoretical perspective adopted (such as grounded theory)
- key objectives and research questions; methods (including analytical and data collection technique)
- key themes/findings (including quotes from participants that illustrate these themes or findings, if appropriate)
- gaps and limitations
- overall comments on quality, based on the critical appraisal and what checklist was used to make this assessment. When study details are inadequately reported, or absent, this should be clearly stated.

Evidence summaries

Full GRADE or GRADE-CERQual tables that present both the results of the analysis and describe the confidence in the evidence should normally be provided (in an appendix).

Additionally, whether GRADE or GRADE-CERQual are used or not, a summary of the evidence should be included within the evidence review document. This summary can be in any format (narrative, tabular, pictorial) but should contain sufficient detail to explain the key findings of the review without needing to refer to the full results in the appendices.

Evidence summaries are structured and written to help committees formulate recommendations, and stakeholders and users of the guidance to understand the reason why those recommendations were made. They are separate to the committee's interpretation of the evidence, which should be covered in the committee discussion section. They can help to understand:

- whether or not there is sufficient evidence (in terms of strength and applicability) to

form a judgement

- whether (on balance) the evidence demonstrates that an intervention, approach or programme is effective or ineffective, or is inconclusive
- the size of effect and associated measure of uncertainty
- whether the evidence is applicable to people affected by the guideline and contexts covered by the guideline.

Structure and content of evidence summaries

Evidence summaries do not need to repeat every finding from an evidence review, but should contain sufficient information to understand the key findings of the review, including:

- Sufficient descriptions of the interventions, tests or factors being reported on to enable interpretation of the results reported.
- The volume of and confidence in the evidence, as well as the magnitude and direction of effects.
- Key strengths and limitations of the evidence that may not be obvious from overall confidence ratings (for example, the countries evidence came from, if that is expected to have a meaningful impact on the results).
- For findings not showing a meaningful benefit or harm between multiple options, it should be clear whether these have been interpreted as demonstrating equivalence, or simply that it is not possible to tell whether there is a difference or not from the available evidence.
- Any outcomes where evidence was searched for but no or insufficient evidence was found.

These summaries can be done in a variety of formats (for example, evidence statement, narrative summaries, tables) provided they cover the relevant information. 'Vote counting' (merely reporting on the number or proportion of studies showing a particular positive or negative finding) is not an acceptable summary of the evidence.

Context- or topic-specific terms (for example, 'an increase in HIV incidence', 'a reduction in injecting drug use' and 'smoking cessation') may be used. Any such terms should be used

consistently in each review and their definitions reported.

6.6 Presenting evidence for reviews other than reviews of primary studies

The principles described above remain relevant when reporting evidence not based on systematic reviews of primary studies done by NICE. A description of some of these alternative approaches and when they may be appropriate is given in the chapter on developing review questions and planning the evidence review. However, additional factors need to be considered in many of these situations and are described in this section. When reviews have used either multiple options described in this section or an option combined with a systematic review of primary studies, the different approaches should be reported separately according to the appropriate reporting approach outlined in this chapter. A description of how these sources of evidence were either combined or interpreted together by the committee should also be given.

Reporting reviews based on a published systematic review or qualitative evidence synthesis

In some cases, evidence reviews may be based on previously published systematic reviews or qualitative evidence syntheses done outside of NICE, rather than an original review. In such cases, where that review is publicly available, presentation of review content in NICE evidence review documents should be limited to those sections where additional material or analysis has been undertaken. If a published and free to access review has been used with no adaptation, it should be cited in the relevant sections and appendices of the NICE evidence review document and a hyperlink to the original review provided, with no reproduction of the review content. If the review used is not free to access, then the relevant content should be summarised within the guideline.

Examples of additions that may be made to published reviews include adding new data to an out-of-date review, including additional outcomes or subgroups, re-analysing data using different statistical strategies, re-evaluating GRADE quality assessments, and combining separate reviews in a network meta-analysis. If we have updated a review to include additional material or analysis, a link should be provided to the relevant original review with a full citation in line with the NICE style guide on referencing and citations. Only the relevant updated sections should be written up in the NICE evidence review document.

An evidence summary should still be provided in the evidence review, which makes clear which parts of the cited reviews were used as evidence within the guideline, and summarises any changes or additional analyses undertaken, if relevant. When considering the confidence we have in the findings of a published review, both the quality of the overall review (as assessed using the checklists recommended in the appendix on appraisal checklists, evidence tables, GRADE and economic profiles), and the quality of the studies within that review should be taken into account.

Reporting reviews based on a published individual participant data meta-analysis

Evidence reviews based on a published individual patient data (IPD) meta-analysis should follow the same principles as reviews based on other published systematic reviews. Reviewers can make use of the PRISMA-IPD checklist to assess the reporting standards of published IPD analyses, and Wang 2021 includes a checklist that can be used for quality assessment of IPD meta-analyses.

In most cases it is not possible to update an IPD meta-analysis within a guideline, and therefore an approach must be decided if there are additional relevant studies not included within the analysis (for example, additional studies published after the searches in the published review). A number of possible approaches can be followed:

- Only include the IPD meta-analysis in the review, and exclude any additional studies.
- Include the IPD meta-analysis review, and additionally report aggregated results for the studies not included in the IPD analysis.
- Include the IPD meta-analysis review, and additionally report aggregated results for all studies within the review, both those included within the IPD meta-analysis and those not included.

The approach taken should be described and justified within the review. It should take into account the number and proportion of studies not included in the IPD meta-analysis, whether those studies are systematically different to the studies included, and whether the studies not included would be likely to lead to different overall conclusions.

Reporting reviews based on multiple published systematic reviews or qualitative evidence syntheses

Sometimes an evidence review may report the results of multiple systematic reviews, either as a result of a review of reviews being done, or because multiple relevant reviews are otherwise identified. Each review should be reported following the advice in the section on reporting reviews based on a published systematic review or qualitative evidence synthesis.

Additionally, the evidence review should report on any overlaps between the included reviews (for example, where multiple included reviews cover the same intervention or include some of the same studies), or any important differences between the methodologies of the included reviews. How these overlaps or differences were dealt with when assessing evidence and making recommendations should be reported.

Reporting reviews based on formal consensus methods

When formal consensus methods, such as Delphi panels or nominal group technique, are used as a way of generating or interpreting evidence, at minimum the following information should be reported in the evidence review document:

- How the participants involved in the formal consensus exercise were selected.
- How the initial evidence or statements presented as part of the formal consensus exercise were derived.
- The methodology used for the formal consensus exercises, including any thresholds used for retaining or discarding statements.
- The results of each round or iteration of the formal consensus exercise.
- How the results of the formal consensus exercise were then used to inform the recommendations made.

Reporting reviews or using recommendations from previously published guidance from other organisations

If systematic reviews or qualitative evidence syntheses done as part of a published non-NICE guideline are used as evidence within a NICE guideline, those reviews should be

assessed following the advice in the section above on reporting reviews based on a published systematic review or qualitative evidence synthesis. No assessment of other aspects of the guideline is needed, because only the evidence from the reviews is being used, not any other part of the non-NICE guideline.

If parts of the non-NICE guideline other than evidence reviews are used (for example, if the recommendations made are themselves used as evidence, not just the underlying reviews) then the guideline should be assessed for quality using the AGREE II instrument. There is no cut-off point for accepting or rejecting a guideline, and each committee needs to set its own parameters. These should be documented in the methods of the guideline, and the full results of the assessment included in the evidence review document. In addition to the assessment of the quality of the guideline, the following should also be included in the review at a minimum:

- A summary of the content from the non-NICE guideline used to inform the NICE guideline (for example, the recommendations considered).
- A description of the justifications presented in the non-NICE guideline (for example, why those recommendations were made).
- A description of how the NICE committee interpreted that content, including any concerns about quality and applicability, and how it informed their own discussions and recommendations.
- A clear link between which parts of the non-NICE guideline informed the final recommendations in the NICE guideline.

Reporting reviews or using recommendations from previously published NICE guidelines

If systematic reviews or qualitative evidence syntheses done as part of published NICE guidelines are considered relevant and appropriate, they can be used as evidence within a different NICE guideline. These reviews can be included as part of the evidence when:

- the review question in the guideline in development is sufficiently similar to the question addressed in the published guideline
- the evidence is unlikely to have changed significantly since the publication of the related published NICE evidence review.

When evidence reviews from another guideline are used to develop new recommendations, the decision should be made clear in the methods section of the guideline in development, and the committee's independent interpretation and discussion of the evidence should be documented in the discussion section. The evidence reviews from the published guideline (including review protocol, search strategy, evidence tables and full evidence profiles [if available]) should be included in the guideline in development. They then become part of the evidence for the new guideline and are updated as needed in future updates of the guideline.

If parts of a published NICE guideline (or multiple guidelines) other than evidence reviews are used (for example, if recommendations made are themselves used as evidence, not just the underlying reviews) and new recommendations are formulated, the committee's discussion and decision should be documented clearly in the review. This should include areas of agreement and difference with the committee for the published guideline (for example, in terms of key considerations – balance of benefits and harms or costs, and interpretation of the evidence).

The following should be included in the review at a minimum:

- A summary of the content from the published NICE guideline used to inform the guideline in development (for example, the recommendations considered).
- A description of the justifications presented in the published NICE guideline (for example, why those recommendations were made).
- A description of how the committee interpreted that content, including any concerns about applicability, and how it informed their own discussions and recommendations, including how the recommendations from the published NICE guideline were extrapolated to the guideline in development. It is not routinely necessary to do an assessment of the published NICE guideline using the AGREE II instrument. However, in certain circumstances such an assessment may be useful (for example, if it is an older NICE guideline that used different methods to those currently in use), and if an assessment is undertaken the results should be reported in the review.
- A clear link between which parts of the published NICE guideline informed the final recommendations in the guideline in development and why new recommendations were needed (including why the original recommendations could not be adopted without change).

Reporting reviews using primary analysis of real-world data

Reviewers should follow the advice in the NICE real-world evidence framework when reporting primary analyses of real-world data done by NICE. At a minimum, the level of detail provided should match that which would be provided in a published research article. It should also be enough to enable an independent researcher with access to the data to reproduce the study, interpret the results, and to fully understand the strengths and limitations of the study.

More information on what is required and links to relevant reporting tools are provided in the NICE real-world evidence framework.

Reporting reviews using calls for evidence or expert witnesses

If evidence for a review has been obtained using either a call for evidence or an expert witness, follow the reporting advice in the appendix on calls for evidence and expert witnesses.

Reporting reviews using additional consultation or commissioned primary research

If evidence for a review has been obtained using either additional consultation or commissioned primary research, follow the reporting advice in the appendix on approaches to additional consultation and commissioned primary research.

6.7 References and further reading

AGREE Collaboration (2003) Development and validation of an international appraisal instrument for assessing the quality of clinical practice guidelines: the AGREE project. Quality and Safety in Health Care 12: 18–23

Booth A, Lewin S, Glenton C. et al. (2018) Applying GRADE-CERQual to qualitative evidence synthesis findings–paper 7: understanding the potential impacts of dissemination bias. Implementation Sci 13:12

Brouwers M, Kho M, Browman G et al. for the AGREE Next Steps Consortium (2010) AGREE II: advancing guideline development, reporting and evaluation in healthcare. Canadian

Medical Association Journal 182: E839–42

Caldwell D, Ades A, Dias S et al. (2016) A threshold analysis assessed the credibility of conclusions from network meta-analysis. Journal of Clinical Epidemiology 80: 68–76

Caldwell D, Welton N (2016) Approaches for synthesising complex mental health interventions in meta-analysis Evidence-Based Mental Health 19:16

Collins G, Reistma J, Altman D et al. (2015) Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): The TRIPOD Statement. Annals of Internal Medicine 162: 55–63

Colvin C, Garside R, Wainwright M et al. (2018) Applying GRADE-CERQual to qualitative evidence synthesis findings-paper 4: how to assess coherence. Implementation Sci 13:13

Glenton C, Carlsen B, Lewin S et al. (2018) Applying GRADE-CERQual to qualitative evidence synthesis findings—paper 4: how to assess coherence. Implementation Sci 13:14

GRADE working group (2004) Grading quality of evidence and strength of recommendations. British Medical Journal 328: 1490–4

The GRADE series in the Journal of Clinical Epidemiology

Guyatt G, Oxman A, Schünemann H et al. (2011) GRADE guidelines: a new series of articles in the Journal of Clinical Epidemiology. Journal of Clinical Epidemiology 64: 380–2

Higgins J, Thomas J, Chandler J et al., editors (2021) Cochrane Handbook for Systematic Reviews of Interventions, version 6.2

Higgins J, López-López J, Becker B, et al. (2019) Synthesising quantitative evidence in systematic reviews of complex health interventions BMJ Global Health 2019;4:e000858

Johnsen J, Biegel D, Shafran R (2000) Concept mapping in mental health: uses and adaptations. Evaluation and Programme Planning 23: 67–75

Lewin S, Bohren M, Rashidian A et al. (2018) Applying GRADE-CERQual to qualitative evidence synthesis findings—paper 2: how to make an overall CERQual assessment of confidence and create a Summary of Qualitative Findings table. Implementation Sci 13:10

Lewin S, Booth A, Glenton C et al. (2018) Applying GRADE-CERQual to qualitative evidence synthesis findings: introduction to the series. Implementation Sci 13:2

Lizarondo L, Stern C, Carrier J, et al. Chapter 8: Mixed methods systematic reviews. In: Aromataris E, Munn Z (Editors), JBI Manual for Evidence Synthesis. JBI, 2020.

Moons K, Altman D, Reistma J et al. (2015) Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): explanation and elaboration. Annals of Internal Medicine 126: W1–W73

Munthe-Kaas H, Bohren M, Glenton C et al. (2018) Applying GRADE-CERQual to qualitative evidence synthesis findings—paper 3: how to assess methodological limitations. Implementation Sci 13:9

NICE Decision Support Unit (2020) Sources and Synthesis of Evidence: Update to Evidence Synthesis Methods (sheffield.ac.uk) [online; accessed 31 March 2022]

NICE Decision Support Unit Evidence synthesis TSD series [online; accessed 31 August 2018]

Noyes J, Booth A, Lewin S et al. (2018) Applying GRADE-CERQual to qualitative evidence synthesis findings–paper 6: how to assess relevance of the data (nih.gov). Implementation Sci 13:4

Phillippo D, Dias S, Ades A et al. (2017) Sensitivity of treatment recommendations to bias in network meta-analysis. Journal of the Royal Statistical Society; Series A

Phillippo D, Dias S, Welton N et al. (2019) Threshold Analysis as an Alternative to GRADE for Assessing Confidence in Guideline Recommendations Based on Network Meta-analyses. Annals of Internal Medicine 170(8): 538-46

Puhan M, Schünemann H, Murad M et al. (2014) A GRADE working group approach for rating the quality of treatment effect estimates from network meta-analysis. British Medical Journal 349: g5630

Salanti G, Del Giovane C, Chaimani A et al. (2014) Evaluating the quality of evidence from a network meta-analysis. PloS one. 9(7): e99682

Thomas J, O'Mara-Eves A, Brunton G. (2014) Using qualitative comparative analysis (QCA) in systematic reviews of complex interventions: a worked example. Systematic Reviews 3:67

Thomas J, Petticrew M, Noyes J, et al. Chapter 17: Intervention complexity. In: Higgins JPT, Thomas J, Chandler J et al (editors), Cochrane Handbook for Systematic Reviews of Interventions version 6.3 (updated February 2022) [online; accessed 31 March 2022]

Viswanathan M, McPheeters M, Murad M et al. (2017) AHRQ series on complex intervention systematic reviews—paper 4: selecting analytic approaches - Journal of Clinical Epidemiology. J Clin Epidemiol 90:28

Wang H, Chen Y, Lin T, et al. (2021) The methodological quality of individual participant data meta-analysis on intervention effects: systematic review BMJ 373:n736

Welton N, Caldwell D, Adamopoulos E, et al. (2009) Mixed Treatment Comparison Meta-Analysis of Complex Interventions: Psychological Interventions in Coronary Heart Disease. Am J Epidemiol 169:1158

Whiting P, Rutjes A, Westwood M et al. and the QUADAS-2 group (2011) QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. Annals of Internal Medicine 155:529

7 Incorporating economic evaluation

7.1 Introduction

This chapter describes the role of economics in developing NICE guidelines, and suggests possible approaches to use when considering and incorporating economic evidence. It also sets out the principles for doing new economic modelling if there is not enough or no applicable published evidence to assess the cost effectiveness of key interventions, services or programmes.

7.2 The role of economic evaluation in guideline development

Economic evaluation compares the costs and consequences of alternative courses of action. Formally assessing the cost effectiveness of an intervention, service or programme can help decision-makers ensure that maximum gain is achieved from limited resources. If resources are used for interventions or services that are not cost effective, the population as a whole gains fewer benefits.

It is particularly important for committee members to understand that economic analysis does not just estimate the resource consequences of a guideline recommendation, but it evaluates costs in relation to benefits (including benefits to quality of life) and harm of alternative courses of action. Our principles set out that our recommendations should not be based on evidence of costs and benefit alone, and that we must take into account other factors when developing our guidance, as well as recognising that decisions about a person's care are often sensitive to their preferences.

Guideline recommendations should be based on the balance between the estimated costs of the interventions or services and their expected benefits compared with an alternative (that is, their 'cost effectiveness'). In general, the committee should be increasingly certain of the cost effectiveness of a recommendation as the cost of implementation increases.

Defining the priorities for economic evaluation should start during scoping of the guideline, and should continue when the review questions are being developed (see the chapter on developing review questions and planning the evidence review). Questions on economic

issues mirror the review questions on effectiveness, but with a focus on cost effectiveness. Health economic input in guidelines typically involves 2 stages. The first is a literature review of published economic evidence to determine whether the review questions set out in the scope have already been assessed by economic evaluations (see the chapter on identifying the evidence: literature searching and evidence submission). Reviews of economic evidence identify, present and appraise data from studies of cost effectiveness. They may be considered as part of each guideline review question. If existing economic evidence is inadequate or inconclusive for 1 or more review questions, the second stage may involve a variety of economic modelling approaches such as adapting existing economic models or building new models from existing data.

The committee may require more robust evidence on the effectiveness and cost effectiveness of recommendations that are expected to have a substantial impact on resources. Economic analysis must be done when there is no robust evidence of cost effectiveness to support these recommendations. Any uncertainties must be offset by a compelling argument in favour of the recommendation. However, the cost impact or savings potential of a recommendation should not be the sole reason for the committee's decision, see our principles.

Resource impact is considered in terms of the additional cost or saving above that of current practice for each of the first 5 years of implementing the guideline. Resource impact is defined as substantial if:

- implementing a single guideline recommendation in England costs more than £1 million per year or
- implementing the whole guideline in England costs more than £5 million per year.

The aim is to ensure that the guideline does not introduce a cost pressure into the health and social care system unless the committee is convinced of the benefits and cost effectiveness of the recommendations (NICE 2021).

Reviews of economic evidence and any economic modelling are quality assured by the development team and a member of staff with responsibility for independent quality assurance, see the introduction chapter. The nature of the quality assurance will depend on the type of economic evaluation, but will consider the evaluation in terms of the appropriate reference case and be based on a methodology checklist (for example, those in the appendix on appraisal checklists, evidence tables, GRADE and economic profiles).

7.3 Economic analyses

Common types of economic analysis are summarised in box 7.1.

Box 7.1 Types of economic analysis

- Cost-minimisation analysis: a determination of the least costly among alternative interventions that are assumed to produce equivalent outcomes
- Cost-effectiveness analysis: a comparison of costs in monetary units with outcomes in quantitative non-monetary units (for example, reduced mortality or morbidity)
- Cost-utility analysis: a form of cost-effectiveness analysis that compares costs in monetary units with outcomes in terms of their utility, usually to the patient, measured in quality-adjusted life years (QALYs)
- Cost-consequence analysis: a form of cost-effectiveness analysis that presents costs and outcomes in discrete categories, without aggregating or weighting them
- Cost-benefit analysis: a comparison of costs and benefits, both of which are quantified in common monetary terms

Cost-minimisation analysis is the simplest form of economic analysis. It can be used when the health effects of an intervention are the same as those of the status quo, and when there are no other criteria for whether the intervention should be recommended. For example, cost-minimisation analysis could be used to decide whether a doctor or nurse should give routine injections when it is found that both are equally effective at giving injections (on average). In cost-minimisation analysis, an intervention is cost effective only if its net cost is lower than that of the status quo. The disadvantage of cost-minimisation analysis is that the health effects of an intervention cannot often be considered equal to those of the status quo.

Cost-effectiveness analysis uses a measure of outcome (a life-year saved, a death averted, a patient-year free of symptoms) and assesses the cost per unit of achieving this outcome by different means. The outcome is not separately valued, only quantified; so, the study takes no view on whether the cost is worth incurring, only focusing on the cost of different methods to achieve units of outcome.

Cost-utility analysis is a form of cost-effectiveness analysis that uses utility as a common outcome, where utilities are measured on a scale between 0 (death) and 1 (full health). It considers people's quality of life (utility) and the length of life (time) they will gain because of an intervention or a programme. The health effects are expressed as quality-adjusted life years (QALYs), which combine both time and utility in a single measure of health gain. QALYs can be used as a common measure across different populations and disease areas. Costs of resources, and their valuation, should be related to the prices relevant to the sector.

We routinely use cost-utility analysis for the economic evaluation of health-related interventions, programmes and services, for several reasons:

- When used in conjunction with an NHS and personal social services (PSS) perspective, it provides a single yardstick or 'currency' for measuring the impact of interventions. It also allows comparison of interventions so that resources may be allocated more efficiently.
- When possible, NICE programmes use a common method of cost-effectiveness analysis that allows comparisons between programmes.

If a cost-utility analysis is not possible (for example, when outcomes cannot be expressed using a utility measure such as the QALY), a cost-consequences analysis may be considered. Cost-consequences analysis can consider all the relevant health and non-health effects of an intervention across different sectors and reports them without aggregation. A cost-consequences analysis that includes most or all of the potential outcomes of an intervention will be more useful than an analysis that only reports 1 or 2 outcomes.

A cost-consequences analysis is useful when different outcomes cannot be incorporated into an index measure. It is helpful to produce a table that summarises all the costs and outcomes and enables the options to be considered in a concise and consistent manner. Outcomes that can be quantified and monetised can be presented in monetary terms. Some effects may be quantified but cannot readily be put into monetary form (for more details see the Department for Transport's Transport Analysis Guidance [TAG] unit A2.1 [2019]). Some effects cannot readily be quantified (such as reductions in the degree of bullying or discrimination) and should be considered by decision-making committees as part of a cost-consequences analysis alongside effects that can be quantified.

All effects (even if they cannot be quantified) and costs of an intervention are considered

when deciding which interventions represent the best value. Effectively, cost-consequences analysis provides a 'balance sheet' of outcomes that decision-makers can weigh up against the costs of an intervention (including related future costs). However, the outcomes are not separately valued, only quantified; so, the study takes no view on whether the cost is worth incurring, only focusing on the cost of different methods to achieve units of outcome.

Cost–benefit analysis considers health and non-health effects but converts them into monetary values, which can then be aggregated. Once this has been done, 'decision rules' are used to decide which interventions to use. Several metrics are available for reporting the results of cost–benefit analysis. Two commonly used metrics are the 'benefit cost ratio' and the 'net present value' – see the Department for Transport's Transport Analysis Guide (TAG) Unit A1.1 (2021) for more information.

7.4 The reference case

A guideline may consider a range of interventions, commissioned by various organisations and resulting in different types of benefits (outcomes). It is crucial that reviews of economic evidence and economic evaluations used to inform guideline development adopt a consistent approach depending on the type of interventions assessed. The 'reference case' specifies the methods considered consistent with the objective of maximising benefits from limited resources. We are interested in benefits to patients (for interventions with health outcomes in NHS and PSS settings), to individuals and community groups (for interventions with health and non-health outcomes in public sector settings), and to people using services and their carers (for interventions with a social care focus).

Choosing the most appropriate reference case depends on whether or not the interventions being evaluated:

- are commissioned by the NHS and PSS alone or by any other public sector body
- focus on social care outcomes.

The reference case chosen should be agreed for each decision problem (relevant to a review question), and be described briefly in the scope and detailed in the economic plan. A guideline may use a different reference case for different decision problems if appropriate (for example, if a guideline reviews interventions with non-health- or social care-related outcomes). This should be agreed with staff who have responsibility for

quality assurance before any economic evaluation is started.

Table 7.1 summarises the reference case according to the interventions being evaluated.

Table 7.1 Summary of the reference case

Element of assessment	Interventions funded by the NHS and personal social services (PSS) with health outcomes	Interventions funded by the public sector with health and non-health outcomes	Interventions funded by the public sector with a social care focus
Defining the decision problem	The scope we develop		
Comparator	Interventions routinely used in the NHS, including those regarded as current best practice	Interventions routinely used in the public sector, including those regarded as best practice	Interventions routinely delivered by the public and non-public social care sector (Social care costs are the costs of interventions which have been commissioned or paid for in full, or in part by non-NHS organisations)
Perspective on costs	NHS and PSS; for PSS include only care that is funded by NHS (such as 'continuing healthcare' or 'funded nursing care') Costs borne by people using services that are reimbursed by the NHS or PSS should also be included	Public sector – often reducing to local government Other (where appropriate); for example, employer Costs borne by people using services that are reimbursed by the public sector should also be included	Public sector – often reducing to local government Other (where appropriate); for example, employer Costs borne by people using services and the value of unpaid care may also be included if they contribute to outcomes

Element of assessment	Interventions funded by the NHS and personal social services (PSS) with health outcomes	Interventions funded by the public sector with health and non-health outcomes	Interventions funded by the public sector with a social care focus
Perspective on outcomes	All direct health effects, whether for people using services and/or, when relevant, other people (principally family members and/or unpaid carers)	All direct health and relevant non-health effects on individuals. For local government and other settings, where appropriate, non-health effects may also be included	All direct health and relevant non-health effects on people for whom services are delivered (people using services and/or carers)
Type of economic evaluation	Cost-utility analysis	Cost-utility analysis (base case) Cost-effectiveness analysis Cost-consequences analysis Cost–benefit analysis Cost-minimisation analysis	Cost-utility analysis (base case) Cost-effectiveness analysis Cost-consequences analysis Cost–benefit analysis Cost-minimisation analysis
Synthesis of evidence on outcomes	Based on a systematic review, with a preference for treatment effects from randomised controlled trials (RCTs)		
Time horizon	Long enough to reflect all important differences in costs or outcomes between the interventions being compared		
Measuring and valuing health effects	Quality-adjusted life years (QALYs): the EQ-5D-3L is the preferred measure of health-related quality of life in adults (See our position statement on the EQ-5D-5L)		

Element of assessment	Interventions funded by the NHS and personal social services (PSS) with health outcomes	Interventions funded by the public sector with health and non-health outcomes	Interventions funded by the public sector with a social care focus
Measure of non-health effects	Not applicable	Where appropriate, to be decided on a case-by-case basis	Capability or social care-related quality of life measures when an intervention results in both health and capability or social care outcomes, such as ICECAP or ASCOT
Source of data for measurement of quality of life	Reported directly by people using service and/or carers		
Source of preference data for valuation of changes in health-related quality of life	Representative sample of the UK population		
Discounting	The same annual rate for both costs and health effects (currently 3.5%)		
Equity considerations: QALYs	A QALY has the same weight regardless of who receives the health benefit		
Equity and health inequality considerations: other	Equity and health inequality considerations relevant to specific topics, and how these were addressed in economic evaluation, must be reported		
Evidence on resource use and costs	Costs should relate to the perspective used and should be valued using the prices relevant to that perspective		

Interventions funded by the NHS and PSS with health outcomes

For decision problems that evaluate an intervention solely commissioned by the NHS and do not have a clear focus on non-health outcomes, the reference case for 'interventions funded by the NHS and PSS with health outcomes' should be chosen.

All relevant NHS and PSS costs that change because of an intervention should be taken into account. Important non-NHS and PSS costs should also be identified and considered for inclusion in sensitivity analysis, or to aid decision making. These may include costs to other central government departments and local government. Service recommendations are likely to have additional costs, which include implementation costs not usually included in the analysis and costs to other government budgets, such as social care. Implementation costs should be included in a sensitivity analysis, where relevant, while costs to other government budgets can be presented in a separate analysis to the base case.

Biosimilar medicines are considered to differ from the original product in price only (in line with the NICE technology appraisal position statement on the use of biosimilars). It may be necessary to consider modelling scenarios for different levels of uptake for biosimilars, and whether there is any relevant Medicines and Healthcare products Regulatory Agency (MHRA) or government guidance on switching to biosimilars.

Productivity costs and costs borne by people using services and carers that are not reimbursed by the NHS or PSS should usually be excluded from any analyses. That is, a societal perspective will not normally be used.

More details on methods of economic evaluation for interventions with health outcomes in NHS and PSS settings can be found in NICE health technology evaluations: the manual (2022).This includes a reference case, which specifies the methods we consider to be the most appropriate for analysis when developing technology appraisal guidance. The reference case is consistent with the NHS objective of maximising health gain from limited resources.

Interventions funded by the public sector with health and non-health outcomes

For decision problems where the interventions evaluated are commissioned in full or in part by non-NHS public sector and other bodies, the reference case for 'interventions

funded by the public sector with health and non-health outcomes' should be chosen. For the base-case analysis, a cost-utility analysis should be done using a cost per QALY where possible.

This reference case may be most appropriate for public health interventions paid for by an arm of government, and would consider all the costs of implementing the intervention, and changes to downstream costs. In some cases, the downstream costs are negative, and refer to cost savings. For example, an intervention such as increasing physical activity, whose effects may include preventing type 2 diabetes, may be paid for by local government, but may result in cost savings to the NHS in the form of fewer or delayed cases of diabetes. A public sector cost perspective would aggregate all these costs and cost savings. A narrower local government cost perspective would consider only the cost of implementation, whereas an NHS cost perspective would consider only the cost savings. When examining interventions that are not paid for by an arm of government (such as workplace interventions), the perspective on costs should be discussed and agreed with staff who have responsibility for quality assurance.

Productivity costs should usually be excluded from both the reference-case and non-reference-case analyses; exceptions (for example, when evaluating interventions in the workplace) can only be made with the agreement of staff with responsibility for quality assurance.

For public health interventions, all direct health effects for people using services or, when relevant, other people such as family members or unpaid carers will be included. Non-health effects may also be included. When needed, the perspective will be widened to include sectors that do not bear the cost of an intervention, but receive some kind of benefit from it.

Interventions with a social care focus

For decision problems where the interventions evaluated have a clear focus on social care outcomes, the reference case on 'interventions with a social care focus' should be chosen. For the base-case analysis, a cost-utility analysis should be done using a cost per QALY approach where possible.

Public sector funding of social care for individual service users is subject to eligibility criteria based on a needs assessment and a financial assessment (means test). Therefore, users of social care may have to fund, or partly fund, their own care. A public sector

perspective on costs should still be adopted, but should consider different scenarios of funding.

A public sector perspective is likely to be a local authority perspective for many social care interventions, but downstream costs that affect other public sector bodies may be considered where relevant, especially if they are a direct consequence of the primary aim of the intervention.

When individuals may pay a contribution towards their social care, 2 further perspectives may also be relevant:

- a wider perspective (which takes account of changes to the amount that individuals and private firms pay towards the cost of care, on top of the public sector contributions) and
- an individual perspective (which accounts for changes in individual payments only).

The value of unpaid care may also be included in sensitivity analysis, or to aid decision making. The value of unpaid care should be set at the market value of paid care. Productivity costs should usually be excluded from both the reference-case and non-reference-case analyses; exceptions can only be made with the agreement of staff with responsibility for quality assurance.

For social care interventions, the usual perspective on outcomes will be all effects on people for whom services are delivered including, when relevant, family members or unpaid carers. When needed, the perspective may be widened to include sectors that do not bear the cost of an intervention, but receive some kind of benefit from it.

Other perspectives

Other perspectives (for example, employers) may also be used to capture significant costs and effects that are relevant to the interventions. Agree use of other perspectives with staff who have responsibility for quality assurance before use.

7.5 Reviewing economic evaluations

Identifying and examining published economic evidence that is relevant to the review questions is an important component of guideline development. The general approach to

reviewing economic evaluations should be systematic, focused and pragmatic. The principal search strategy (see the section on developing search strategies in the chapter on identifying the evidence: literature searching and evidence submission), including search strategies for health economic evidence, should be posted on the NICE guideline's webpage 6 weeks before consultation on the draft guideline.

Searching for economic evidence

Advice on the approach to searching for economic evidence is detailed in the chapter on identifying the evidence: literature searching and evidence submission.

Selecting relevant economic evaluations

The process for sifting and selecting economic evaluations for assessment is essentially the same as for effectiveness studies (see the section on identifying and selecting relevant evidence in the chapter on reviewing evidence). It should be targeted to identify the papers that are most relevant to current UK practice and therefore likely to inform the committee's decision making.

Inclusion criteria for sifting and selecting papers for each review should specify populations and interventions relevant to the review question. They should also specify:

- An appropriate date range, because older studies may reflect outdated practices.
- The country or setting, because studies done in other countries might not be relevant to the UK. In some cases it may be appropriate to limit consideration to the UK or countries with similar healthcare systems.

The review should also usually focus on economic evaluations that compare both the costs and consequences of the alternative interventions under consideration. Cost-utility, cost-benefit, cost-effectiveness, cost-minimisation or cost-consequences analyses (see box 7.1 in the section on economic analyses) can be considered depending on what the committee deems to be the most relevant perspective and likely outcomes for the question. Non-comparative costing studies, 'burden of disease' studies and 'cost of illness' studies should usually be excluded. However, non-comparative costing studies (such as econometric, efficiency, simulation, micro-costing and resource use, and time-series) may be included for some service delivery questions, or flagged if they might be useful for economic modelling. Sometimes, the published economic evidence is extremely sparse. In such cases, the inclusion criteria for studies may be broadened. The development team

will decide to do this in consultation with staff who have responsibility for quality assurance and, when appropriate, with the committee or its chair.

Assessing the quality of economic evaluations

All economic evaluations relevant to the guideline should be appraised using the methodology checklists (see the appendix on appraisal checklists, evidence tables, GRADE and economic profiles). These should be used to appraise published economic evaluations, as well as unpublished papers, such as studies submitted by registered stakeholders and academic papers that are not yet published. The same criteria should be applied to any new economic evaluations done for the guideline (see the section on approaches to original economic evaluation).

Consideration and inclusion of economic evaluations will depend on the applicability of evidence to our decision-making context (usually the reference case), the amount of higher-quality evidence and the degree of certainty about the cost effectiveness of an intervention (when all the evidence is considered as a whole). Lower-quality studies are more likely to be excluded when cost effectiveness (or lack of it) can be reliably established without them. The reasons for such exclusions should be explained in the guideline's evidence review documents.

Sometimes reported sensitivity analyses show whether the results of an evaluation or study are robust despite methodological limitations. If there is no sensitivity analysis, judgement is needed to assess whether a limitation would be likely to change the results and conclusions. If necessary, a checklist, such as the health technology assessment checklist for decision-analytic models (Philips et al. 2004), may also be used to give a more detailed assessment of the methodological quality of economic evaluations and modelling studies. Judgements made, and reasons for these judgements, should be recorded in the guideline's evidence review documents.

Summarising and presenting results for economic evaluations

Cost-effectiveness or net benefit estimates from published or unpublished studies, or from original economic evaluations done for the guideline, should be presented in the guideline's evidence review documents, for example, using an 'economic evidence profile' (see the appendix on appraisal checklists, evidence tables, GRADE and economic profiles), as well as a statement on the cost effectiveness with respect to our decision threshold. This should include relevant economic information (applicability, limitations, costs, effects,

cost effectiveness or net benefit estimates, as appropriate). Costs do not need to be adjusted to present value, but costs from other countries should be converted to pound sterling using an exchange rate from an appropriate and current source (such as HM Revenue and Customs or Organisation for Economic Co-operation and Development).

It should be explicitly stated if economic information is not available or if it is not thought to be relevant to the review question.

Separate health economic evidence statements are no longer required.

7.6 Prioritising questions for further economic analysis

If a high-quality economic analysis that addresses a key issue and is relevant to current practice has already been published, then further modelling may not be needed. However, often the economic literature is not sufficiently robust or applicable. Original economic analyses should only be done if an existing analysis cannot easily be adapted to answer the question.

Economic plans

The full economic plan initially identifies key areas of the scope as priorities for further economic analysis and outlines proposed methods for addressing review questions about cost effectiveness. The full economic plan may be modified during development of the guideline; for example, as evidence is reviewed, it may become apparent that further economic evaluation is not needed or may not be possible for some areas that were initially prioritised. A version of the economic plan setting out the questions prioritised for further economic analysis, the population, the interventions and the type of economic analysis is published on the NICE guideline's webpage at least 6 weeks before the guideline goes out for consultation (see the chapter on developing review questions and planning the evidence review). The reasons for the final choice of priorities for economic analysis should be explained in the guideline's evidence review documents.

Discussing the economic plan with the committee early in guideline development is essential to ensure that:

- the most important questions are selected for economic analysis

- the methodological approach is appropriate (including the reference case)
- all important effects and resource costs are included
- effects and outcomes relating to a wider perspective are included if relevant
- additional effects and outcomes not related to health or social care are included if they are relevant
- economic evidence is available to support recommendations that are likely to lead to substantial costs.

The number and complexity of new analyses depends on the priority areas and the information needed for decision making by the committee. Selection of questions for further economic analysis should be based on systematic consideration of the potential value of economic analysis across all key issues.

Economic analysis is potentially useful for any question in which an intervention, service or programme is compared with another. It may also be appropriate in comparing different combinations or sequences of interventions, as well as individual components of the service or intervention. However, the broad scope of some guidelines means that it may not be practical to conduct original economic analysis for every component.

The decision about whether to carry out an economic analysis therefore depends on:

- the potential overall expected benefit and resource implications of an intervention, both for individual people and the population as a whole
- the degree of uncertainty in the economic evidence review and the likelihood that economic analysis will clarify matters.

Economic modelling may not be warranted if:

- It is not possible to estimate cost effectiveness. However, in this case, a 'scenario' or 'threshold' analysis may be useful.
- The intervention has no likelihood of being cost saving and its harms outweigh its benefits.
- The published evidence of cost effectiveness is so reliable that further economic analysis is not needed.

- The benefits sufficiently outweigh the costs (that is, it is obvious that the intervention is cost effective) or the costs sufficiently outweigh the benefits (that is, it is obvious that the intervention is not cost effective).
- An intervention has very small costs, very small benefits and very small budget impact.

7.7 Approaches to original economic evaluation

General principles

Regardless of the methodological approach taken, the general principles described in this section should be observed. Any variation from these principles should be described and justified in the guideline's modelling report or evidence review documents. The decision problem should be clearly stated. This should include a definition and justification of the interventions or programmes being assessed and the relevant groups using services (including carers).

Developing conceptual models linked to topic areas or review questions may help the health economist to decide what key information is needed for developing effectiveness and cost-effectiveness analyses (see the chapter on the scope for details). Models developed for public health and service delivery topics are likely to relate to several review questions, so most recommendations will be underpinned by some form of modelled analysis.

The choice of model structure is a key aspect of the design-oriented conceptual model. Brennan's taxonomy of model structures (Brennan et al. 2006) should be considered for guidance on which types of models may be appropriate to the decision problem.

Even if a fully modelled analysis is not possible, there may be value in the process of model development, because this will help to structure committee discussions. For example, a model might be able to demonstrate how a change in service will affect demand for a downstream service or intervention.

For service delivery questions, the key challenge is linking changes in service to a health benefit. This obviously poses a challenge when doing health economic analyses, but it will also be difficult finding high-quality evidence of effectiveness. Modelling using scenario analysis is usually needed to generate the health effects used within the health economic

analyses. Because of the considerable resource and health impact of any recommendations on service delivery, its cost effectiveness must be considered, either analytically or qualitatively (see the appendix on service delivery – developing review questions, evidence reviews and synthesis).

Economic analysis should include comparison of all relevant alternatives for specified groups of people affected by the intervention or using services. Any differences between the review questions and the economic analysis should be clearly acknowledged, justified, approved by the committee and explained in the guideline's modelling report or evidence review documents. The interventions or services included in the analysis should be described in enough detail to allow stakeholders to understand exactly what is being assessed. This is particularly important when calculating the cost effectiveness of services.

An economic analysis should be underpinned by the best-quality evidence. The evidence for treatment effectiveness should be based on and be consistent with that identified for the relevant effectiveness review question. However, sometimes models may differ from the guideline review, such as when additional outcomes or additional timepoints might be extracted from the effectiveness evidence, that were not prioritised in the effectiveness review.

Where there is not enough evidence in the effectiveness review, usually a research recommendation will be made and modelling will not go ahead. Occasionally, a model will be developed based upon indirect evidence, bespoke analysis of real-world data or expert opinion. This should be discussed with staff responsible for quality assurance and clearly explained and justified in the guideline.

The structure of any economic model should be discussed and agreed with the committee early in guideline development. The reasons for the structure of the model should be clear. Potential alternatives should be identified and considered for use in sensitivity analysis. If existing economic models are being used, or are informing a new analysis, particularly economic models used previously in the guideline's development, the way these models are adapted or used should be clear.

Clinical endpoints that reflect how a patient feels, functions, or how long a patient lives are considered more informative than surrogate outcomes. When using 'final' clinical endpoints in a model is not possible and data on other outcomes are used to infer the effect on mortality and health-related quality of life, there should be evidence supporting the

outcome relationship (see NICE health technology evaluations: the manual [2022] for the 3 levels of evidence for surrogate relationships that can be considered in decision making). The uncertainty associated with the relationship between the surrogate endpoints and the final outcomes should be quantified and captured in the model's probabilistic analysis.

For evaluations of diagnostic testing, there may be some direct benefits from the knowledge gained and some direct harm from the testing, but most of the outcomes come downstream because of treatment or preventive measures being started, modified or stopped. Diagnostic tests can sometimes be evaluated using clinical trials, but this is unusual. If direct data on the impact of a diagnostic test strategy on final outcomes is not available, it will be necessary to combine evidence from different sources. A linked-evidence modelling approach should be used that captures the following components:

- the pre-test prevalence of the disease
- diagnostic test accuracy statistics for the test or tests being evaluated
- diagnostic test accuracy statistics for subsequent tests
- the treatment pathway for each test outcome
- treatment outcomes
- adverse events associated with testing including the impact of radiation exposure from an imaging test.

Clinical trial populations often differ from the population of interest and so might not capture the absolute effects of an intervention. Quantifying the baseline risk of health outcomes and how the condition would naturally progress with the comparator or comparators can be a useful step when estimating absolute health outcomes in the economic analysis. This can be informed by observational studies. Relative treatment effects seen in randomised trials may then be applied to data on the baseline risk of health outcomes for the populations or subgroups of interest. The methods used to identify and critically evaluate sources of data for these estimates should be reported.

When outcomes in an economic evaluation are known to be related, a joint synthesis of structurally related outcomes is recommended whenever possible, to account for correlation and to increase precision. Examples of these relations include:

- network meta-analysis – the correlation between different pair-wise comparisons (see

NICE's Decision Support Unit technical support document 6)

- diagnostic meta-analysis – the inverse correlation between sensitivity and specificity
- utility mapping equations – the different coefficients in the equation will be correlated (see NICE's Decision Support Unit technical support document 10).

Studies using survival outcomes, or time-to-event outcomes, often measure the relative effects of treatments using hazard ratios (HRs), which may either be constant over time (proportional hazards) or change over time. When incorporating such outcomes into a model, the proportional hazards assumption should always be assessed (see NICE's Decision Support Unit technical support document 14).

Sometimes there will be a lack of robust quantitative evidence for a key model parameter. In such situations, informal or formal expert opinion might sometimes be used to identify a plausible distribution of values.

For service delivery questions, any analysis will need to consider resource constraints. These might be monetary, but might also be resources such as staff, beds, equipment and so on. However, affordability should not be the sole consideration for service recommendations; the impact of any proposed changes on quality of care needs to be considered.

Before presenting final results to a committee for decision making, all economic evaluations should undergo rigorous quality assessment and validation to assess inputs, identify logical, mathematical and computational errors, and review the plausibility of outputs. The HM Treasury's review of quality assurance of government models (2013) provides guidance on developing the environment and processes required to promote effective quality assurance. This process should be recorded within the guideline's modelling report or evidence review documents.

Quality assurance of an economic evaluation may take various forms at different stages in development, as detailed in the HM Treasury Aqua Book (2015). It can range from basic steps that should always occur, such as disciplined version control, extensive developer testing of their own model, and independent testing by a colleague with the necessary technical knowledge, to external testing by an independent third party and independent analytical audit of all data and methods used. For developer health economists testing their own evaluation, or those of others ('model busting'), useful and practical validation methods include:

- 1-way and n-way sensitivity analyses, including null values and extreme values
- ensuring that the model results can be explained, for example, the logic and reason underlying the effect of a particular scenario analysis on results
- ensuring that predictions of intermediate endpoints (for example, event rate counts) and final endpoints (for example, undiscounted life expectancy) are plausible, including comparison with source materials.

Results should be reported of any analyses done to demonstrate external validity. However, relevant data should not be omitted just to enable external validation (for example, not including trials so that they can be used for subsequent validation).

Conventions on reporting economic evaluations should be followed (see Husereau et al. 2022) to ensure that reporting of methods and results is transparent. For time horizons that extend beyond 10 years, it may be useful to report discounted costs and effects for the short (1 to 3 years) and medium (5 to 10 years) term. The following results should be presented where available and relevant:

- endpoints from the analysis, such as life years gained, number of events and survival
- disaggregated costs
- total and incremental costs and effects for all options.

When comparing multiple mutually exclusive options, an incremental approach should be adopted. This should be done by:

- comparing the interventions sequentially in rank order of cost or outcome, with each strategy compared with the next non-dominated alternative in terms of (for cost-utility analyses) the 'incremental cost per QALY gained'
- ranking the interventions in order of mean 'net health benefit' (or 'net monetary benefit'), where the opportunity cost is specified by the incremental (monetary) cost divided by a specific cost per QALY threshold.

Comparisons with a common baseline intervention should not be used for decision making (although should be included in the incremental analysis if it reflects a relevant option).

Any comparison of interventions in an economic model that are not based on head-to-head trial comparisons should be carefully evaluated for the between-study heterogeneity,

and potential for modifiers of treatment effect should be explored. Limitations should be noted and clearly discussed in the guideline's modelling report or evidence review documents. Ideally, when options have not been directly compared, and, more generally, when there are more than 2 options, a network meta-analysis should be considered as the best option for evidence synthesis to inform the model (see the chapter on reviewing evidence).

Economic models developed for the guideline are available to registered stakeholders during consultation on the guideline. These models should be fully executable and clearly presented.

Different approaches to economic analysis

There are different approaches to economic analysis (see box 7.1 in the section on economic analyses for examples). If economic analysis is needed, the most appropriate approach should be considered early during the development of a guideline, and reflect the content of the guideline scope.

There is often a trade-off between the range of new analyses that can be done and the complexity of each piece of analysis. Simple methods may be used if these can provide the committee with enough information on which to base a decision. For example, if an intervention is associated with better effectiveness and fewer adverse effects than its comparator, then an estimate of cost may be all that is needed. Or a simple decision tree may provide a sufficiently reliable estimate of cost effectiveness. In other situations, a more complex approach, such as Markov modelling or discrete event simulation, may be warranted.

The type of economic analysis that should be considered is informed by the economic perspective specified in the scope of the guideline, and the extent to which the effects resulting from the intervention extend beyond health.

Measuring and valuing effects for health interventions

The measurement of changes in health-related quality of life should be reported directly from people using services (or their carers). The value placed on health-related quality of life of people using services (or their carers) should be based on a valuation of public preferences obtained from a representative sample of the UK population, using a choice-based valuation method such as the time trade-off or standard gamble. We prefer the

QALY as the measure of health effects, and the EQ-5D as the instrument to measure health-related quality of life in adults.

For some economic analyses, a flexible approach may be needed, reflecting the nature of effects delivered by different interventions or programmes. If health effects are relevant, the EQ-5D-based QALY should be used. When EQ-5D data are not available from the relevant clinical studies included in the clinical evidence review, EQ-5D data can be sourced from the literature. The methods used for identifying the data should be systematic and transparent. The justification for choosing a particular data set should be clearly explained. When more than 1 plausible set of EQ-5D data is available, sensitivity analyses should be carried out to show the impact of the alternative utility values.

When EQ-5D data are not available, published mapped EQ-5D data should be used, or they may be estimated by mapping other health-related quality-of-life measures or health-related effects observed in the relevant studies to the EQ-5D if data are available. The mapping function chosen should be based on data sets containing both health-related quality-of-life measures. The statistical properties of the mapping function should be fully described, its choice justified, and it should be adequately demonstrated how well the function fits the data. Sensitivity analyses exploring variation in the use of the mapping algorithms on the outputs should be presented.

In some circumstances, EQ-5D data may not be the most appropriate or may not be available. Qualitative empirical evidence on the lack of content validity for the EQ-5D should be provided, demonstrating that key dimensions of health are missing. This should be supported by evidence that shows that EQ-5D performs poorly on tests of construct validity and responsiveness in a particular patient group. This evidence should be derived from a synthesis of peer-reviewed literature. In these circumstances, alternative health-related quality of life measures may be used and must be accompanied by a carefully detailed account of the methods used to generate the data, their validity, and how these methods affect the utility values.

When necessary, consideration should be given to alternative standardised and validated preference-based measures of health-related quality of life that have been designed specifically for use in children. The standard version of the EQ-5D has not been designed for use in children. We do not recommend specific measures of health-related quality of life in children and young people. A generic measure that has shown good psychometric performance in the relevant age ranges should be used. Technical support document 8 by NICE's Decision Support Unit (Brazier et al. 2011) summarises the psychometric

performance of several preference-based measures.

As outlined in NICE health technology evaluations: the manual (2022) and the accompanying NICE position statement on use of the EQ-5D-5-level (5L) valuation set for England (updated October 2019), we do not currently recommend using the EQ-5D 5L valuation set. Guideline development teams should use the 3-level (3L) valuation set for reference-case analyses, when available.

The QALY remains the most suitable measure for assessing the impact of services, because it can incorporate effects from extension to life and experience of care. It can also include the trade-offs of benefits and adverse events. However, if linking effects to a QALY gain is not possible, links to a clinically relevant or a related outcome should be considered. Outcomes should be optimised for the lowest resource use. The link (either direct or indirect) of any surrogate outcome, such as a process outcome (for example, bed days), to a clinical outcome needs to be justified. However, when QALYs are not used, issues such as trade-offs between different beneficial and harmful effects need to be considered.

Measuring and valuing effects for non-health interventions

For some decision problems (such as for interventions with a social care focus), the intended outcomes of interventions are broader than improvements in health status. Here broader, preference-weighted measures of outcomes, based on specific instruments, may be more appropriate. For example, social care quality-of-life measures are being developed and we will consider using 'social care QALYs' if validated, such as the ASCOT (Adult Social Care Outcome Toolkit) set of instruments used by the Department of Health and Social Care in the Adult Social Care Outcomes Framework indicator on social care-related quality of life.

Similarly, depending on the topic, and on the intended effects of the interventions and programmes, the economic analysis may also consider effects in terms of capability and wellbeing. For capability effects, we may consider use of the ICECAP-O (Investigating Choice Experiments for the Preferences of Older People CAPability measure for Older people) or ICECAP-A (Investigating Choice Experiments for the Preferences of Older People CAPability measure for Adults) instruments when developing methodology in the future. If an intervention is associated with both health- and non-health-related effects, it may be helpful to present these elements separately.

Economic analysis for interventions funded by the NHS and PSS with health outcomes

Economic analyses done for decisions about interventions with health outcomes funded by the NHS and PSS should usually follow the reference case in table 7.1 in the section on the reference case. Advice on how to follow approaches described in NICE health technology evaluations: the manual (2022) is provided in the NICE Decision Support Unit's technical support documents. Departures from the reference case may sometimes be appropriate; for example, when there are not enough data to estimate QALYs gained. Any such departures must be agreed with members of staff with responsibility for quality assurance and highlighted in the guideline's modelling report or evidence review documents with reasons given.

Economic analysis for interventions funded by the public sector with health and non-health outcomes

The usual perspective for the economic analysis of public health interventions is that of the public sector. This may be simplified to a local government perspective if few costs and effects apply to other government agencies.

As public health and wellbeing programmes are funded by the public sector, in particular by local authorities, we have broadened our approach for the appraisal of interventions in these areas. Local government is responsible not only for the health of individuals and communities, but also for their overall welfare. The tools used for economic evaluation must reflect a wider remit than health and allow greater local variation in the outcomes captured. The nature of the evidence and that of the outcomes being measured may place more emphasis on cost-consequences analysis and cost–benefit analysis for interventions in these areas.

Whenever there are multiple outcomes, a cost-consequences analysis is usually needed, and the committee weighs up the changes to the various outcomes against the changes in costs in an open and transparent manner. However, for the base-case analysis, a cost-utility analysis should be undertaken using a cost per QALY approach where possible.

A wider perspective may be used, and will usually be carried out using cost–benefit analysis. When a wider perspective is used, it must be agreed with staff who have responsibility for quality assurance and highlighted in the guideline's modelling report or evidence review documents with reasons given.

Economic analysis for interventions with a social care focus

For social care interventions, the perspective on outcomes should be all effects on people for whom services are delivered (people using services or carers). Effects on people using services and carers (whether expressed in terms of health effects, social care quality of life, capability or wellbeing) are the intended outcomes of social care interventions and programmes. Although holistic effects on people using services, their families and carers may represent the ideal perspective on outcomes, a pragmatic and flexible approach is needed to address different perspectives, recognising that improved outcomes for people using services and carers may not always coincide.

Whenever there are multiple outcomes, a cost-consequences analysis is usually needed, and the committee weighs up the changes to the various outcomes against the changes in costs in an open and transparent manner. However, for the base-case analysis, a cost-utility analysis should be done using a cost per QALY approach where possible.

Any economic model should take account of the proportion of care that is publicly funded or self-funded. Scenario analysis may also be useful to take account of any known differences between local authorities in terms of how they apply eligibility criteria. Scenario analysis should also be considered if the cost of social care varies depending on whether it is paid for by local authorities or by individual service users; the value of unpaid care should also be taken into account where appropriate. A range of valuation methods exists to cost this type of care, so methods chosen should be clearly described and sensitivity analyses using other methods should be considered.

It is expected that the analytical difficulties involved in creating clear, transparent decision rules around the costs that should be considered, and for which interventions and outcomes, will be particularly problematic for social care. These should be discussed with the committee before any economic analysis is undertaken and an approach agreed.

Identification and selection of model inputs

An economic analysis uses decision-analytic techniques with outcome, cost and utility data from the best available published sources.

The reference case across all perspectives (see table 7.1 in the section on the reference case) states that evidence on effects should be obtained from a systematic review with a preference on randomised controlled trials. Some inputs, such as costs, may have

standard sources that are appropriate, such as national list prices or a national audit, but for others appropriate data will need to be sourced.

Additional searches may be needed; for example, if searches for evidence on effects do not provide the information needed for economic modelling. Additional information may be needed on:

- disease prognosis
- the relationship between short- and long-term outcomes
- quality of life
- adverse events
- antimicrobial resistance
- resource use or costs.

Although it is desirable to conduct systematic literature reviews for all such inputs, this is time-consuming and other pragmatic options for identifying inputs may be used. Informal searches should aim to satisfy the principle of 'saturation' (that is, to 'identify the breadth of information needs relevant to a model and sufficient information such that further efforts to identify more information would add nothing to the analysis' [Kaltenthaler et al. 2011]). Studies identified in the review of evidence on effects should be scrutinised for other relevant data, and attention should be paid to the sources of parameters in analyses included in the systematic review of published economic evaluations. Alternatives could include asking committee members and other experts for suitable evidence or eliciting their opinions, this could include for example, using formal consensus methods such as the Delphi method or the nominal-group technique. If a systematic review is not possible, transparent processes for identifying model inputs (such as effectiveness, adverse events, diagnostic accuracy, prognosis, quality of life, resource use and costs) should be reported; the internal quality and external validity of each potential data source should be assessed, and their selection justified. If more than 1 suitable source of evidence is found, consideration should be given to synthesis or exploration of alternative values in sensitivity analyses.

For some questions, there may be good reason to believe that relevant and useful information exists outside of literature databases or validated national data sources. Examples include ongoing research, a relatively new intervention and studies that have

been published only as abstracts. Typically, the method for requesting information from stakeholders is through a call for evidence (see the appendix on call for evidence and expert witnesses).

For some guidelines, econometric studies provide a supplementary source of evidence and data for bespoke economic models. For these studies, the database 'Econlit' should be searched as a minimum.

Real-world data

'Real-word data' is defined in the NICE real-world evidence framework (2022) as 'data relating to patient health or experience, or care delivery collected outside the context of a highly controlled clinical trial'. Data from electronic health records, registries, audits and other sources of real-world data may be used to better define and inform parameter estimates for economic models. For some parameters, real-world data are considered the most appropriate evidence source (for example, jurisdiction-specific administrative databases can be used to estimate control or untreated event probabilities, resource use counts and unit costs), or such data may be more up to date and closer to reality (for example, in relationship to real-world costs and populations), so they should be the main source of such data in economic models. Real-world evidence should not be considered a substitute for published evidence from randomised controlled trials when assessing differences in outcomes between interventions. However, real-world evidence may be used where evidence from trials is absent or insufficient, or to answer different, but related, research questions alongside good published randomised controlled trial evidence (for example, effectiveness versus efficacy, effects in subgroups of interest, UK-specific head-to-head data and to derive final from surrogate trial outcomes).

To obtain real-world data, it may be necessary to negotiate access with the organisations and individuals that hold the data, or to ask them to provide a summary for inclusion in the guidance if published reports are insufficient. Any processes used for accessing data will need to be reported in the economic plan and in the guideline. Given the difficulties that organisations may have in extracting audit data, such requests should be focused and targeted: for example, identifying a specific audit and requesting results from the previous 3 years.

Additional guidance on the use of non-randomised studies and real-world data is provided by the NICE real-world evidence framework (2022), which:

- identifies when real-world data can be used to reduce uncertainties and improve guidance
- provides guidance on searching for and selecting appropriate real-world data sources
- describes best-practices for planning, doing and reporting real-world evidence studies to improve the quality and transparency of evidence.

Further guidance on searching and selecting evidence for key model inputs is also provided by Kaltenthaler et al. (2011) and Paisley (2016).

Cost data

Some information on unit costs may be found in the Personal Social Services - Research Unit (PSSRU) report on unit costs of health and social care or NHS England's National Cost Collection. Information on resource impact costings can be found in NICE's process guide on resource impact assessment. The NHS Supply Chain catalogue provides costs of clinical consumables, capital medical equipment and non-medical products. Some information about public services may be better obtained from national statistics or databases, rather than from published studies. Philips et al. (2004) provides a useful guide to searching for data for use in economic models.

In cases where current costs are not available, costs from previous years should be adjusted to present value using inflation indices appropriate to the cost perspective, such as the hospital and community health services index and the PSS pay and prices index, available from the PSSRU report on unit costs of health and social care (Jones 2021), or the Office for National Statistics (ONS) consumer price index.

Wherever possible, costs relevant to the healthcare system in England should be used. However, in cases where only costs from other countries are available these should be converted to Pounds Sterling using an exchange rate from an appropriate and current source (such as HM Revenue and Customs or Organisation for Economic Co-operation and Development).

Usually the public list prices for technologies (for example, medicines or medical devices) should be used in the reference-case analysis. However, as outlined in NICE health technology evaluations: the manual (2022), reference-case analyses should be based on prices that reflect as closely as possible the prices that are paid in the NHS for all technology evaluations. When there are nationally available price reductions (for example,

for medicines procured for use in secondary care through contracts negotiated by the NHS Commercial Medicines Unit), the reduced price should be used in the reference-case analysis to best reflect the price relevant to the NHS. The Commercial Medicines Unit publishes information on the prices paid for some generic medicines by NHS trusts through its Electronic Market Information Tool (eMIT), focusing on medicines in the 'National Generics Programme Framework' for England. Analyses based on price reductions for the NHS will be considered only when the reduced prices are transparent and can be consistently available across the NHS, and when the period for which the specified price is available is guaranteed. When a reduced price is available through a patient access scheme that has been agreed with the Department of Health and Social Care, the analyses should include the costs associated with the scheme. If the price is not listed on eMIT, then the current price listed on the British National Formulary (BNF) should be used. For medicines that are predominantly dispensed in the community, prices should be based on the Drug Tariff. In the absence of a published list price and a price agreed by a national institution (as may be the case for some devices), an alternative price may be considered, provided that it is nationally and publicly available. If no other information is available on costs, local costs obtained from the committee may be used.

Quality of life data

Preference-based quality-of-life data are often needed for economic models. Many of the search filters available are highly sensitive and so, although they identify relevant studies, they also detect a large amount of irrelevant data. An initial broad literature search for quality of life data may be a good option, but the amount of information identified may be unmanageable (depending on the key issue being addressed). It may be more appropriate and manageable to incorporate a quality of life search filter when performing additional searches for key issues of high economic priority. When searching bibliographic databases for health-state utility values, specific techniques outlined in Ara (2017), Golder et al. (2005) and Papaioannou et al. (2010) may be useful, and specific search filters have been developed that may increase sensitivity (Arber et al. 2017). The provision of quality of life data should be guided by the health economist at an early stage during guideline development so that the information specialist can adopt an appropriate strategy. Resources for identifying useful utility data for economic modelling are the dedicated registries of health-state utility values such as ScHARRHUD and Tufts CEA Registry and the NICE Decision Support Unit's technical support documents.

Exploring uncertainty

The committee should discuss any potential bias, uncertainties and limitations of economic models. Sensitivity analysis should be used to explore the impact that potential sources of bias and uncertainty could have on model results.

Deterministic sensitivity analysis should be used to explore key assumptions and input parameters used in the modelling, as well as to test any bias resulting from the data sources selected for key model inputs. This should test whether and how the model results change under alternative, plausible scenarios. Threshold analysis should also be considered, particularly for highly uncertain parameters, to explore the impact of the parameter on the incremental cost-effectiveness ratio (ICER) or net-health benefit.

'Tornado' histograms may be a useful way to present deterministic results. Deterministic threshold analysis might inform decision making when there are influential but highly uncertain parameters. However, if the model is non-linear, deterministic analysis will be less appropriate for decision making.

Probabilistic sensitivity analysis should be used to account for uncertainty arising from imprecision in model inputs. The use of probabilistic sensitivity analysis will often be specified in the health economic plan. Any uncertainty associated with all inputs can be simultaneously reflected in the results, so the preferred cost-effectiveness estimates should be those derived from probabilistic analyses when possible. In non-linear decision models where outputs are a result of a multiplicative function (for example, in Markov models), probabilistic methods also provide the best estimates of mean costs and outcomes. The choice of distributions used should be justified. When doing a probabilistic analysis, use enough model simulations to minimise the effect of Monte Carlo error. Reviewing the variance around probabilistic model outputs (net health benefits or ICERs) as the number of simulations increases can provide a way of assessing whether the model has been run enough times or more runs are needed. Presentation of the results of probabilistic sensitivity analysis could include scatter plots or confidence ellipses, with an option for including cost-effectiveness acceptability curves and frontiers.

When probabilistic methods are unsuitable, or not possible, the impact of parameter uncertainty should be thoroughly explored using deterministic sensitivity analysis, being mindful of correlated parameters. The decision not to use probabilistic methods should be justified in the guideline's modelling report or evidence review documents.

Consideration can be given to including structural assumptions and the inclusion or exclusion of data sources in probabilistic sensitivity analysis. In this case, the method used to select the distribution should be outlined in the guideline (Jackson et al. 2011).

Discounting

Cost-effectiveness results should reflect the present value of the stream of costs and benefits accruing over the time horizon of the analysis. For the reference case, the same annual discount rate should be used for both costs and benefits. We consider that it is usually appropriate to discount costs and health effects at the same annual rate of 3.5%.

Sensitivity analyses using 1.5% as an alternative rate for both costs and health effects may be presented alongside the reference-case analysis, particularly for public health guidance. When treatment restores people who would otherwise die or have a very severely impaired life to full or near full health, and when this is sustained over a very long period (normally at least 30 years), cost-effectiveness analyses are very sensitive to the discount rate used. In this circumstance, analyses that use a non-reference-case discount rate for costs and outcomes may be considered. A discount rate of 1.5% for costs and benefits may be considered by the committee if it is highly likely that, based on the evidence presented, long-term health benefits are likely to be achieved. However, the committee will need to be satisfied that the recommendation does not commit the funder to significant irrecoverable costs.

Subgroup analysis

The relevance of subgroup analysis to decision making should be discussed with the committee. When appropriate, economic analyses should estimate the cost effectiveness of an intervention in each subgroup.

Local considerations

For service delivery questions, cost-effectiveness analyses may need to account for local factors, such as the expected number of procedures and the availability of staff and equipment at different times of the day, week and year. Service delivery models may need to incorporate the fact that each local provider may be starting from a different baseline of identified factors (for example, the number of consultants available at weekends). It is therefore important that these factors are identified and considered by the committee. Where possible, results obtained from the analysis should include both the national

average and identified local scenarios to ensure that service delivery recommendations are robust to local variation.

Service failures

Service designs under consideration might result in occasional service failure – that is, where the service does not operate as planned. For example, a service for treating myocardial infarction may have fewer places where people can be treated at weekends compared with weekdays because of reduced staffing. Therefore, more people will need to travel further away by ambulance and the journey time will also be longer. Given the limited number of ambulances, a small proportion may be delayed, resulting in consequences in terms of costs and QALYs. Such possible service failures should be taken into account in effectiveness and economic modelling. This effectively means that analyses should incorporate the 'side effects' of service designs.

Service demand

Introducing a new service or increasing capacity will often result in an increase in demand. This could mean that a service does not achieve the predicted effectiveness because there is more demand than was planned for. This should be addressed either in the analysis or in considerations.

7.8 Using economic evidence to formulate guideline recommendations

For an economic analysis to be useful, it must inform the guideline recommendations. The committee should discuss cost effectiveness in parallel with general effectiveness when formulating recommendations (see the chapter on interpreting the evidence and writing the guideline).

Within the context of our principles on social value judgements, the committee should be encouraged to consider recommendations that:

- increase effectiveness at an acceptable level of increased cost **or**
- are less effective than current practice, but free up sufficient resources that can be re-invested in public sector care or services to increase the welfare of the population

receiving care.

The committee's interpretations and discussions should be clearly presented in the guideline evidence review documents. This should include a discussion of potential sources of bias and uncertainty. It should also include the results of sensitivity analyses in the consideration of uncertainty, as well as any additional considerations that are thought to be relevant. It should be explicitly stated if economic evidence is not available, or if it is not thought to be relevant to the question.

Recommendations for interventions informed by cost-utility analysis

If there is strong evidence that an intervention dominates the alternatives (that is, it is both more effective and less costly), it should normally be recommended. However, if 1 intervention is more effective but also more costly than another, then the ICER should be considered.

Health effects

The cost per QALY gained should be calculated as the difference in mean cost divided by the difference in mean QALYs for 1 intervention compared with the other.

If 1 intervention appears to be more effective than another, the committee must decide whether it represents reasonable 'value for money' as indicated by the relevant ICER. In doing so, the committee should also refer to the principles that guide the development of NICE guidance and standards (also see box 7.2).

Box 7.2 Comparing the cost effectiveness of different interventions

If possible, we consider value for money by calculating the incremental cost-effectiveness ratio (ICER). This is based on an assessment of the intervention's costs and how much benefit it produces compared with the next best alternative. It is expressed as the 'cost (in £) per quality-adjusted life year (QALY) gained'. This takes into account the 'opportunity cost' of recommending one intervention instead of another, highlighting that there would have been other potential uses of the resource. It includes the needs of other people using services now or in the future who are not known and not represented. The primary consideration underpinning our guidance and standards is the overall population need. This means that sometimes we do not recommend an intervention because it does not provide enough benefit to justify its cost. It also means that we cannot apply the 'rule of rescue', which refers to the desire to help an identifiable person whose life is in danger no matter how much it costs. Sometimes we use other methods if they are more suitable for the evidence available, for example when looking at interventions in public health and social care.

Interventions with an ICER of less than £20,000 per QALY gained are generally considered to be cost effective. Our methods manuals explain when it might be acceptable to recommend an intervention with a higher cost-effectiveness estimate. A different threshold is applied for interventions that meet the criteria to be assessed as a 'highly specialised technology'.

When assessing the cost effectiveness of competing courses of action, the committee should not give particular priority to any intervention or approach that is currently offered. In any situation where 'current practice', compared with an alternative approach, generates an ICER above a level that would normally be considered cost effective, the case for continuing to invest in it should be carefully considered, based on similar levels of evidence and considerations that would apply to an investment decision. The committee should be mindful of whether the intervention is consuming more resource than its value is contributing based on our cost per QALY threshold.

Equity considerations

NICE guideline economic evaluations of healthcare, social care and public health interventions does not include any equity weighting – a QALY has the same weight for all population groups.

The estimation of QALYs implies a particular position regarding the comparison of health

gained between individuals. Normally, an additional QALY is of equal value regardless of other characteristics of the individuals, such as their socio-demographic characteristics, their age, or their level of health.

It is important to recognise that care provision, specifically social care, may be means tested, and that this affects the economic perspective in terms of who bears costs – the public sector or the person using services or their family. Economic evaluation should reflect the intentions of the system.

One of the principles that guide the development of NICE guidance and standards is the aim to reduce health inequalities (principle 9), and that NICE guidance should support strategies that improve population health as a whole, while offering particular benefit to the most disadvantaged. Our recommendations should not be based on evidence of costs and benefit alone, so equity considerations relevant to specific topics, and how these were addressed in economic evaluation, must be reported in the guideline.

Initially, the severity modifiers introduced by the Centre for Health Technology Evaluation (CHTE) for technology appraisal guidance will not be applied to NICE guideline health economic analyses. For NICE guidelines, the severity of the condition should be captured within the QALY benefits and then deliberatively within decision making. However, to enable consistent decision making across NICE guidelines and technology appraisals and to foster better integration of NICE recommendations across these programmes, we are currently exploring approaches on how the severity modifier could be applied within NICE guidelines. We would consult with stakeholders ahead of any implementation.

Considering health inequalities

We recognise the important role NICE guidance can play in the national drive to reduce health inequalities, defined by the UK Government and the NHS as unfair differences in health between more and less socially disadvantaged groups.

To support our commitment to addressing health inequalities, we have commissioned a prototype tool to explore the approach of providing quantitative estimates of the impact of NICE recommendations on health inequalities. The tool uses distributional cost-effectiveness analysis to model changes in health inequalities between 5 socioeconomic groups in England based on the neighbourhood index of multiple deprivation.

We encourage piloting the tool, when data allows, to determine its usefulness in informing

committee consideration of health inequalities during guideline development. Piloting will also enable an exploration of operational considerations, possible trade-offs between cost-effectiveness and health inequality effects, and identify any limitations of the tool.

Non-health effects

Outside the health sector, it is more difficult to judge whether the benefits accruing to the non-health sectors are cost effective, but it may be possible to undertake cost-utility analysis based on measures of social care-related quality of life. The committee should take into account the factors it considers most appropriate when making decisions about recommendations. These could include non-health-related outcomes that are valued by the rest of the public sector, including social care. It is possible that over time, and as the methodology develops (including the establishment of recognised standard measures of utility for social care), there will be more formal methods for assessing cost effectiveness outside the health sector.

Recommendations for interventions informed by cost-effectiveness analysis

If there is strong evidence that an intervention dominates the alternatives, it should normally be recommended. However, if 1 intervention is more effective but also more costly than another, then the ICER should be considered. If 1 intervention appears to be more effective than another, the committee must decide whether it represents reasonable 'value for money' as indicated by the relevant ICER.

The committee should use an established ICER threshold (see the section on recommendations for interventions informed by cost-utility analysis). In the absence of an established threshold, the committee should estimate a threshold it thinks would represent reasonable 'value for money' as shown by the relevant ICER.

The committee should take account of our principles when making its decisions.

Recommendations for interventions informed by cost-consequences analysis

The committee should ensure that, when possible, the different sets of consequences do not double count costs or effects. The way that the sets of consequences have been

implicitly weighted should be recorded as openly, transparently and accurately as possible. Cost-consequences analysis then requires the decision-maker to decide which interventions represent the best value using a systematic and transparent process. Various tools, such as multi-criteria decision analysis (MCDA), are available to support this part of the process. However, attention needs to be given to any weightings used, particularly with reference to our reference case and our principles.

Recommendations for interventions informed by cost–benefit analysis

When considering cost–benefit analysis, the committee should be aware that an aggregate of individual 'willingness to pay' (WTP) is likely to be more than public-sector WTP, sometimes by quite a margin. If a conversion factor has been used to estimate public sector WTP from an aggregate of individual WTP, the committee should take this into account. In the absence of a conversion factor, the committee should consider the possible discrepancy in WTP when making recommendations that rely on a cost–benefit analysis.

The committee should also attempt to determine whether any adjustment should be made to convert 'ability-to-pay' estimates into those that prioritise on the basis of need and the ability of an intervention to meet that need.

The committee should not recommend interventions with an estimated negative net present value (NPV) unless other factors such as social value judgements are likely to outweigh the costs. Given a choice of interventions with positive NPVs, committees should prefer the intervention that maximises the NPV, unless other objectives override the economic loss incurred by choosing an intervention that does not maximise NPV.

Care must be taken with published cost–benefit analyses to ensure that the value of all the health and relevant non-health effects have been included. Older cost–benefit analyses, in particular, often consist of initial costs (called 'costs') and subsequent cost savings (called 'benefits') and fail to include monetarised health effects and all relevant non-health effects.

Recommendations for interventions informed by cost-minimisation analysis

Cost minimisation can be used when the difference in effects between an intervention and

its comparator is known to be small and the cost difference is large (for example, whether doctors or nurses should give routine injections). If it cannot be assumed from prior knowledge that the difference in effects is sufficiently small, ideally the difference should be determined by an equivalence trial, which usually requires a larger sample than a trial to determine superiority or non-inferiority. For this reason, cost-minimisation analysis is only applicable in a relatively small number of cases.

Recommendations when there is no economic evidence

When no relevant published studies are found, and a new economic analysis is not prioritised, the committee should make a qualitative judgement about cost effectiveness by considering potential differences in resource use and cost between the options alongside the results of the review of evidence of effectiveness. This may include considering information about unit costs, which should be presented in the guideline. The committee's considerations when assessing cost effectiveness in the absence of evidence should be explained in the guideline.

Further considerations

Decisions about whether to recommend interventions should not be based on cost effectiveness alone. The committee should also take into account other factors, such as the need to prevent discrimination and to promote equity. The committee should consider trade-offs between efficient and equitable allocations of resources. These factors should be explained in the guideline.

7.9 References

Ara RM, Brazier J, Peasgood T et al. (2017) The identification, review and synthesis of HSUV's from the literature. Pharamcoeconomics 35 (Suppl 1):43–55

Arber M, Garcia S, Veale T et al. (2016) Performance of search filters to identify health state utility studies. Value in Health 19: A390–1

Arber M, Garcia S, Veale T et al. (2017) Performance of Ovid MEDLINE search filters to identify health state utility studies. International Journal of Technology Assessment in Health Care 33: 472–80

Ayiku L, Levay P, Hudson T (2021) The NICE OECD countries geographic search filters: Part 2 – Validation of the MEDLINE and Embase (Ovid) filters. Journal of Medical Library Association 109 No. 2: 258–266

Brazier J, Longworth L (2011) NICE DSU Technical support document 8: An introduction to the measurement and valuation of health for NICE submissions.

Brennan A, Chick SE, Davies R (2006) A taxonomy of model structures for economic evaluation of health technologies. Health Economics 15: 1295–1310

Department for Transport (2019). Transport Analysis Guidance (TAG) unit A2.1

Department for Transport (2021). Transport Analysis Guide (TAG) Unit A1.1

Dias S, Sutton AJ, Welton NJ, Ades AE (2011) NICE Decision Support Unit technical support document 6: Embedding evidence synthesis in probabilistic cost-effectiveness Analysis: software choices

Golder S, Glanville J, Ginnelly L (2005) Populating decision-analytic models: the feasibility and efficiency of database searching for individual parameters. International Journal of Technology Assessment in Health Care 21: 305–11

HM Treasury (2015) The Aqua Book: guidance on producing quality analysis for government. [online; accessed 3 September 2018]

HM Treasury (2013) Review of quality assurance of government analytical models: final report. [online; accessed 3 September 2018]

Husereau D, Drummond M, Augustovski F et al. (2022) Consolidated Health Economic Evaluation Reporting Standards 2022 (CHEERS 2022) Statement: Updated Reporting Guidance for Health Economic Evaluations. Value Health 25(1): 3-9.

Jackson CH, Bojke L, Thompson G et al. (2011) A framework for addressing structural uncertainty in decision models. Medical Decision Making 31: 662–74

Jones K, Burns A. (2021) Unit Costs of Health and Social Care 2021 (Personal Social Services Research Unit [PSSRU]), University of Kent, Canterbury

Kaltenthaler E, Tappenden P, Paisley S (2011) NICE Decision Support Unit technical support document 13: identifying and reviewing evidence to inform the conceptualisation and population of cost-effectiveness models. [online; accessed 3 September 2018]

Latimer N (2013) NICE Decision Support Unit technical support document 14: Survival analysis for economic evaluations alongside clinical trials – extrapolation with patient-level data

Longworth L, Rowen D (2011) NICE Decision Support Unit technical support document 10: the use of mapping methods to estimate health state utility values. [online; accessed 3 September 2018]

National Institute for Health and Care Excellence. The principles that guide the development of NICE guidance and standards

National Institute for Health and Care Excellence (2021) Assessing resource impact process manual: guidelines

National Institute for Health and Care Excellence (2022) NICE real-world evidence framework

NICE Decision Support Unit (2011) Technical support document series [accessed 3 September 2018]

Paisley S (2016) Identification of evidence for key parameters in decision-analytic models of cost-effectiveness: a description of sources and a recommended minimum search requirement. Pharmacoeconomics 34: 597–608

Papaioannou D, Brazier JE, Paisley S (2010) NICE Decision Support Unit technical support document 9: the identification, review and synthesis of health state utility values from the literature. [online; accessed 3 September 2018]

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: 1–Wood H, Arber M, Isojarvi J et al. (2017) Sources used to find studies for systematic review of economic evaluations. Presentation at the HTAi Annual Meeting. Rome, Italy, June 17 to 21 2017

8 Linking to other guidance

This chapter describes how guidelines link to other NICE guidance in the topic area and to guidance from other developers.

Related NICE guidelines and guidance on health technologies should be identified during scoping of a guideline (see the chapter on the scope), or when checking if a guideline needs updating (see the chapter on ensuring that published guidelines are current and accurate). This includes identifying any related guidance or guidelines in development, as well as those that are published.

When a guideline topic covers guidance on health technologies that is published or in development, NICE staff with a quality assurance role work closely with NICE staff from the health technologies team to agree the approach to take.

8.1 Related NICE technology appraisal guidance

A guideline committee cannot usually publish its own recommendations on health technologies covered by published or in development health technologies guidance. In rare cases, technology appraisal guidance will be updated in the guideline (see the section on updating technology appraisal guidance in a guideline).

Technology appraisal being developed or updated alongside the guideline

When a technology appraisal is developed at the same time as a related NICE guideline, development should ideally be coordinated so that the appraisal recommendations are published in time for a link to be included in the consultation draft of the guideline. This is not always possible (for example, if the technology has not yet received a marketing authorisation). In this case, the guideline consultation draft should link to the appraisal consultation document or final appraisal determination. If guideline committee members wish to comment on the appraisal, they do this through the developer (see NICE health technology evaluations: the manual).

The guideline committee chair and the developer may attend relevant appraisal committee meetings as advisers, but do not take part in the formal decision-making.

Updating technology appraisal guidance in a guideline

If there is new evidence for an existing NICE technology appraisal, it can be updated within the context of a guideline, as part of guideline development.

See our NICE health technology evaluations: the manual for more information on how NICE technology appraisals are developed.

Early planning is essential to identify how the guideline developer will update technology appraisal guidance within a guideline. The developer should consider whether there are any data not in the public domain that are likely to be useful. If so, they should call for evidence from registered stakeholders, using the procedures described in the appendix on call for evidence and expert witnesses.

If there is significant new evidence or a change in costs since the technology appraisal guidance was published, the developer assesses cost effectiveness (performs an economic analysis) to determine whether a change in the recommendations is appropriate. In exceptional circumstances, it may not be clear that an economic analysis is needed until the evidence is reviewed and discussed by the committee. Nevertheless, the developer should start planning for any economic analysis at an early stage. The intended approach to cost-effectiveness (economic) analysis for technology appraisal updates should be included in the economic plan and discussed with the committee and a member of NICE staff with responsibility for quality assurance.

The approach should follow the principles described in the chapter on incorporating economic evaluation and should be similar to that used in the technology appraisal. Any differences must be justified by changes in the evidence base or the decision context (for example, a broader range of comparators in the guideline).

The developer may sometimes consider that cost effectiveness can best be assessed by modifying or building on the architecture of an existing economic analysis from the technology appraisal. If so, this should be discussed with a member of NICE staff with responsibility for quality assurance during development of the economic plan.

New medicines

Assessment of new medicines and significant new indications will be carried out in line with section 3.17 of the 2019 voluntary scheme for branded medicines pricing and access

published on the GOV.UK website.

New technology appraisal referral during development of the guideline

When a new related technology appraisal is referred after the guideline has started development, and will be developed alongside the guideline, NICE's technology appraisals team informs the developer and NICE staff with responsibility for guideline quality assurance.

Referring to technology appraisals in recommendations

Please see NICE's interim process and methods statement for bringing together NICE guidance.

Sometimes a guideline covers a medicine for which there is technology appraisal guidance, but for a different population or indication (condition). In these cases, the committee developing the guideline recommendation should assess evidence of effectiveness using methods described in this manual. The guideline recommendations may be different from the technology appraisal recommendations if there is evidence of differing safety or effectiveness for the population or indications covered by the guideline.

8.2 Related NICE guidelines

Related published or in-development NICE guidelines should be identified by the developer at the scoping stage, or by the NICE surveillance programme when checking if a guideline needs updating (see the chapter on ensuring that published guidelines are current and accurate). The scope should document which NICE guidelines are considered relevant for the guideline that is being developed. Any other related NICE guidelines that are highlighted during guideline development should be discussed, and a joint approach agreed with NICE staff with responsibility for quality assurance.

During scoping, identified related guidelines should be reviewed to determine whether the guideline in development will link to existing published recommendations (that is, the key issues will be excluded from the scope of the guideline in development) or whether the guideline in development will consider similar review questions (that is, the key issues are included in the scope of the guideline in development).

When a guideline in development will link to recommendations in a published guideline, this will be specified in the scope.

Similar review question covered in another guideline

When a similar review question is identified in a published NICE guideline, and the evidence review underpinning any recommendations is considered appropriate, the committee developing the guideline can choose to link to the recommendations in the published guideline or to draft new recommendations, based on the evidence review for the published guideline.

In these cases, the committee should discuss and document whether:

- the review question in the guideline in development is similar enough to the question addressed in the published guideline
- the evidence base underpinning any recommendations is unlikely to have changed significantly since the publication of the related guideline
- the evidence review for the review question in the published guideline is relevant and appropriate to the question in the guideline in development.

If the committee agree that the evidence review is relevant for the guideline in development, they consider the recommendations based on the evidence review.

There are 3 options in these circumstances and these are described in the following sections.

Link to the recommendations in the other guideline

If the committee are happy to accept the intent and exact wording of the recommendations, and understand that any future changes (for example, as part of an update) would also apply to their guideline, then the guideline will link to the recommendations in the published guideline. If needed, a recommendation can be added to the guideline explaining anything users will need to take into account when applying the recommendations from the published guideline (for example, if some parts of a recommendation do not apply to the population or setting covered by the new guideline). The committee's discussion and decision should be documented clearly in the guideline. This should include areas of agreement and difference with the committee for the

published guideline (for example, in terms of key considerations – balance of benefits and harms or costs, and interpretation of the evidence).

Linking to other guidelines is preferred to copying recommendations into the guideline in development. In rare cases, recommendations may be copied into the guideline, for example, if this is essential for readability or there is a need to highlight specific recommendations.

Use the evidence review to make new recommendations

If the committee considers that the intent and wording of the published recommendations are not appropriate for their guideline, or they are not prepared to agree that any updates to recommendations would also apply to their guideline, they may make new recommendations based on the existing evidence review. When evidence reviews from another guideline are used to develop new recommendations, the decision should be made clear in the methods section of the guideline, and their independent interpretation and discussion of the evidence should be documented in the discussion section. The evidence reviews from the published guideline (including review protocol, search strategy, evidence tables and full evidence profiles [if available]) should be included in the guideline. They then become part of the whole evidence base for the guideline, and are updated as needed in future updates of the guideline.

Request a new evidence review

If the review question is not considered similar enough or the evidence review is not considered appropriate, the committee may request that a new evidence review is conducted. This should follow NICE's standard processes and methods, as described in this manual. The decision to conduct a new review should be agreed by NICE staff with responsibility for quality assurance, and documented in the methods section and the discussion section for the review question.

8.3 Related interventional procedures guidance

Interventional procedures (IP) guidance differs from other NICE guidance in that it addresses the safety and efficacy of interventions, and not their effectiveness or cost effectiveness. (For more details see the NICE interventional procedures programme manual.)

Any related published IP guidance should be identified during the scoping of a guideline or by the surveillance review. The approach depends on whether the recommendation in the IP guidance is for 'standard' (previously known as 'normal') or 'special' arrangements for clinical governance, consent and audit or research. Because guidelines focus on placing established treatments in the care pathway, they will generally only include IP guidance that recommends 'standard' arrangements.

If IP guidance on a procedure relevant to a guideline starts development during development of the guideline, the IP programme will send the finalised scopes for the procedures to NICE staff with responsibility for quality assurance of the guideline.

IP guidance with recommendations for 'standard' arrangements

Review question is not justified

If the scoping group for a guideline decides that IP guidance for which 'standard' arrangements (previously known as normal arrangements) are recommended is relevant to the guideline but does not justify a review question, the guideline simply links to the IP guidance pages. The developer does not search for new evidence on procedures that are not incorporated into a review question. However, if in the course of their search for evidence for the guideline the developer finds new relevant evidence on that procedure, they inform the IP team at NICE.

Review question is justified

If the scoping group for a new guideline considers that IP guidance with 'standard' arrangements is likely to justify a review question, this is highlighted in the scope for the guideline. For example, comparisons of clinical and cost effectiveness could be included if there is IP guidance covering several procedures for the same condition. The IP team is contacted by NICE staff with responsibility for guideline quality assurance during scoping to seek their input on the proposed approach.

If the procedure is included in the final scope of the guideline, the developer considers its effectiveness using standard methods for guideline development (see the chapters on reviewing evidence and incorporating economic evaluation).

If a guideline is being updated and the surveillance review identifies that IP guidance with 'standard' arrangements is likely to justify a review question in the updated guideline, this

is highlighted in the surveillance report.

When a review question in a guideline relates to IP guidance, the IP guidance remains active unless the IP team and NICE staff with responsibility for guideline quality assurance have identified reasons why not. This is because the IP guidance relates to the efficacy and safety of the procedure, and so the guideline and the IP guidance address different questions. Therefore IP guidance remains current even if the recommendations are supplemented by a guideline recommendation on the effectiveness of a procedure for 1 or more indications. IP guidance may also contain more detailed information that may be of value to practitioners and people considering having the procedure. Importantly, the IP guidance may also specify conditions for use of the procedure; for example, that the surgeon should have training, or that the procedure should be carried out within the context of a multidisciplinary team. The guideline will link to the IP guidance pages.

Other approaches

If NICE staff with responsibility for guideline quality assurance and the IP team agree that a different approach is needed, they present a paper to NICE's guidance executive. Decisions to take other approaches are made on a case-by-case basis. Examples include:

- Appraisal of the evidence indicates that a procedure with IP guidance is not effective, and the guideline recommends that it should not be used.
- There is considerable uncertainty about the clinical or cost effectiveness of a procedure.
- The guideline committee makes a 'research only' recommendation for a procedure with published IP recommendations for 'standard' arrangements.

IP guidance with recommendations for 'special' arrangements

If the guideline committee considers that a procedure with recommendations for 'special' arrangements has become part of NHS practice and falls into the area of a review question, the IP team is notified for potential review of the IP guidance. If the guidance is reviewed and the procedure's status is changed to 'standard' arrangements, the developer considers the procedure's effectiveness and cost effectiveness. If the procedure retains its 'special' arrangements status, the guideline should link to the IP guidance pages.

IP guidance with recommendations for 'research only' or 'do not use'

Sometimes IP guidance recommends that a procedure should only be carried out in research, or that it should not be used. A recommendation not to use a procedure is made if there is no evidence of efficacy and/or safety, or evidence of a lack of efficacy and/or safety. A 'research only' recommendation is made if the evidence shows that there are important uncertainties. The evidence base for procedures with recommendations for 'research only' or 'do not use' usually reflects the fact that they are not established procedures. As such, they would not normally form part of a review question in a guideline.

Developing a guideline and IP guidance concurrently

When a newly notified procedure has been scoped and it has been agreed that it will be assessed by the IP team, the IP team informs the developer and NICE staff with responsibility for guideline quality assurance that the notified procedure is relevant to the guideline. The guideline can link to the IP guidance when it is published.

8.4 Guidance from other developers

Sometimes relevant guidelines published by organisations other than NICE are identified in the search for evidence and considered appropriate for inclusion in the evidence base. Intellectual property issues must be considered when using content from other organisations. Published guidelines are assessed for quality using the AGREE II instrument. There is no cutoff point for accepting or rejecting a guideline, and each committee needs to set its own parameters. These should be documented in the methods of the guideline, and a summary of the assessment included in the evidence review. The full results of the assessment should be presented with the guideline. The evidence reviews in the guideline need to be critically appraised (using checklists in the appendix on appraisal checklists, evidence tables, GRADE and economic profiles) to ensure that the quality of the evidence is appropriate for deriving NICE recommendations.

Reviews of evidence from guidelines that directly cover review questions being addressed by the committee may be considered as evidence if they:

- are assessed as being of high quality using the appropriate methodology checklist (see the appendix on appraisal checklists, evidence tables, GRADE and economic profiles)

- are accompanied by a GRADE table, GRADE-CERQual table or evidence statement and evidence tables
- are assessed as being sufficiently up to date
- do not have any intellectual property restrictions that would make it impossible to follow NICE's normal guideline development process.

If using evidence from published guidelines, the guideline development team should create its own evidence summaries or statements (see the section on summarising evidence in the chapter on reviewing evidence). Evidence tables from guidelines published by other organisations should be referenced with a direct link to the source website or a full reference of the published document. The committee should formulate its own recommendations, taking into consideration the whole body of evidence.

Recommendations from guidelines published by another organisation should not be quoted or linked to, unless the guideline has been assessed for quality using the AGREE II instrument and the evidence reviews have been critically appraised to verify quality.

8.5 References and further reading

Brouwers MC, Kho ME, Browman GP et al. (2010) AGREE II: advancing guideline development, reporting and evaluation in health care. Canadian Medical Association Journal 182: E839–42

9 Interpreting the evidence and writing the guideline

This chapter gives guidance on how the committee should interpret the evidence and decide what recommendations to make. It also gives some advice on how to word guideline recommendations, although content designers will help committees with this.

9.1 Interpreting the evidence to make recommendations

Assessment and interpretation of the evidence to inform guideline recommendations is at the heart of the work of the committee.

Recommendations are developed using a range of evidence from the literature searches and other evidence – such as real world data and expert testimony (see the appendix on call for evidence and expert witnesses), views of stakeholders, people using services and practitioners, health inequalities briefings (if available) and the committee's discussions and debate (see the chapter on decision-making committees).

The committee should use its judgement to decide what the evidence means in the context of the scope of the guideline or area(s) for update. The quality of the evidence will have been assessed for both internal and external validity (see the chapter on reviewing evidence), but also needs interpretation. If a conceptual framework or logic model is being used to develop the guidance, the committee should consider this when interpreting the evidence.

The committee should decide what action to recommend and keep in mind which sectors (including which practitioners or commissioners within those sectors) should act on the recommendations. This will identify the likely impact of the recommendations on practice or services and the committee can decide whether to stipulate who the recommendation is aimed at.

The committee should discuss how they moved from the evidence to each recommendation, including the relative value placed on the agreed outcomes, the benefits and harms of any interventions, resource use, and the overall quality of the evidence, as

well as other factors they took into consideration.

For each recommendation or group of recommendations, the committee should discuss and agree their rationale for making the recommendations and the likely impact of the recommendations on practice or services. They should also discuss how the recommendations address any equality issues or health inequalities identified during the guideline development process.

Quality of the evidence

Evidence review documents summarise the evidence obtained from the results of evidence searches. Depending on the topic and type of evidence, they may include GRADE tables, GRADE-CERQual tables or (if GRADE or GRADE-CERQual are not used) evidence statements.

The committee should ensure that the reviews are a fair summary of the evidence and should discuss any uncertainty in the review findings (including limitations of individual studies and inconsistency across studies).

For details, see the chapter on reviewing evidence.

Trade-off between benefits and harms of an intervention

A key stage in moving from evidence to recommendations is weighing up the size and importance of the benefits and harms of an intervention, compared with the alternatives specified in the review protocol, and the potential for unintended consequences. This may be done qualitatively or quantitatively.

The committee should discuss the extent to which the effects seen in the evidence are representative of what would happen in the real world.

The committee should also assess the extent to which the recommendations may impact on health inequalities. This needs to be made clear, regardless of whether the recommendation is aimed at the whole population, specific subgroups or a combination of both. If there is potential to increase health inequalities, the committee should consider whether they can do anything to prevent this from happening or reduce the impact.

Trade-off between economic considerations and resource use

The committee should discuss cost effectiveness at the same time as effectiveness when formulating recommendations. Interventions that are not considered cost effective should not usually be offered.

The evidence review document should explain how costs, resource use and economic considerations were taken into account in determining the cost effectiveness of an intervention. This may be informal, or may be more formal and include economic modelling (see the chapter on incorporating economic evaluation).

If several possible interventions are being considered, the committee should consider sequencing them in terms of their cost effectiveness. This usually means preferring the most cost-effective intervention first, although other factors such as availability and acceptability need to be considered.

Considerations about equity may also affect the decision whether to recommend the intervention (see the section on equity considerations in the chapter on incorporating economic evaluations).

Use of indirect evidence

Sometimes, when there is no evidence directly relevant to a specific population, indirect evidence from other populations may be considered. For example, evidence on treating absence seizures in children and young people was extrapolated to adults because the disease has a similar pathophysiology in all 3 populations.

This needs careful consideration by the committee, with discussion of the features of the condition or interventions that allow extrapolation to a different population.

This also applies when extrapolating findings from evidence in different care settings (for example, between primary and secondary care). The committee should consider the similarities and differences in case mix, staffing, facilities and processes between the settings before extrapolating evidence in this way.

Consider the feasibility of putting recommendations into practice

The committee should judge to what extent it will be feasible to put the recommendations

into practice.

The committee should consider the extent of change in practice that will be needed to implement a recommendation, staff training needs, policy levers and funding streams, and the possible need for carefully controlled implementation with, for example, training programmes. This should be documented in the guideline and in any resources to support implementation (see the chapter on support for putting the guideline recommendations into practice).

Wider basis for making recommendations

The committee should take into account a range of issues, including any ethical issues, equity considerations, health inequalities and national priorities for health and care, as well as equality legislation, to ensure that the guideline recommendations are ethical, practical and specific.

There are no hard-and-fast rules or mechanisms for doing this: the committee should make conscious and explicit use of its members' skills and expertise. All evidence needs interpretation: evidence alone cannot determine the content of a recommendation.

Developing evidence-based recommendations involves:

- using what is known (inductive reasoning) while accepting that there is uncertainty about what is likely to happen because of implementing a recommendation
- drawing on theory or methodological principles (deductive reasoning).

Alongside this manual, committees should use our principles and our charter to inform their decisions. The committee may also draw on the principles outlined in the report on ethical issues in public health by the Nuffield Council on Bioethics.

Promoting equality and reducing health inequalities

The equality and health inequalities assessment (EHIA) form should document how the committee's responsibilities under equality legislation and our equality scheme have been discharged in reaching the recommendations (see the section on key principles for developing NICE guideline recommendations in the introduction chapter), and how the recommendations address equality issues and health inequalities.

The committee needs to consider whether:

- the evidence review has found evidence to support recommendations to address any equality issues and health inequalities identified during guideline development (if not, consider other sources of information for example expert testimony or health inequality briefings, if available)
- criteria for access to an intervention might be discriminatory (for example, through membership of a particular group, or by using an assessment tool that might discriminate unlawfully)
- any groups of people might find it impossible or difficult to receive or access an intervention.

Ideally, recommendations should be formulated to promote equality and reduce health inequalities (for example, by making access more likely for certain groups, or by tailoring the intervention to specific groups). If this is not possible, the committee should consider whether it is appropriate to make a research recommendation (for further details see the section on formulating recommendations for research).

Strength of recommendations

The concept of the 'strength' of a recommendation (Guyatt et al. 2008) is key to translating evidence into recommendations. This takes into account the quality of the evidence but is conceptually different.

If the committee believes that the vast majority of practitioners or commissioners and people using services would, based on the evidence seen by the committee, choose a particular intervention, they should make a strong recommendation for the intervention. This is generally the case if the benefits clearly outweigh the harms for most people and the intervention is likely to be cost effective. If the opposite is true, they should make a strong recommendation against the intervention (see below).

If the committee concludes, based on the evidence, that there is a closer balance between benefits and harms, and some people would not choose an intervention whereas others would, they should make a weak recommendation for the intervention.

If a specific subgroup of people is likely to benefit from an intervention and others are not, it may be possible to make a strong recommendation for that subgroup of people.

The committee should be aware that we reflect the strength of the recommendation in the wording (see the section on wording the recommendations).

Strong recommendations against an intervention

Reasons for the committee to make a strong recommendation against an intervention include:

- potential harms outweigh the potential benefits
- the intervention has no reasonable prospect of providing cost-effective benefits
- stopping the intervention is not likely to cause harm
- good-quality evidence shows a lack of efficacy or effectiveness
- there is a lack of evidence of efficacy or effectiveness for an intervention, or the quality of the evidence is too low or too uncertain
- the intervention has a major national safety warning, for example from the Medicines and Healthcare products Regulatory Agency (MHRA) or NHS England.

If most people are likely to experience no benefit or to experience harm but there may be a benefit for some, the committee can make a strong recommendation against the intervention but with a caveat. In this case, they should be as specific as possible about the circumstances under which, or population for whom, the intervention is appropriate.

'Only in research' recommendations

The committee can make an 'only in research' recommendation if the necessary research can realistically be set up or is already planned, or people using services are already being recruited for a study. The following criteria may also apply:

- the intervention has a reasonable prospect of providing cost-effective benefits
- there is a real prospect that the research will inform future NICE guidelines.

Little evidence of difference between interventions

There might be little evidence of differences in effectiveness or cost effectiveness

between interventions. In this case, all effective or cost-effective interventions may be recommended.

Insufficient evidence

If published evidence of efficacy or effectiveness for an intervention is lacking, too low quality, or too uncertain for firm conclusions to be reached, the committee may use its experience and knowledge to do 1 of the following:

- make recommendations by consensus
- make research recommendations, or
- make no recommendation.

The last option should be used sparingly on the basis that scoping will have shown that guidance was needed.

The principles in the section on wording the recommendations should be used.

Recording the committee's discussion and rationale for the recommendations

The committee's justifications for making the recommendation, and its strength, should be summarised in the rationale for the recommendation and fully explained in the committee discussion section of the relevant evidence review document.

The committee discussion follows a structured format, to ensure transparency about the issues considered. In most cases the committee reaches decisions through a process of informal consensus. If formal voting procedures are used, this is also recorded.

Principles of person-centred care

All NICE guidelines advocate the principles of person-centred care: people using services and the wider public should be informed of their options and be involved in decisions about their care, as described in our webpage on making decisions about your care.

The following foundational guidelines include general recommendations on the principles of person-centred care, such as communication, providing information and shared

decision making, which should not be restated in topic-specific guidelines (see table 1):

- patient experience in adult NHS services
- babies, children and young people's experience of healthcare
- service user experience in adult mental health
- people's experience in adult social care services
- shared decision making
- medicines adherence
- medicines optimisation
- multimorbidity
- transition from children's to adults' services
- decision making and mental capacity.

Recommendations from these guidelines can be cross-referred to when needed.

Recommendations on person-centred care can be included in topic-specific guidelines if there is evidence of specific need for the topic.

Topic areas we do not usually make recommendations on

Table 1: Topic areas we do not usually make recommendations on

Topic area	What to do instead	Exceptions
General principles of care covered in foundational guidelines	Link to the relevant foundational guideline: patient experience in adult NHS services service user experience in adult mental health people's experience in adult social care services shared decision making babies, children and young people's experience of healthcare	If there is evidence of issues specific to the topic of the guideline recommendations
Repeating recommendations from another NICE guideline	Link to the other guideline (see the chapter on linking to other guidance)	If linking between specific recommendations would be cumbersome for users
Recommendations on general lifestyle advice	Link to relevant public health guidelines	If there is evidence and a strong rationale to include a recommendation specific to the topic of the guideline recommendations

Topic area	What to do instead	Exceptions
Recommendations on good practice or general principles of care that are not linked to review questions or evidence	Do not include	If there is evidence and a strong rationale to include a recommendation specific to the topic of the guideline recommendations
Recommendations on general medicines optimisation issues	Link to the relevant guideline, for example: medicines adherence medicines optimisation controlled drugs	If there is evidence and a strong rationale to include a recommendation specific to the topic of the guideline recommendations
Prescribing information covered by the BNF (for example, dosing, monitoring, adverse effects, contraindications)	Nothing – this is covered by the BNF	If the recommendation would not make sense without reference to the prescribing information
National patient safety advice on medicines and devices	The NICE medicines adviser will work with the MHRA if needed	If there is a significant safety risk and clear evidence that safety advice is not routinely implemented in practice, if the recommendation will not make sense without the information
Training or competency in areas that are the responsibility of professional bodies	The implementation team can work with professional bodies to identify training needs	Training or competency in areas that are not the responsibility of professional bodies, and are identified as being important to cover in guideline recommendations
Service configuration or service delivery	Do not include	Recommendations on service delivery or service configuration that are evidence based or address system priorities

Topic area	What to do instead	Exceptions
Following laws or statutory guidance	Do not include	If there is evidence that guidance is needed on how to follow the law or statutory guidance

9.2 Recommendations on medicines

When making decisions about treatment options, users of our guidelines are expected to take note of prescribing information, such as dosage, duration of treatment, contraindications, adverse effects, warnings, safety advice and any monitoring requirements for a medicine. This is available in the British National Formulary (BNF) or BNF for Children (BNFC), as well as the medicine's summary of product characteristics (available on the electronic medicines compendium). We do not usually include prescribing information in our recommendations though there are some exceptions to this (see table 1 and the section on off-label use of licensed medicines). For more information on prescribing, see our webpage on making decisions using NICE guidelines.

Overprescribing

Overprescribing is when people are given medicines they do not need or want, or where the harm outweighs the benefits. In line with the Department of Health and Social Care's national overprescribing review, we should include recommendations for reviewing and stopping medicines if overprescribing is a concern. Also see NICE's guidelines on medicines optimisation and medicines associated with dependence or withdrawal symptoms.

National medicines safety advice

National medicines safety advice includes national patient safety alerts and the MHRA's drug safety updates. We do not usually include patient safety information in our recommendations though there are some exceptions to this. See table 1 for details.

Antimicrobials and antimicrobial stewardship

Recommendations on antimicrobials should:

- take account of antimicrobial resistance and the principles of good antimicrobial

stewardship

- name the specific antibiotic or class of antibiotics being recommended
- include information on reviewing and stepping down treatment when recommending intravenous or prophylactic antibiotics.

Guidelines that cover antimicrobial prescribing may include prescribing tables that detail choice of antimicrobials, dosages, duration of treatment and routes of administration (for an example of an antimicrobial prescribing table, see the section on choice of antibiotic in NICE's guideline on Clostridioides difficile infection).

Off-label use of licensed medicines

Recommendations are usually about using medicines within their licensed indications. However, there are clinical situations in which recommending an off-label use of a licensed medicine may be in the best interests of the person, in line with the MHRA guidance (see appendix 2 of the MHRA guidance on the supply of unlicensed medicinal products). For example, this may happen if the clinical need cannot be met by using a licensed product within the terms of its marketing authorisation and there is enough evidence or experience of using the medicine to support its safety and efficacy.

Dosage information for off-label use of a licensed medicine is not usually included in the summary of product characteristics (SPC). If off-label use is being recommended, we will check whether there is any relevant dosage information in the BNF or BNF for Children. If there is none, we will work with the BNF to add the necessary information if needed. If we are recommending off-label use of a medicine, this will be clearly stated in the guideline.

Unlicensed medicines

The MHRA states that: If a UK licensed medicine can meet the person's clinical need (even if it is used off-label), it should be recommended instead of an unlicensed product. An unlicensed medicine should not be recommended if a product available and licensed within the UK could be used to meet the person's clinical need.

Committees should take account of the MHRA guidance (see appendix 2 of the MHRA guidance on the supply of unlicensed medicinal products) when making recommendations but consider each situation on its own merit.

Medical devices, including off-label use

Recommendations are usually about using devices within the terms of the instructions for their use. However, there are clinical situations in which the off-label use of a device may be in the best interests of the person. For example, when using a device outside the time period specified in the instructions for use.

Committees should take account of the MHRA guidance on the off-label use of medical devices.

9.3 Wording the recommendations

This section gives the key principles for writing recommendations. Following these principles helps ensure that recommendations meet user needs. The content designer works with the rest of the development team and committee throughout guideline development to ensure that recommendation wording reflects the committee's intent and is clear and easy to follow.

The recommendations should be in line with our style and principles, and accessibility regulations.

For information on NICE style, and using clear English and person-centred language, see our style guide and guide on writing for NICE.

For information on accessibility, see our webpages on accessibility and accessibility changes: notes for developers.

Focus on the action and what readers need to know

Recommendations should be clear about what needs to be done, without the reader having to read the rationale or committee's discussion in the evidence review document. When writing recommendations, keep in mind the following:

- a reader asking, 'What does this mean for me?'
- how a health and care professional will be able to implement them with an individual person, in a way that supports shared decision making.

Include only 1 action per recommendation or bullet point unless it is clearer to include a closely linked action in the same recommendation.

Be specific about actions and use direct instructions in recommendations wherever possible because these are easier to follow. Recommendations should start with a verb such as 'offer' (or 'do not'), 'consider', 'measure', 'advise', 'discuss', 'ask about'.

Exceptions to this principle include:

- Recommendations that specify who should take action, or cover service organisation. For example: A multidisciplinary team should provide care.
- Recommendations that use 'must' or 'must not' (because of a legal duty or serious consequences of not following the recommendation).

Think carefully about how much detail to include. Recommendations should be clear and concise. Including a lot of detail can reduce the impact and make them harder to understand.

Reflect the strength of the recommendation

In 'strong' recommendations (see the section on strength or recommendations) for actions that should (or should not) be offered, use directive language such as 'offer' (or 'do not offer'), 'advise', or 'ask about'. In keeping with the principles of shared decision making, people may choose whether or not to accept what they are offered or advised.

If there is a closer balance between benefits and harms (activities or interventions that could be used), use 'consider' to reflect that the recommendation is 'weak'.

Use 'person-centred', precise, concise, clear English

Key principles include using language that is person-centred, using clear and consistent wording, and using bullet lists and tables if they make recommendations easier to follow.

Language that is person-centred acknowledges the experience of people who are directly affected by the recommendations (and family members, carers or advocates), and their role in decision making. For more information, see the section on talking about people in the NICE style guide.

9.4 Supporting shared decision making

Identify preference-sensitive decision points

Guidelines should be written to support shared decision making between people and their health or social care practitioners (see the recommendations on supporting people to make decisions about their care in NICE's guidelines on shared decision making, patient experience in adult NHS services, service user experience in adult mental health, people's experience in adult social care services, multimorbidity and babies, children and young people's experience of healthcare).

The committee should identify recommendations where someone's values and preferences are likely to be particularly important in their decision about the best course of action for them.

These 'highly preference-sensitive decision points' occur when the committee recommends 2 or more options that deliver similar outcomes but have different types of harms and benefits or different practicalities (such as a choice between medicine and surgery, or differing burden of treatment) that people may value differently.

Alternatively, a highly preference-sensitive decision point may occur if the choice between 1 or more investigation, treatment or care option and 'doing nothing new or different' is finely balanced.

These decision points may be identified as early as the guideline scoping stage, or when the committee reviews the evidence.

Summarise information to support decisions

When a highly preference-sensitive decision point is identified, create a summary of the evidence to make it easy for professionals and practitioners to discuss the options with the person.

Base the summary on the evidence review documents underpinning the recommendations, and explain the benefits, risks, alternative options, and what might happen if the person decides not to have the intervention. The BRAN format is an example of how to do this:

- **benefits** of each recommended option
- **risks** and consequences of each option (including adverse effects and consequences of treatment such as the need for regular monitoring with warfarin, or implications for driving with insulin treatment)
- **alternatives** to the main option(s)
- option of doing **nothing** new or different – what might happen if I decide against the option(s) and remain on my current treatment (if any).

Medicines advisers can help development teams with questions such as how to apply BRAN to a particular decision point, and how much information to include on adverse effects of treatments.

Occasionally, we will develop an additional decision aid (see the chapter on support for putting the guideline recommendations into practice and the NICE decision aid process guide).

9.5 Formulating recommendations for research

The committee is likely to identify areas for which there are uncertainties or for which robust evidence is lacking. They can suggest up to 5 key recommendations for research that are likely to inform future decision-making (based on a systematic assessment of gaps in the current evidence). The committee should justify and document why they have made each recommendation for research. For further information, see the NICE research recommendations process and methods guide.

9.6 References and further reading

Alonso-Coello P, Oxman AD, Moberg J et al. for the GRADE working group (2016) GRADE Evidence to Decision (EtD) frameworks: a systematic and transparent approach to making well informed healthcare choices. 2: Clinical practice guidelines. BMJ 353: i2089

Claxton K, Sculpher MJ (2006) Using value of information analysis to prioritise health research: some lessons from recent UK experience. Pharmacoeconomics 24: 1055–68

Glasziou P, Del Mar C, Salisbury J (2003) Evidence-based medicine workbook. London: British Medical Journal Books

Guyatt GH, Oxman AD, Vist GE et al. for the GRADE working group (2008) GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. BMJ 336: 924 (see also the GRADE website)

Joint Royal College of Paediatrics and Child Health/Neonatal and Paediatric Pharmacists Group Standing Committee on Medicines (2013) The use of unlicensed medicines or licensed medicines for unlicensed applications in paediatric practice

Kelly MP, Moore TA (2012) The judgement process in evidence-based medicine and health technology assessment. Social Theory and Health 10:1–19

Michie S, Johnston M (2004) Changing clinical behaviour by making guidelines specific. British Medical Journal 328: 343–5

Nuffield Council on Bioethics (2007) Public health: ethical issues. London: Nuffield Council on Bioethics

Sackett DL, Straus SE, Richardson WS (2000) Evidence-based medicine: how to practice and teach EBM. Edinburgh: Churchill Livingstone

Schünemann HJ, Mustafa R, Brozek J et al. for the GRADE working group (2016) GRADE Guidelines: 16. GRADE evidence to decision frameworks for tests in clinical practice and public health. Journal of Clinical Epidemiology 76: 89–98

Scottish Intercollegiate Guidelines Network (2019) SIGN 50. A guideline developer's handbook, revised edition. Edinburgh: Scottish Intercollegiate Guidelines Network

Tannahill A (2008) Beyond evidence – to ethics: a decision making framework for health promotion, public health and health improvement. Health Promotion International 23: 380–90

Weightman A, Ellis S, Cullum A et al. (2005) Grading evidence and recommendations for public health interventions: developing and piloting a framework. London: Health Development Agency

10 The validation process for draft guidelines, and dealing with stakeholder comments

Consultation with stakeholders is an integral part of the guideline development process. Comments received from registered stakeholders are a vital part of the quality-assurance and peer-review processes, and it is important that they are addressed appropriately. Registered stakeholders are notified of the consultation dates in advance via the guideline page on the NICE website, and are reminded by email. (See NICE's webpage on registering as a stakeholder for more information.)

This chapter describes the validation process for draft guidelines. It includes information on what happens during the consultation, the principles of responding to stakeholder comments after the consultation and when a second consultation may be needed.

Before the draft guideline is signed off for consultation, an equality impact assessment is completed by the developer and the committee chair to show which equality issues have been identified and considered during guideline development. The equality impact assessment is signed off by a member of NICE staff with responsibly for quality assurance, and published on the NICE website with the draft guideline. The assessment is updated by the developer and the committee chair after the consultation.

10.1 What happens during consultation

Commenting on the draft guideline

The draft version of the guideline (recommendations, rationales, committee discussions, evidence reviews and methods) is posted on the NICE website for consultation with registered stakeholders. Stakeholders can register at any point during guideline development. NICE informs registered stakeholders that the draft is available, via email and through its promotional channels, and invites them to comment by the deadline. Questions for stakeholders are posted with the draft guideline. The purpose of these questions is to seek stakeholder views on factors such as the potential equality impact. NICE also asks stakeholders to comment on recommendations identified as likely to

substantially increase costs, and their justification, and to consider whether any other draft recommendations are expected to add substantial costs. Questions related to implementation may also be included to identify practitioners or services who are already implementing the draft recommendations, or resources that may support implementation.

The length of time for consultation depends on the size of the guideline and the number of review questions. Consultation on a new guideline or full update consisting of 15 to 20 review questions usually lasts for 6 weeks. A 4-week consultation may be used for partial updates of guidelines with less than 15 review questions, while small updates with 1 or 2 review questions will normally have a 2-week consultation. NICE staff with responsibility for quality assurance will decide how long the consultation will last, and stakeholders will be told well in advance.

NICE is unable to accept:

- more than 1 set of comments from each registered stakeholder organisation
- comments that are not presented correctly on the form provided
- comments with attachments such as research articles, letters or leaflets.

In these cases, NICE will invite a registered stakeholder to resubmit a single set of comments with no attachments before the consultation deadline. NICE is unable to accept any comments received after the deadline.

Comments should be constructed as reasoned argument and be submitted for the purpose of improving the draft guideline. NICE reserves the right not to respond to comments that are hostile or inappropriate.

Stakeholders should make sure that any confidential information or information that the owner would not wish to be made public is clearly underlined and highlighted. Confidential information should be kept to a minimum. Stakeholders should explain why the information is confidential and if and when it will become publicly available.

Where views on the guideline are shared by more than 1 stakeholder organisation, NICE encourages these organisations to work together to produce a joint response. This should be submitted by 1 registered stakeholder; other stakeholders supporting the joint response should respond to the consultation noting their endorsement.

When registering, and when commenting on the draft scope and draft guideline, stakeholders are asked to disclose whether their organisation has any direct or indirect links to, or receives or has ever received funding from, the tobacco industry. Disclosures will be included with the published consultation responses. Tobacco companies and those who speak for them or are funded by them (collectively referred to as 'tobacco organisations') cannot register as stakeholders. This is in line with NICE's obligation under Article 5.3 of the WHO Framework Convention on Tobacco Control (FCTC) to protect public health policies from the commercial and other vested interests of the tobacco industry. Tobacco organisations are simply referred to as 'respondents' and any comments received during consultation are reviewed for factual inaccuracy claims and are made public along with any responses.

Approaches to additional consultation and commissioned primary research

Additional consultation is a targeted engagement exercise to obtain a range of views, experiences and expertise, independent from the committee. An additional consultation for a guideline is considered only on an exceptional basis and is additional to the routine stakeholder consultations.

Commissioned primary research can occasionally be useful for addressing topic areas where there is limited or no evidence. Commissioned primary research should start during the development of the guideline, and the findings should be used by the committee as part of the evidence base for developing recommendations.

For more information, see the appendix on approaches to additional consultation and commissioned primary research.

External expert peer review

Although NICE does not routinely commission peer review from external experts, members of NICE staff with responsibility for quality assurance, or the developer, may occasionally consider arranging additional external expert peer review of part or all of a guideline, or an evidence review, executable model or economic analysis. For example, review by external experts may be valuable if novel methods have been used in developing an evidence review.

External expert peer reviewers may include practitioners, those commissioning care,

academics (for example, with expertise in economic or meta-analysis), or people with a lay perspective. Experts are selected on the basis of their experience in the particular issue under review.

External expert peer review may take place during guideline development or during consultation on the draft guideline. If it occurs during development the comments are not published, but the reviewers should be named in the guideline. Comments from external expert peer reviewers during the development of the guideline should be discussed by the committee, and their decisions fully documented in the committee discussion section of the relevant evidence review. If the reviewers also comment during consultation, their comments are responded to in the same way as comments from registered stakeholders and are published in the guideline consultation table on the NICE website under 'external expert peer reviewers'. All external expert peer reviewers are required to complete a declaration of interests form (see the section on code of conduct and declaration of interests in the chapter on decision-making committees).

10.2 Principles of responding to stakeholder comments

After consultation the committee discusses the comments received during consultation, proposes any changes needed to the guideline, and agrees the final wording of the recommendations.

This section describes how developers should respond to consultation comments. The same principles apply when responding to comments on the draft scope (see the chapter on the scope).

Developers must take the following key points into account when responding to comments from registered stakeholders:

- Each comment must be acknowledged and answered as directly, fully and with as much information as possible.
- For a draft guideline, the committee must consider whether changes to the guideline are needed as a result of consultation comments; any changes to the guideline must be agreed by the committee before publication.
- If changes are made to a guideline as a result of a consultation comment, this must be

made clear in the response to the comment. If no changes have been made, it should be clear from the response why not.

- Developers should maintain an audit trail of any changes made to the guideline.

Registered stakeholders who have commented on the draft guideline may be sent the final guideline, and comments and responses, in confidence before publication (see the chapter on finalising and publishing the guideline recommendations). Comments and responses are made available on the NICE website when the final guideline is published.

NICE reserves the right to summarise and edit comments received during consultations, or not to publish them at all, if we consider the comments are too long, or publication would be unlawful or otherwise inappropriate.

Comments received from non-registered stakeholders and individuals are reviewed to decide whether they have raised any key issues that the committee need to consider. A formal response is not given and these comments are not made available on the NICE website. Comments received from 'respondents' are reviewed for factual inaccuracy claims and are made public along with any responses.

Comments received after the deadline are not considered and are not responded to; in such cases the sender will be informed.

When evidence is highlighted by stakeholders during consultation, this should be considered for inclusion in the guideline. The developer will take the evidence into account:

- if it meets all of the inclusion criteria for the relevant review (as set out in the review protocol), and should have been identified in the guideline searches/screening
- if it falls within the timeframe for the guideline search parameters.

Any effects on the guideline of including new evidence will be considered, and any further action agreed between the developer and NICE staff with responsibility for quality assurance.

If the new evidence falls outside of the timeframe for the guideline searches, the impact on the guideline will still need to be considered, and any further action agreed between the developer and NICE staff with responsibility for quality assurance.

10.3 When a second consultation may be needed

In exceptional circumstances, NICE may consider the need for a further stakeholder consultation after the first consultation. This additional consultation may be needed if either:

- information or data that would significantly alter the guideline were omitted from the first draft **or**
- evidence was misinterpreted in the first draft and the amended interpretation significantly alters the draft recommendations.

NICE staff with responsibility for guideline quality assurance make the final decision on whether to hold a second consultation, how long it should be.

11 Finalising and publishing the guideline recommendations

The guideline recommendations, rationales and evidence reviews are quality assured, as described in the chapter on the validation process for draft guidelines, and dealing with stakeholder comments. This chapter covers the sign-off process, publication and promoting guideline awareness through press and other communication events and channels.

11.1 Equality and health inequalities assessment

After consultation, and throughout the work done until the finalised guideline recommendations are submitted for sign-off, the equality and health inequality assessment is updated by the development team and the committee to show whether any additional equality issues have been identified, and how these have been addressed. The equality and health inequality assessment is also quality assured and signed off by relevant staff as specified in the equality and health inequalities assessment form. The assessment is then published on the NICE website with the final guideline recommendations.

11.2 Signing off the guideline recommendations

The guidance executive, made up of executive directors, guidance centre directors and senior team members, considers and approves guideline recommendations for publication on behalf of the Board. The guidance executive reviews a report from staff with responsibility for guideline quality assurance. The report details whether the guideline recommendations:

- addresses all the issues identified in the scope
- is consistent with the evidence
- was developed using the agreed processes and methods
- was developed with due regard to the need to eliminate discrimination, advance equality and foster good relations (see section above on equality and health

inequalities assessment)

- will lead to a resource impact when implemented.

When assessing the report, the guidance executive may ask the development team to answer specific queries. The guidance executive may also refer any matter to the executive team or Board for resolution if it considers a matter to be of particular significance or concern.

11.3 Embargoed release (releasing an advance copy to stakeholders)

Registered stakeholders who have commented on the draft guideline recommendations (see the chapter on the validation process for draft guidelines, and dealing with stakeholder comments), and agreed to conditions of confidentiality, may be sent the final guideline recommendations, the evidence reviews and a copy of the responses to stakeholder consultation comments 2 weeks before publication. This information is confidential until the guideline recommendations are published. This step allows registered stakeholders to highlight to NICE any substantive errors, and to prepare for publication and implementation. It is not an opportunity to comment further on the guideline recommendations. NICE should be notified of any substantive errors at least 1 week before publication of the guideline recommendations.

11.4 Publication

The guideline recommendations and its evidence reviews, methods, equality and health inequality assessment, responses to stakeholder comments, and support tools (see the chapter on support for putting the guideline recommendations into practice) are usually published at the same time.

11.5 Promoting awareness of the guideline recommendations

The development team and committee work with the media relations team and with the implementation lead, where relevant, to disseminate and promote awareness of the guideline recommendations from the time of publication. Each topic area is different and activities for raising awareness will vary depending on the type and content of the

guideline recommendations.

We use a range of methods to raise awareness of the guideline recommendations, including:

- notifying registered stakeholders of guideline recommendation publication
- publishing news articles, blogs, newsletters and alerts
- issuing a press release
- using social media channels
- promoting the guideline recommendations within NICE.

We may also use other means of raising awareness of the guideline recommendations – for example, training programmes, conferences, or implementation workshops.

Press and other communication events

A structured and considered exchange of information between NICE and the media helps promote awareness of the guideline recommendations and allows any potentially controversial aspects of the guideline recommendations to be explained and set in context.

Press meetings or conferences

At publication, a press release may be issued or a press meeting may be held, or both, but only if heightened media interest in the topic is likely.

Interviews and filming

At or outside a press conference, the media relations team may set up interviews or filming with:

- people involved in developing the guideline recommendations (such as committee members) or
- people with personal experience in the area the guideline recommendations cover (on seeking advice from the people and communities team) or

- representatives from voluntary and community sector organisations and other stakeholders.

Participants are provided with any support and training needed.

Social media communication

Guideline recommendation publication is usually accompanied by activity on social media which may include graphics, animations, videos and quotes from key committee members or NICE directors. In most cases, this work will be prepared ahead of publication.

Rules about press and other communication events

Communication from NICE and committee members to the media

Information may be provided to the media under embargo until publication of the guideline recommendations. Committee members and development teams should not answer any press enquiries they receive before the guideline recommendations are published without involving the media relations team.

Media promotion of the guideline recommendations by external organisations

Committee members and development teams should tell the media relations team if any external organisations are planning their own media promotion for the guideline recommendations (for example by issuing their own press release).

Separate events arranged by committee members

Development teams should tell committee members to notify the media relations team at the earliest possible opportunity if they wish to arrange separate events at which practitioners, providers, commissioners, people using services and the public can learn more about the guideline recommendations.

Material developed from guideline recommendation content

Any materials developed from guideline recommendation content should be submitted to staff with responsibility for quality assurance before submission or presentation to external audiences.

Committee members who wish to publish their materials for a UK audience only may do so under the NICE UK open content licence. This is a self-assessment exercise and no fee is involved.

The international use, adaptation or contextualisation of NICE content is subject to a formal licensing agreement or contract (see NICE's webpage on reusing our content). For advice and support on adapting and contextualising NICE content, see NICE International's webpage on adapting NICE guidelines.

12 Support for putting the guideline recommendations into practice

12.1 Introduction

Our purpose is to help practitioners and commissioners get the best care to people fast, while ensuring value for the taxpayer. The NICE strategy for 2021 to 2026 outlines the importance for our guidance to provide useful and useable advice and for it to purposefully influence the health and care system to adopt the best possible care. To achieve this, we will transform by targeting 3 key areas:

- focusing on what matters most
- creating advice that's useful and useable
- continually learning from data and implementation.

We support the implementation of guidelines by gathering system intelligence, understanding uptake, disseminating guidelines through networks, sharing actionable insight and best practice, and responding to implementation challenges. We work in partnership with key stakeholders to provide a focused number of implementation support resources that have been tested by our users and will have the greatest impact on the system.

Implementation is considered from the guideline monitoring and scoping stages and throughout the guideline development process.

We work with committees, organisations and stakeholders from all relevant sectors to identify potential implementation challenges and consider solutions to address them. For some topics, we work with organisations to develop resources to support implementation, to signpost to existing resources, to embed recommendations in policy documents and to raise awareness with key stakeholders. This chapter outlines some of the resources and support available to help health and care practitioners, commissioners, patients, the public and the voluntary and community sector to use NICE guidelines.

12.2 Tools for planning and resource impact assessment

We provide a baseline assessment tool for each guideline at the time of publication. Organisations can use the tool to identify whether they are in line with NICE guideline recommendations, and to help them plan and record activity to implement them.

We have developed a resource planner to help users plan for and implement our guidance by listing forthcoming guidance, and summarising the resource impact of published guidance and when available those in development.

NICE resource impact assessment tools help organisations assess the potential costs, savings and capacity impacts associated with implementing a guideline. For guidelines that will have a significant resource impact, a resource impact report and resource impact template are produced, where data allows. A guideline's resource impact is significant if the national cost is more than £1 million per year for a single recommendation or £5 million per year for the whole guideline. When costs and savings cannot be quantified but the resource impact may be significant, a resource impact summary report is produced. If the guideline's resource impact is not significant, a 1-page resource impact statement is produced.

12.3 Into practice resources

Visual summaries

For some guidelines, a visual summary of part of the guideline is produced for health or care practitioners. This might happen if practice needs to change, a practitioner needs to make quick decisions, or a specific audience needs support in implementing the recommendations. Examples include a summary showing how the technology appraisals on lung cancer drugs fit together, treatments for more severe depression in adults and antimicrobial prescribing for Clostridioides difficile infection.

Quick guides

We have previously produced quick guides for some social care topics. They aim to help practitioners to implement recommendations (for example, in care homes) or to support people using services to understand what to expect and make decisions about their care.

Patient decision aids

If a guideline contains a highly preference-sensitive decision point, we include information in the guideline to make it easier for professionals and practitioners to discuss options with the person making a decision on care (see the section on supporting shared decision making in the chapter on interpreting the evidence and writing the guideline). Occasionally we develop a separate patient decision aid to support shared decision making by the person and their health or care practitioner. See our web page on making decisions about your care for the process guide and a list of our patient decision aids.

Case studies

NICE case studies support NICE guideline implementation in areas where we have identified a specific implementation or health inequalities challenge. They show how a team or individual has implemented recommendations, addressing these challenges and the learning from their experience. They are published on the tools and resources tab of the guideline and are reviewed in line with any guideline updates.

Into practice guide

Our into practice guide shows how to use evidence to improve care and services. It sets out the most common steps taken when putting evidence-based guidance into practice.

12.4 Assessing and measuring the use of NICE guidance

Several tools are available which aim to address variation in care:

- Our quality standard service improvement template helps providers do an initial assessment of how their service compares with a range of quality statements.
- The innovation scorecard reports on the use of medicines and medical technologies in the NHS in England that we have positively appraised.

We use a range of data to measure the uptake and impact of NICE guidance. Examples include NHS England's Secure Data Environment (SDE), other national data collections, clinical audit, and data from journals. Intelligence and feedback from users of our guidance and strategic partners also provide insights on the use of our guidance.

12.5 Working with organisations to support implementation

We are strengthening external collaboration so that opportunities for implementation are maximised, for example, by aligning with organisational regulation, monitoring and improvement frameworks.

If other organisations produce resources to support guideline implementation we work with them to ensure that the NICE guideline is correctly referenced and embedded within the resource and signpost to it from the NICE website, if possible.

Implementation support is prioritised to reflect the priorities and needs of the health and care system.

Organisations from all sectors and individuals, both lay and practitioner, play a key role in supporting guideline implementation. We work with external organisations in the following ways to help to put all NICE guidance and quality standards into practice:

- The national implementation team work with key national partners to identify barriers to, and priorities for, implementation. The team develop implementation support resources and solutions to help practitioners and commissioners to use NICE products.
- The system implementation team supports regional and local health and care systems to implement NICE guidance and use quality standards. The team are regionally based and work with partners to provide feedback and intelligence to NICE on the views of stakeholders.
- The medicines optimisation team supports the implementation of NICE products and provides feedback and intelligence to NICE from the NHS and other stakeholders, with a specific focus on medicines optimisation. It does this through national, regional and local networks, such as the NICE medicines and prescribing associates.
- An implementation strategy group comprised of external academics and practitioners provides expertise and feedback on our implementation activities and keeps the organisation updated on new and ongoing developments in implementation science.
- The people and communities team works with national and local voluntary and community sector organisations and members of the public to raise awareness of our

guidance and standards, promote their use, and support implementation.

We also actively seek feedback from people who use our guidelines to improve guidelines and implementation resources. We welcome system intelligence and information about any of our guidelines or broader topic areas and use this to inform our monitoring and topic intelligence work. This enables us to prioritise our guideline updates and implementation support work.

13 Ensuring that published guidelines are current and accurate

This chapter describes the process and methods for checking that published guidelines are current and deciding whether updates are needed (surveillance).

13.1 Aims of surveillance

The aim of surveillance is to check that guidelines are up to date. This is done by exploring if there is any new evidence to contradict, reinforce or clarify guideline recommendations. Surveillance also identifies new interventions that may need to be considered within the guideline. Finally, it explores changes in context that may mean modifications are needed, for example, changes in policy, infrastructure, legislation or costs.

NICE has more than 350 published guidelines, so the number of checks needed is considerable. A proactive approach (with an assessment of priority) is used to respond to events that may impact guideline recommendations at any time after guideline publication (for example a safety alert, or publication of a key study)

13.2 Proactive surveillance

Some topic areas change frequently, and this increases the risk of guidelines having out-of-date recommendations. NICE takes a proactive approach to surveillance, and monitors key events (such as ongoing studies) that are judged to be relevant to the guideline.

Events are identified through constant intelligence gathering. This starts during initial guideline development, as the guideline committee and stakeholders can flag up future events that need to be monitored for impact. Ongoing studies are typically identified through discussions with the National Institute for Health Research. This approach means that NICE can quickly identify changes in the evidence base, and assess the impact on recommendations and the need for any changes.

An event that could affect the guideline could include:

- publication of a study that is directly relevant to NICE guidance and has the potential

to affect recommendations

- substantial changes in policy or legislation (an example includes changes to the UK physical activity guidelines by the Chief Medical Office)
- development of a related piece of NICE guidance that contradicts recommendations in another NICE guideline
- withdrawal of a drug from the market, or a clinically significant drug safety update from the Medicines and Healthcare products Regulatory Authority (MHRA) or the Commission on Human Medicines.

This list is not exhaustive and individual events will be considered on a case-by-case basis. To make the most efficient use of resource, events are triaged, to determine whether surveillance assessment is needed

- If an assessment is needed these are prioritised based on:
  - safety (always prioritised first)
  - health and social care system priorities
  - burden on services
  - population impact
  - potential impact on addressing health inequalities
  - evidence base: is it changing frequently and what is the degree of uncertainty?
  - what value NICE could add by incorporating the new information into a guideline.

13.3 Surveillance assessment process

The NICE surveillance programme considers how an event could affect a guideline. This involves checking how the event could affect the guideline recommendations, and taking feedback from topic experts in the area. The check may include intelligence gathering and literature searches, if needed.

Stakeholders are not normally consulted on the decision to update (or not update) a guideline in response to a surveillance check. This is because these checks usually focus

only on one important event and potentially a small section of a guideline. The decisions are published on the NICE website.

If the guideline needs updating, registered stakeholders are informed of the planned approach.

Topic expert engagement

Topic experts (including members of NICE's Expert Advisers Panel) are invited to participate in surveillance. They provide their views about how an event affects the recommendations, and their knowledge of recent developments in the topic area. If the response from topic experts is limited, or further specialist input is needed, we may seek input from other experts. This could include other government organisations, or representatives from a Quality Standards Advisory Committee.

Intelligence gathering

If needed, additional intelligence may be sought. This might include:

- feedback from internal teams within NICE who have expertise in the topic area (for example the NICE medicines or social care teams, if these are major areas in the guideline)
- asking stakeholders for their views, including organisations representing the interests of patients, people using services, carers, and the public
- external queries and comments received since publication of the guideline (these are collated in an issues log for consideration during surveillance)
- related NICE guidance and quality standards (including placeholder statements in NICE quality standards) developed since the guideline was published
- information about guideline implementation, including evidence derived from analysis of primary data on the uptake of recommendations
- information about important ongoing studies in the area covered by the guideline (identified through searches of trial databases)
- changes in licensing status of medicines

- updated or new national policy.

Literature searching

If needed, published evidence is identified by searching a range of bibliographic databases relevant to the topic. Surveillance searches generally use the same databases that were used during the development of the original guideline. The sources searched may vary depending on the topic. In general, MEDLINE, MEDLINE in Process, Embase, Cochrane Database of Systematic Reviews and Cochrane Central Register of Controlled Trials (CENTRAL) would be considered.

Search approaches

The search approach will vary between topics. The following search approaches can be used:

- population or population/intervention search as needed for the guideline scope with:
  - randomised controlled trials (RCTs) and systematic reviews as a default
  - if RCTs are not appropriate because of the topic or guideline (for example, purely diagnostic), then other study types will be considered
- focused searches for a specific question or a new question, meaning that the study type searched for (RCTs or observational studies) should reflect the type expected to address the question
- citation search forward or back (this option would be supplemented with either a restrictive full scope search or focused searches).

Other considerations

It may be appropriate to consider setting limits for the searches, which could include, but are not limited to:

- study design using appropriate search filters
- date
- location

- populations and subpopulations
- intervention
- service delivery aspect
- prognostic factors.

Search period

The search period will start at the:

- end of the search for the last update of the guideline
- end of the search for the last standard check.

The search date ends on the date the search is conducted.

Decision making

Proposals on the need to update a guideline are based on:

- an assessment of the event and any other relevant evidence published since guideline publication (abstracts of primary or secondary evidence)
- information from topic expert engagement
- if relevant, intelligence gathering and feedback from stakeholder consultation.

The decision also includes an element of judgement.

The possible decisions are:

- update (reviewing the evidence and producing either a full update that replaces the original guideline, or a partial update of defined sections of the guidelines)
- no update (this also includes a decision to defer the update until later)
- amend (making changes to the guideline without reviewing the evidence)
- withdraw (either some recommendations or the whole guideline).

Unless the guideline is fully updated or completely withdrawn, there may also be a refresh (for more information on refreshing, see the section on refreshing the guideline recommendations in the chapter on updating guideline recommendations).

When a guideline is being updated, the original scope may be used (unchanged), the original scope may be modified (for example, where new areas have been identified that require an extension to the scope) or a new scope may be developed. See the chapters on the scope and on updating guideline recommendations for further details.

Stakeholder consultation

Consultations on proposed decisions will be held if there is value in doing so (for example, when it is not clear to NICE if an update is needed or not). The consultation period will usually be 2 weeks.

Signing off the final decision

All surveillance proposals go through a validation and approval process at NICE, which includes signoff by the associate director and centre director.

13.4 References and further reading

Alderson LJ, Alderson P, Tan T (2014) Median life span of a cohort of National Institute for Health and Care Excellence clinical guidelines was about 60 months. Journal of Clinical Epidemiology 67: 52–5

Shekelle P, Eccles MP, Grimshaw JM et al. (2001) When should clinical guidelines be updated? British Medical Journal 323: 155–7

Shekelle PG, Ortiz E, Rhodes S et al. (2001) Validity of the Agency for Healthcare Research and Quality clinical practice guidelines: how quickly do guidelines become outdated? JAMA 286: 1461–7

14 Updating guideline recommendations

14.1 Scheduling updates

When scheduling updates of guideline recommendations, NICE prioritises topic areas according to need for both new and updated guidelines.

14.2 Full updates of guidelines

If a full update of a guideline is needed either:

- a new scope is prepared, following the process described in the chapter on the scope **or**
- the scope of the published guideline is used and registered stakeholders are informed.

Sometimes an existing topic-specific committee is asked to update a guideline in their topic area. Sometimes a new topic-specific committee is set up for the update. Recruitment of committee members follows the usual process (see the chapter on decision-making committees). Where possible, the developer informs all members of the topic-specific committee, or topic-expert members of the standing committee, for the published guideline if a new committee is being recruited. The composition of the committee should be tailored to new requirements if a new scope has been developed. The guideline is developed using the same methods and process as for a new guideline and the draft is subject to the normal 4- to 6-week consultation period (see the chapter on the validation process for draft guidelines, and dealing with stakeholder comments). The developer should maintain records appropriate for audit (see the section on committee meetings in the chapter on decision-making committees). The usual process for finalising and publishing the guideline is followed (see the chapter on finalising and publishing the guideline recommendations).

14.3 Updates of topic areas in guidelines

If only some topic areas of a guideline need to be updated, either:

- a new scope is prepared, following the process described in the chapter on the

scope **or**

- parts of the scope of the published guideline are used (as determined by the check of the need for an update; see the chapter on ensuring that published guidelines are current and accurate), and registered stakeholders are informed. No new scope is produced.

In both cases, the scope is clear about exactly which sections of the guideline are being updated and which are not, including any sections that may be withdrawn (for example, if they are now covered in another guideline). Recommendations that are outside the scope of an update may be refreshed (see the section on refreshing the guideline).

The update is developed using the same methods and process as for a new guideline. Updates of some topic areas using the scope of the published guideline use the review questions and review protocols already defined by the existing guideline. However, if the review questions and/or protocols are unavailable, need refinement, or if there is ambiguity in the published guideline, the developer may approach the committee members with topic expertise for advice before starting the evidence review.

Update of topic areas in a guideline are subject to the same level of scrutiny as full updates and new guidelines. The underlying principles of transparency of process and methodological rigour continue to hold. The draft is subject to a consultation period of up to 6 weeks, depending on length and complexity (see the chapter on the validation process for draft guidelines, and dealing with stakeholder comments). The developer should maintain records appropriate for audit (see the section on committee meetings in the chapter on decision-making committees). The usual process for finalising and publishing the guideline is followed (see the chapter on finalising and publishing the guideline recommendations).

14.4 Refreshing the guideline recommendations

Refreshing guideline recommendations allow us to improve the usability of recommendations without changing the intent and therefore without the need for an evidence review. Refreshing changes can be made to guideline recommendations even when the surveillance decision is not to update the guideline. All changes to recommendations made as part of the surveillance process should be agreed by the NICE surveillance programme (see the chapter on ensuring that published guidelines are current and accurate).

When it has been agreed which topic areas need updating, the publishing team also identifies recommendations that may need refreshing to feed into the scoping process. Occasionally during development of the update, additional recommendations that are not part of the update may be identified for refreshing by the committee or the publishing team.

Refreshing might involve:

- amending or adding cross references to other NICE guidance or hyperlinks to other NICE-endorsed tools or resources
- adding or amending wording to reflect changes to a medicine's marketing authorisation, to reflect changes in service configuration (for example, a change from primary care trusts to clinical commissioning groups) or a change to an organisation's name
- ensuring recommendations take into account the latest government policy or guidelines, for example, on alcohol consumption
- amending recommendations to reflect the current practice context, for example, removing references to tools or resources that no longer exist
- bringing recommendations in line with NICE's current policy on wording without affecting the intent, for example:
  - reflecting the involvement of people in decisions about their care
  - using person-centred language.

Refreshing changes that are made during scoping and guideline development should be agreed with NICE staff with responsibility for quality assurance.

14.5 Presenting updates

A full update replaces an existing guideline and has a new set of recommendations, a new set of rationale and impact sections, new evidence reviews and new sections detailing the committee's discussion of the evidence. When a full update is published the old guideline is withdrawn.

When presenting updates of topic areas within guidelines, the aim is to ensure that there is

a single set of publications that bring together the updated information and relevant information from all previous versions of the guideline. In this way, readers of the updated guideline will be able to easily identify what has changed. The rest of this section covers general principles to be used when part of a guideline has been updated.

Preparing an update of topic areas for consultation

Before consultation on an update of topic areas within a guideline, the developer should check the following:

- All sections have been updated as agreed.
- It is clear which sections have been updated and are open for comment during consultation.
- Recommendations from sections which have not been updated have been checked to determine whether any changes are essential (for example, if a medicine is no longer available).
- Refreshing changes (see the section on refreshing the guideline) to recommendations in sections that have not been updated are kept to a minimum (for example, changing from the passive voice to direct instructions).
- A summary of changes to recommendations is included.
- The status of any guidance incorporated in the previous version of the guideline has been confirmed with NICE. For example, has the other guidance been updated by the guideline update?
- All recommendations (new, updated and unchanged) have been assessed with respect to NICE's equality duties.

Preparing the final version of an update of topic areas for publication

The developer should check the following:

- It is clear which sections have been updated, and whether the recommendations have been updated or amended.

- The summary of changes to recommendations has been revised in line with the final recommendations.

Resources to support implementation are also checked for current relevance.

14.6 Post-publication changes

Measures are in place throughout the development of a guideline to avoid errors in the collection, synthesis, interpretation or presentation of the evidence as far as possible. On rare occasions errors are found after publication of the guideline, or users may ask for clarification.

Corrections or changes to published guideline recommendations are made if an error or lack of clarity:

- puts users of health or care services at risk, or affects their care or provision of services **or**
- damages NICE's reputation **or**
- significantly affects the meaning of a recommendation.

Errors or clarifications that do not warrant immediate changes to the recommendations are logged for consideration when the guideline undergoes surveillance (see the chapter on ensuring that published guidelines are current and accurate).

If an error or clarification meets the criteria for changing a published guideline recommendation, NICE's process for dealing with post-publication changes is followed. An explanation of the decisions and actions taken is sent to the person or organisation that reported the error or requested clarification.

Sometimes recommendations need to be removed because a medicine has been removed from the market or a few recommendations have been updated or replaced by recommendations in another guideline.

The guideline is amended. Resources to support implementation are also amended if necessary. The changes are explained in the guideline. Depending on the nature and significance of the change and the time since publication of the guideline, registered stakeholders may also be notified.

Routine maintenance

Routine maintenance changes may also be made after publication or update of a guideline. These include minor changes such as updating or fixing broken links or updating standard text in line with agreed template changes.

14.7 References and further reading

Clark E, Donovan EF, Schoettker P (2006) From outdated to updated, keeping clinical guidelines valid. International Journal for Quality in Health Care 18: 165–6

Eccles M, Rousseau N, Freemantle N (2002) Updating evidence-based clinical guidelines. Journal of Health Services Research and Policy 7: 98–103

Shojania KG, Sampson M, Ansari MT et al. (2007) Updating systematic reviews. AHRQ Technical Reviews and Summaries, technical review 16. Rockville, MD: Agency for Healthcare Research and Quality

Turner T, Misso M, Harris C et al. (2008) Development of evidence-based clinical practice guidelines (CPGs): comparing approaches. Implementation Science 3: 45–52

15 Appendices

Appendices A to P can be found via tools and resources.

- Appendix A: Service delivery – developing review questions, evidence reviews and synthesis
- Appendix B: Approaches to additional consultation and commissioned primary research
- Appendix C: Key roles and responsibilities of committee members
- Appendix D: Guideline committee Terms of Reference and Standing Orders
- Appendix E: Code of conduct for committee members
- Appendix F: Suggested sources for scoping
- Appendix G: Sources for evidence reviews
- Appendix H: Appraisal checklists, evidence tables, GRADE and economic profiles
- Appendix I: Review protocol template
- Appendix J: Call for evidence and expert witnesses
- Appendix K: Network meta-analysis reporting standards
- Appendix L: Process and methods for guidelines developed in response to health and social care emergencies
- Appendix M: Interim principles for methods and processes for supporting digital living guideline recommendations
- Appendix N: Surveillance decision framework and multi-criteria decision framework for deciding whether to develop or update recommendations and which methods to use
- Appendix O: Surveillance - interim principles for monitoring approaches of guideline recommendations
- Appendix P: Updating guideline recommendations

Glossary

Abstract

Summary of a study, which may be published alone or as an introduction to a full scientific paper.

AGREE (Appraisal of Guidelines for Research and Evaluation)

An international collaboration of researchers and policy makers whose aim is to improve the quality and effectiveness of practice guidelines. The AGREE II instrument, developed by the group, is designed to assess the quality of guidelines.

Allocation

The process by which study participants are allocated to a treatment group.

Applicability

How well an observation or the results of a study or review are likely to hold true in a particular setting.

Association

Statistical relationship between 2 or more events, characteristics or other variables. The relationship may or may not be causal.

Audit trail

Clear record of actions so that the reasons for the actions are apparent to a third party. For example, the reasons for changes to a draft guideline should be clearly recorded.

Baseline

A set of measurements before any intervention starts (after any initial 'run-in' period with no intervention), with which subsequent results are compared.

Bias

Systematic (as opposed to random) deviation of the results of a study from the 'true' results, caused by the way the study is designed or conducted.

Case–control study

An observational study to find out the possible cause(s) of a disease or condition. This is done by comparing a group of patients who have the disease or condition (cases) with a group of people who do not have it (controls) but who are otherwise as similar as possible (in characteristics thought to be unrelated to the causes of the disease or condition). This means the researcher can look for aspects of their lives that differ to see if they may have caused the condition.

For example, a group of people with lung cancer might be compared with a group of people the same age who do not have lung cancer. The researcher could compare how long both groups had been exposed to tobacco smoke. Such studies are retrospective because they look back in time from the outcome to the possible causes of a disease or condition.

Citation searching

Citation searching (also known as 'snowballing') can help to identify additional research. It has 2 dimensions:

- Backward citation searching is reviewing references cited in studies identified for inclusion in the review.
- Forward citation searching involves searching for additional studies that cite articles known to be relevant (such as those identified for inclusion in the review).

Cohort study

An observational study with 2 or more groups (cohorts) of people with similar characteristics. One group has a treatment, is exposed to a risk factor or has a particular symptom and the other group does not. The study follows their progress over time and records what happens.

Committee

The advisory group that considers the evidence and develops the recommendations, taking into account the views of stakeholders. NICE has standing committees (which work on multiple guidelines) and topic-specific committees (which are put together for a single guideline topic or to work on multiple guidelines within a topic area). Members include practitioners and professionals (both specialists and generalists, and/or academics), care providers and commissioners, people using health and care services and/or their family members or carers, or people from communities affected by the guideline.

Committee chair

A member of the committee who leads committee meetings, and ensures that the committee keeps to the scope of the guideline, works collaboratively and adheres to NICE's equality policy and principles on social value judgements. The chair completes the equality impact assessment with the developer at scoping and final guideline stages, approves the draft guideline for consultation, and advises the developer on responses to comments from registered stakeholders.

Comparator

The standard (for example, another intervention or usual care) against which an intervention is compared in a study. The comparator can be no intervention (for example, best supportive care).

Conceptual framework

A theoretical structure of assumptions, principles and rules, which holds together the ideas comprising a broad concept. A conceptual model has been defined as the

abstraction and representation of complex phenomena of interest in some readily expressible form, such that the individual stakeholders' understanding of the parts of the system, and/or the mathematical representation of that system, can be shared, questioned, tested and ultimately agreed.

Confidence interval

The confidence interval is a way of expressing how certain we are about the findings from a study, using statistics. It gives a range of results that is likely to include the 'true' value for the population. A wide confidence interval indicates a lack of certainty about the true effect of the test or treatment – often because a small group of patients has been studied. A narrow confidence interval indicates a more precise estimate (for example, if a large number of patients have been studied).

Consultation table

A table of all the comments received by NICE during consultation on a scope or draft guideline. The committee considers the comments received, and the developer then responds to the comments in the table.

Contractors

Organisations contracted to do some aspects of guideline development for NICE. This might include doing evidence reviews or fieldwork, or the developer role.

Co-opted members

An expert invited to 1 or more meetings to contribute to formulating recommendations in a specific part of the guideline. They take part fully in discussions, but do not have voting rights or count towards quorum. Co-opted members can include people with expertise in user, carer or community experience and views, as well as those with professional or practitioner expertise.

Core members (standing committee)

The core members of a standing committee include at least 1 practitioner and 1 lay

member, and may include an economist. A standing committee usually has between 6 and 12 core members. They serve for an initial period of up to 3 years and work on all guidelines developed by the committee during that period.

Correlates review

Correlates reviews describe relationships between epidemiological factors and outcomes.

Cost–benefit analysis

This is a type of economic evaluation in which the costs and benefits are measured using the same monetary units (for example, pounds sterling) to see whether the benefits exceed the costs.

Cost–consequences analysis

This is a type of economic evaluation in which the costs (such as treatment and hospital care) and the consequences (such as health outcomes) of a test or treatment are compared with those for a suitable alternative. Unlike cost–benefit analysis or cost-effectiveness analysis, it does not attempt to summarise outcomes in a single measure (such as the quality-adjusted life year) or in financial terms. Instead, outcomes are shown in their natural units (some of which may be monetary) and it is left to decision-makers to determine whether, overall, the treatment is worth carrying out.

Cost-effectiveness analysis

This is a type of economic evaluation in which the benefits are expressed in non-monetary terms related to health, such as symptom-free days, heart attacks avoided, deaths avoided or life years gained (that is, the number of years by which the intervention extends life). Cost-effectiveness analysis assesses the cost of achieving the same benefit by different means. Cost-effectiveness analysis is also used as an umbrella term to cover all types of economic evaluation.

Cost-minimisation analysis

In a cost-minimisation analysis, the costs of different interventions that provide the same

benefits are compared. If they are equally effective, only the costs are compared because the cheapest intervention will provide the best value for money. In practice, there are relatively few cost-minimisation analyses because it is rare for 2 healthcare interventions to provide exactly the same benefits.

Cost–utility analysis

This is a type of economic evaluation in which the benefits are assessed in terms of both quality and duration of life, and expressed as quality-adjusted life years (QALYs).

Cross-sectional survey

An observational study in which a population is examined to see what proportion has a particular outcome or has been exposed to a specific risk factor, or both. Cross-sectional surveys are usually used to determine the prevalence of outcomes or exposures to risk factors in populations. This type of survey may also be called a cross-sectional study or a prevalence study. Although cross-sectional surveys often provide useful estimates of disease burden for a particular population, they are less reliable for determining the prevalence of very rare conditions or conditions of short duration. Because cross-sectional surveys are descriptive rather than analytical, they cannot be used to estimate the relationship between cause and effect.

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). See also Markov modelling.

Delphi technique

A technique used for reaching agreement on a particular issue, without the participants meeting or interacting directly. It involves sending participants a series of questionnaires asking their views. After completing each questionnaire, participants are asked to give further views in the light of the group feedback until the group reaches a predetermined level of agreement. The judgements of the participants may be analysed statistically.

Design-oriented conceptual model

This is an explicit simplification and abstraction of the problem-oriented conceptual model, mediated by what is feasible and by the availability of evidence and data.

Developer

The team responsible for scoping the guideline, identifying and reviewing the evidence, undertaking economic analyses, supporting the committee and writing the guideline in light of the committee's discussions and decisions. The team includes administrators, coordinators and project managers who provide administrative and management support to the committee, plan and schedule the work, arrange meetings, and liaise with stakeholders, and all other people and organisations contributing to guideline development.

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.

Dosage

The amount of a medicine to be taken, including the size and timing of the doses.

Economic evaluation

The comparative analysis of alternative courses of action in terms of both their costs and consequences.

Economist

A person with skills in economic analysis whose role is to advise on economic aspects of the key issues or questions, review economic literature, prioritise topics for further analysis and carry out additional cost-effectiveness analyses.

Effect (as in treatment effect, effect size)

The observed association between interventions and outcomes, or a statistic to summarise the strength of the observed association.

Effectiveness

The extent to which an intervention produces an overall benefit under usual or everyday conditions. In this manual effectiveness includes cost effectiveness unless otherwise indicated.

Endorsement

The NICE endorsement programme formally endorses resources produced by external organisations that support the implementation of NICE guidance and the use of quality standards in part or in full.

Epidemiological review

Epidemiological reviews describe a problem in terms of its causes, distribution, control and prevention, and can be used to help focus the review questions. For example, an epidemiological review of accidents would provide information on the most common accidents, morbidity and mortality statistics, and data on inequalities in the impact of accidents.

Equity

Fair distribution of resources or benefits.

Evidence

Information on which a decision or recommendation is based. Evidence can be obtained from a wide range of sources, including randomised controlled trials, observational studies and expert opinion (of practitioners, people using services, family members and carers).

Evidence from practice

Information on context and current practice, which could be in the form of published implementation trials, audit data, manuscripts of case studies or service evaluation reports, or accounts from experts.

Evidence review

Identifying and reviewing the evidence, and undertaking economic analyses:

- The information specialist identifies relevant literature to answer the review questions, creates databases to manage the search results and keeps a log of search results and strategies.
- The technical analyst critically appraises the evidence, distils it into tables and writes brief summaries (including GRADE tables, GRADE-CERQual or evidence statements, if used). The reviewer also summarises the main issues for the committee and contributes to its discussions.
- The economist identifies potential economic issues to be considered in the guideline and performs economic analyses.

Exceptional update

Update of a guideline carried out sooner than originally planned because new data have become available.

Exclusion criteria (literature review)

Explicit criteria used to decide which studies should be excluded from consideration as potential sources of evidence.

Exclusion criteria (study participants)

Criteria that define who is not eligible to participate in a study.

Expert Advisers Panel

The Expert Advisers Panel provides a single repository of experts and practitioners (who have been through a robust recruitment process) for the NICE Centre for Guidelines, and can be called upon to contribute to various guideline development activities, such surveillance reviews and guideline updates.

Expert witness

An expert invited to attend a committee meeting to provide evidence from their experience and specific expertise. Expert witnesses answer questions from committee members and may be invited to present evidence in the form of expert testimony, which is published on the NICE website when the guideline is published. Expert witnesses are not members of the committee. They have expert knowledge of 1 or more of the following areas: experience and views of practitioners; people using services; carers or the community and voluntary sector; government and policy; or research and practice.

External validity

The degree to which the results of a study hold true in non-study situations (for example, in routine NHS practice). It may also be referred to as the generalisability of study results to non-study populations. For example, the external validity of a study that took place in Spain may be questioned if the results are applied to people in Australia.

Extrapolation

In data analysis, predicting the value of a parameter outside the range of observed values.

Follow-up

Observation over a period of time of a person, group or defined population to observe changes in health status or health- and social care-related variables.

Forest plot

A type of graph used to display the results of a meta-analysis.

Formal consensus methods

Formal consensus methods are techniques that can be used to enable a committee to reach an agreement on a particular issue. Methods include Delphi and nominal-group techniques, and consensus development conferences. These methods may be used during guideline development when there is a lack of strong research evidence in a particular area.

Free-text terms

Terms used for searching that are not controlled vocabulary as used in the database or information source, but standard terms used in natural language.

Full update of a guideline

When a guideline is identified for a full update, the existing guideline with its recommendations, are stood down and a replacement guideline is developed with new recommendations.

Generalisability

The extent to which the results of a study based on measurements in a particular population or a specific context hold true for another population or in a different context.

GRADE (Grading of Recommendations Assessment, Development and Evaluation)

A systematic and explicit approach to grading the quality of evidence and the strength of recommendations. GRADE is an evolving system and is continuously being adapted and extended to cover different areas and types of evidence; for example, CERQUAL for qualitative evidence and GRADE for diagnostic studies. See the GRADE working group for the latest news and publications.

GRADE table/GRADE profile

A table summarising, for each important outcome, the quality of the evidence and the outcome data (used as part of the GRADE approach to assessing the quality of the evidence).

Grey literature

Literature that is not formally published or that has a limited distribution, such as institutional reports. Grey literature may not be easily identified through standard bibliographic retrieval systems.

Health inequalities

The Public Health England report on place-based approaches for reducing health inequalities defines health inequalities as unfair and avoidable differences in health across the population, and between different groups within society. They categorise health inequalities across 4 dimensions:

- socio-economic status and deprivation (for example, unemployment, poor housing, poor education, low income or people living in deprived areas)
- protected characteristics (for example, age, sex, race, sexual orientation, and disability)
- vulnerable groups of society, or 'inclusion health' groups (for example, vulnerable migrants, homeless people, sex workers, and Gypsy, Roma and Travellers)
- geography (for example, urban or rural areas).

Health-related quality of life

A combination of a person's overall physical, mental and social wellbeing; not merely the absence of disease.

Health Technology Assessment

Independent research 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).

Implementation

The process of putting guideline recommendations into practice.

In confidence material

Information (for example, the findings of a research project) defined as 'confidential' because its public disclosure could affect 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 cost-effectiveness ratio (ICER)

The difference in the mean costs between 2 interventions, strategies or programmes in the population of interest divided by the differences in the mean outcomes between the 2 interventions, strategies or programmes in the population of interest.

Index test

The test in a study which is being compared with the best available test (the reference standard).

Indication (specific)

The defined use of a medicine as licensed by the Medicines and Healthcare products Regulatory Agency (MHRA).

Indirect treatment comparison

An analysis to compare interventions that have not been compared directly in a head-to-head trial.

Internal validity

A measure of how well a research study has been designed and how well it avoids bias. That is, the extent to which the cause-and-effect relationships in a study are true for the people and conditions of the study.

Key issues

Key issues are included in the scope of a guideline and broadly define aspects of care or service provision for which most advice is needed.

Key questions

Key questions are included in the scope of a guideline and are broad questions related to the areas defined by the key issues. Key questions relate to the effectiveness and cost effectiveness of interventions that are being considered for a given population. Key questions are then used to develop more detailed review questions.

Lay member

A member of the committee who has personal experience of using health or care services, or who is from a community affected by the guideline. A lay member can also be someone with experience as a carer, an advocate, or a member or officer of a voluntary or community organisation.

Literature review

A summary of the evidence from several studies, with conclusions about the findings. It may or may not be systematically researched and developed.

Logic model

A model that incorporates the assumed relationships between action and outcomes as described in the conceptual framework.

Marketing authorisation

This was previously known as a product licence. Marketing authorisation is granted to medicines that meet the standards of safety, quality and efficacy set by a medicines regulator (for example, the Medicines and Healthcare products Regulatory Agency [MHRA] or the European Medicines Agency). It is normally necessary before a medicine can be prescribed or sold.

Markov modelling

A decision-analytic technique that predicts future events occurring in a group over a period of time by assigning group members to a fixed number of health states and then modelling transitions among the 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 and Social Care that is 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.

Meta-analysis

A method often used in systematic reviews to combine results from several studies of the same test, treatment or other intervention to estimate the overall effect of the treatment.

Meta-ethnography

A process for sorting and combining the findings from qualitative studies.

Model inputs

Information needed for economic modelling. This may include information about effectiveness, adverse events, diagnostic accuracy, prognosis, quality of life, resource use and costs.

Narrative summary

Summary of findings presented as a written description rather than, for example, as a graph or table.

Net benefit estimates

In cost-effectiveness and cost–utility analysis, the net benefit estimate can be expressed in outcomes (for example, using quality-adjusted life years [QALYs]) or monetary terms. The net health (or outcome) benefit is the difference between the total expected QALYs (or outcome) and the health (or outcomes) expected to be forgone elsewhere (the total expected costs divided by the maximum acceptable incremental cost-effectiveness ratio [ICER] value). The net monetary benefit is the difference between the monetary value of total expected QALYs (our outcome) multiplied by the maximum acceptable ICER value [ICER] and total expected costs. In cost–benefit analysis, the net benefit estimate is the estimate of the amount of money remaining after all payments made are subtracted from all payments received. This is used in the economic evidence profile for guidelines.

Network meta-analysis

An analysis that compares 3 or more interventions using a combination of direct evidence (from studies that directly compare the interventions of interest) and indirect evidence (from studies that do not compare the interventions of interest directly).

NICE guidance

Recommendations produced by NICE. There are 5 types of guidance:

- guidelines covering clinical topics, medicines practice, public health and social care
- interventional procedures guidance
- technology appraisals guidance
- medical technologies guidance
- diagnostics guidance.

All guidance is developed by independent committees and is consulted on.

NICE guidelines

Recommendations (and the evidence they are based on) on broad topics covering health, public health and social care in England. NICE guidelines include clinical, medicines practice, public health and social care guidelines.

Non-randomised controlled trial

These are trials in which participants (or groups) are allocated to receive either the intervention or a control (or comparison intervention) but the allocation is not randomised. This type of study is often called a controlled before-and-after (CBA) study.

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

An odds ratio compares the odds of something happening in one group with the odds of it happening in another. An odds ratio of 1 shows that the odds of the event happening (for example, a person developing a disease or a treatment working) is the same for both groups. An odds ratio of greater than 1 means that the event is more likely in the first group than the second. An odds ratio of less than 1 means that the event is less likely in the first group than in the second group.

Sometimes probability can be compared across more than 2 groups – in this case, one of the groups is chosen as the 'reference category', and the odds ratio is calculated for each group compared with the reference category.

P value

The p value is a statistical measure that is used to indicate whether or not an effect is statistically significant.

People using services and the public

Anyone who is using health or care services, or a member of the public affected by a guideline.

Personal social services

Care services for vulnerable people, including those with special needs because of old age or physical disability and children in need of care and protection. Examples are residential care homes for older people, home help and home care services, and social workers who provide help and support for a wide range of people (Department of Health and Social Care definition).

PICO (population, intervention, comparison and outcome) framework

A structured approach for developing review questions about interventions. The PICO framework divides each question into 4 components: the population (the population being studied), the interventions (what is being done), the comparators (other main treatment options) and the outcomes (measures of how effective the interventions are).

Placeholder statements

In NICE quality standards, placeholder statements are used for areas of care in need of quality improvement but for which there is no evidence-based guidance available to formulate quality statements or measures.

Practitioner

A healthcare, social care or public health worker.

Problem-oriented conceptual model

This is a simplified, diagrammatic representation of the framework that describes the resources, processes and interactions in the delivery of interventions.

Prognosis

A probable course or outcome of a disease. Prognostic factors are characteristics of a patient or disease that influence the disease course. A good prognosis is associated with a low rate of undesirable outcomes; a poor prognosis is associated with a high rate of undesirable outcomes.

Project manager

The staff member who oversees and facilitates the guideline development process.

Proprietary name

The brand name a manufacturer gives to a medicine or device it produces.

QUADAS-2 (Quality Assessment of Diagnostic Accuracy Studies-2)

A tool for assessing the quality of studies of the accuracy of diagnostic tests.

Qualitative research

Qualitative research explores people's beliefs, experiences, attitudes, behaviour and interactions. It asks questions about how and why, rather than how much. It generates non-numerical data, such as a person's description of their pain rather than a measure of pain. Qualitative research techniques include focus groups and in-depth interviews.

Quality-adjusted life year (QALY)

A measure of the state of health of a person or group in which the benefits, in terms of length of life, are adjusted to reflect the quality of life. One QALY is equal to 1 year of life in perfect health.

Quality assurance

NICE staff carry out quality assurance of the guideline, including reviews of the evidence and any economic analysis, to ensure that it is up-to-date, credible, robust and relevant.

- The centre director is responsible for ensuring that the guideline is produced in accordance with this manual. The centre director is also responsible for appointing the committee chair and committee members.
- The topic lead is responsible for the development and quality of the guideline (including the scope), and has delegated responsibility for approving the consultation draft, the final guideline, and other documents, before final approval by NICE's guidance executive. The topic lead also advises the chair of the committee and the developer on matters of process.

- The NICE clinical, medicines, nursing, public health and social care advisers are responsible for providing advice during all stages of guideline development. Additional specialist adviser roles may be appointed by NICE as required to support guideline development.
- The technical adviser is responsible for the technical quality assurance of the evidence reviews and other work undertaken by the technical analyst. The technical adviser commissions, coordinates and quality assures any fieldwork.
- The economics lead is responsible for ensuring the technical quality of the economic evidence and any economic analysis.

Quality assurance of guideline surveillance reflects quality assurance of guideline development. The NICE associate director – surveillance is responsible for ensuring that processes are followed and that decisions to update or not update guidelines are robust and fit for approval by NICE's guidance executive.

Quality of life

See Health-related quality of life.

Quality standards

Quality standards set out the priority areas for quality improvement in health and social care. They cover areas where there is variation in care. Each standard includes a set of statements to help improve quality, and information on how to measure progress.

Quorum

The smallest number of group members that must be present for a valid meeting. The quorum of a committee is 50% of the total potential membership. No recommendations should be confirmed unless the quorum is reached.

Randomised controlled trial

Trials in which participants (or clusters) are randomly allocated to receive either intervention or control. If well implemented, randomisation should ensure that intervention

and control groups differ only in their exposure to treatment.

Recommendations

Specific advice in NICE guidelines on the care and services that are suitable for most people with a specific condition or need, or for particular groups or people in particular circumstances (for example, when being discharged from hospital). Recommendations may also cover ways to promote good health or prevent ill health, or how organisations and partnerships can improve the quality of care and services.

Reference case

The reference case specifies the methods considered by NICE to be the most appropriate for estimating clinical and cost effectiveness when developing guidance. These are also consistent with an NHS objective of maximising health gain from limited resources.

Reference standard (or gold standard)

A method, procedure or measurement that is widely accepted as being the best available to test for or treat a disease.

Research recommendations

Recommendations for future research that cover areas of uncertainty or gaps in the evidence identified during guideline development.

Respondent

Tobacco organisations (for example, tobacco companies, those who speak for them or are funded by them) with an interest in a particular topic. The term 'respondent' acknowledges NICE's commitment to Article 5.3 of the WHO Framework Convention on Tobacco Control. This sets out an obligation to protect the development of public health policy from any vested interests of the tobacco industry.

Review protocol

A document that outlines the background, objectives and planned methods for an evidence review.

Review questions

Review questions guide a systematic review of the literature. They address only the key issues and questions covered in the scope of the guideline, and will usually be structured with a framework (for example, using PICO or SPICE).

Scoping search

A search of key sources at the scoping stage to identify previous guidelines, health technology assessment reports, key systematic reviews, randomised controlled trials and economic evaluations relevant to the guideline topic. The search also includes the NICE website, government, charity, and other community and voluntary sector websites to identify relevant policies and documents.

Scoping workshop

The scoping workshop is attended by registered stakeholders and is held when key issues that need discussion have been identified by the developer. The workshop may be held before during or after consultation.

Search filter

A collection of search terms designed to retrieve certain types of study (for example, those using a specific study design or on a specific topic).

Sensitivity (of a test)

This refers to how well a test detects what it is testing for. It is the proportion of people with the disease or condition that are correctly identified by the study test.

Sensitivity analysis

A means of exploring uncertainty in the results of economic evaluations. There may be uncertainty because data are missing, estimates are imprecise or there is controversy about methodology. Sensitivity analysis can also be used to see how applicable results are to other settings. The analysis is repeated using different assumptions to examine the effect of these assumptions on the results.

- Deterministic sensitivity analysis investigates how bias in selecting data sources for key model parameters might affect the results.
- One-way sensitivity analysis (univariate analysis) varies each parameter individually to investigate how this affects the results.
- Probabilistic sensitivity analysis assigns probability distributions to uncertain parameters and incorporates these into models using decision-analytic techniques (for example, Monte Carlo simulation).

Shared learning examples

These show how NICE guidance and standards have been put into practice by a range of health, local government and social care organisations.

Social care

Social care generally refers to all forms of personal care and other practical assistance for children, young people and adults who need extra support. This includes:

- children and young people who are at risk of, or who are already experiencing, social and emotional problems
- children, young people and adults with learning or physical disabilities or mental health problems
- people who misuse drugs or alcohol
- older people.

Social value judgements

The decisions in NICE guidance are based on the best available evidence. Sometimes the available evidence is not of good quality or can be incomplete, so the committees involved have to make scientific value judgements and social value judgements. Social value judgements take account of society's expectations, preferences, culture and ethical principles when making recommendations.

Specificity (of a test)

This refers to how well a test detects what it is testing for. The proportion of people classified as negative by the reference standard who are correctly identified by the study test.

SPICE framework

A structured approach for developing review questions that divides each question into 5 components: setting, perspective, intervention, comparison and evaluation (SPICE).

Stakeholders

Stakeholders are organisations with an interest in a particular guideline topic; they may represent people whose practice or care is directly affected by the guideline.

They include: national organisations for people who use health and social care services, their families and carers, and the public; local Healthwatch organisations; national organisations that represent health and social care practitioners and other people whose practice may be affected by the guideline, or who can influence uptake of the guideline recommendations; public sector providers and commissioners of care or services; private, voluntary sector and other independent providers of care or services; companies that manufacture drugs, devices, equipment or adaptations, and commercial industries relevant to public health; organisations that fund or carry out research; government departments and national statutory agencies.

Stakeholders are encouraged get involved at all stages. Registered stakeholders comment on the draft scope and draft guideline, may provide evidence, and support implementation of the guideline.

See also respondent.

Standing committee

A committee consisting of core members who work on multiple guidelines. Topic expert members are brought in to work on specific guidelines.

Survey

See cross-sectional study.

Surveillance report

A report that summarises the evidence and intelligence identified through the surveillance process and explains the reasons for updating or not.

Surveillance review

The process of checking whether a guideline needs to be updated. This generally includes consideration of new evidence and intelligence such as topic expert feedback, changes to legislation or policy, and information on implementation.

Systematic review

A review that summarises the evidence on a clearly formulated review question according to a predefined protocol, using systematic and explicit methods to identify, select and appraise relevant studies, and to extract, analyse, collate and report their findings. It may or may not use statistical meta-analysis.

Time horizon

The time period over which the main differences between interventions in effects and the use of resources in health and social care are expected to be experienced, taking into account the limitations of the supporting evidence.

Topic adviser (topic-specific committee)

A member of the committee who also works closely with the developer to provide topic-specific support.

Topic expert members (of a standing committee)

Experts on the topic of a guideline who join a standing committee to work on that guideline. They may include lay members, practitioners, providers and commissioners.

Topic-specific committee

A committee consisting of members appointed for the development of a specific guideline or to work on multiple guidelines within a topic area.

Treatment options

The choices of intervention available.

Update information

Major changes since publication

May 2024: We updated section 1.4 on choice of guideline topics to reflect the publication of NICE-wide topic prioritisation: the manual.

January 2024: Major changes from the 2023 guidelines manual are as follows:

Chapter 2: The scope

We included details of scopes for topic suites, added additional details on NICE's updated approach to considering health inequalities, information on consultation time of scopes was changed to 2 to 4 weeks (except for small updates), simplified the process for scopes for updates and streamlined the chapter including the removal of operational detail.

Chapter 5: Identifying the evidence: literature searching and evidence submission

We added information outlining living evidence surveillance, search sources were updated, more information was added on real world data and NICE's updated approach to considering equalities and health inequalities was added.

Chapter 9: Interpreting the evidence and writing the guideline

We added a table outlining 'topic areas we do not usually make recommendations on', the chapter was restructured to provide a more logical flow and NICE's updated approach to considering equalities and health inequalities was added.

Chapter 11: Finalising and publishing the guideline recommendations

We added NICE's updated approach to considering equalities and health inequalities, clarified the information around embargoed release of guideline recommendations and sections on promoting awareness of guidelines were updated to match current practice.

Chapter 12: Support for putting the guideline recommendations into practice

We have added emphasis on partnership working and supporting implementation, updated the section on tools and resources to reflect current practice, added additional detail on assessing, and measuring the use of guidance, and using that system intelligence in the guideline lifecycle and removed details on endorsement.

Appendix L: Process and methods for guidelines developed in response to health and social care emergencies

We removed duplicate information which repeated content in the main guideline manual and added additional text on how recommendations on medicines are dealt with in partnership with CHTE.

December 2023: We added a new appendix P on updating guideline recommendations, which includes information on types of update, identification of topics, full updates and routine editorial maintenance.

November 2023: We added a new appendix O on interim principles for monitoring approaches of guideline recommendations through surveillance. We also added information on how we will consolidate and streamline our guideline content to appendix M.

August 2023: Major changes from the 2022 guidelines manual are shown below.

Chapter 1: Introduction

We have described the future approach to developing guidelines, including producing digital living guideline recommendations for some topics. We have also made changes to reflect updated processes and methods covered in chapters 4, 6 and 7, and added information on our approach to reducing inequalities.

Chapter 4: Developing review questions and planning the evidence review

We have added a new section on considering health inequalities when preparing review questions, and signposted our real-world evidence framework.

We have added details on writing review protocols to support mixed-methods reviews, qualitative methods and other newer evidence review methods, and on considering medicines safety information and antimicrobial stewardship in protocols.

Chapter 6: Reviewing evidence

We have added details of our approach to equalities and health inequalities when developing recommendations, and signposted our real-world evidence framework. We have also included advice on reusing existing systematic reviews or recommendations from non-NICE producers in evidence reviews to reduce the time and cost of developing recommendations.

We have added details on mixed-methods reviews, qualitative methods and other newer methods, and on considering medicines safety information and antimicrobial stewardship.

Chapter 7: Incorporating economic evaluation

We have made changes to align with the Centre for Health Technology Evaluation methods and processes. The exception is the severity modifier. We are exploring approaches to applying the severity modifier in NICE guidelines.

We have signposted our real-world evidence framework, and added information about a tool to explore providing quantitative estimates of the impact of guideline recommendations on health inequalities. We encourage guideline developers to pilot this tool.

Appendices: We have made minor changes to appendices B, H, I, J and K, to make them accessible, and add examples of additional consultation (appendix B) and expert testimony (appendix J).

January 2022: Major changes from the 2018 guidelines manual are shown below.

Chapter 1: introduction

We have put more emphasis on addressing health inequalities.

Chapter 2: the scope

We have added a new shorter process for scoping, with more emphasis on addressing health inequalities.

Chapter 4: Developing review questions and planning the evidence review

We have added more information on mixed methods reviews and prediction models.

Chapter 5: Identifying the evidence: literature searching and evidence submission

We have added an option for the use of RCT classifiers, and removed OpenGrey as a potential source.

Chapter 6: Reviewing evidence

We have made searching for conference abstracts 'optional' and made it an option for there to be no distinction between 'critical' or 'important' outcomes. We have specified that data extraction for complex analyses (such as network meta-analyses) should be checked by 2 reviewers. We have added instructions for mixed methods reviews. Example evidence statements have been removed (apart from prognostic examples).

Chapter 7: Incorporating economic evaluation

We have removed the end-of-life modifier section as this was a technology appraisal policy applied to NICE guidelines, which the Centre for Health Technology Evaluations (CHTE) have now withdrawn. NICE guidelines will also discontinue its application.

Chapter 8: Linking to other guidance

We have added information on updating NICE technology appraisals in guidelines.

Chapter 9: Interpreting the evidence and writing the guideline

We have added clarification for 'offer', 'consider' and 'do not offer' recommendations. We have given more details on medicines, including dosing information and therapeutic monitoring. We have added options for not making a recommendation when there is insufficient evidence.

Chapter 10: The validation process for draft guidelines, and dealing with stakeholder comments

We have added an option for a 2-week consultation for a small update of 1 or 2 review questions, and added information to differentiate between additional consultation and commissioned primary research.

Chapter 11: Support for putting the guideline recommendations into practice

We have added information about press launches of guidelines.

Chapter 13: Ensuring that published guidelines are current and accurate

We have emphasised proactive surveillance, and removed 5-year routine surveillance.

Appendix B: Approaches to additional consultation and commissioned primary research

We have added information and guidance on additional consultation and commissioned primary research, and given more information on the differences between these 2 processes.

Appendix G: Sources for evidence reviews

We have added sources of primary or real-world data.

Appendix H: Appraisal checklists, evidence tables, GRADE and economic profiles

We have added a checklist for mixed methods reviews.

Appendix J: Call for evidence and expert witnesses

We have added more guidance on conducting a call for evidence and involving expert witnesses.

Appendix L: Process and methods for guidelines developed in response to health and social care emergencies

We have provided a new process and methods for guidelines developed in response to health and social care emergencies.

January 2021: We amended the text on assessing new medicines and significant new indications to clarify that this will be carried out in line with the 2019 voluntary scheme for branded medicines pricing and access published on the GOV.UK website.

October 2020: We amended the text on topic-specific committees to indicate that they

may work on multiple guidelines within a topic area, with membership subject to renewal for a total period of up to 10 years.

July 2020: We added process and methods for guidelines developed in response to health and social care emergencies (appendix L).

October 2018: Major changes from the 2014 guidelines manual are shown below.

Service delivery

Methods for developing recommendations on service delivery have been incorporated into the manual, with information added to the chapters on scoping, searching, evidence submission and economics. We have added a new appendix (appendix A) with detailed advice on developing review questions in this area.

Primary data analytics

NICE is currently exploring the place of primary data analytics in our work and further advice will be shared as this develops.

Chapter 2 – the scope

We encourage developers to list areas where evidence is lacking and details of stakeholders who might provide information in a call for evidence or who might identify expert witnesses.

We are clear that guidelines don't usually include key issues covered by bodies such as the Department of Health and Social Care, NHS England or Public Health England.

We remind developers that guidelines don't usually cover training requirements. However, recommendations may cover the need for specific knowledge and skills for a particular aspect of care.

Chapter 3 – decision-making committees

We encourage developers to think about other related NICE guidance in development and promote cross-representation across committees when topics are closely related.

We include advice for developers about seeking expert testimony from children and vulnerable groups, including use of video recording and giving testimony in private session.

We have made changes to ensure consistency with the updated code of practice for declaring and dealing with conflicts of interest. We have also clarified the involvement of tobacco companies as respondents rather than stakeholders.

Chapter 4 – developing review questions and planning the evidence review

We indicate that core outcome sets (one source is the COMET database) should be used if suitable based on quality and validity. We give standards for assessing the suitability of core outcome sets.

We include review questions that assess diagnostic prediction models and prognostic prediction models and link to external sources of further advice.

We have included a standard template for review protocols as an appendix (appendix I). Registration of the review protocol on the PROSPERO database is now mandatory.

Chapter 5 – identifying the evidence

We have included new sources, tools and approaches to searching, as well as a new prompt for identifying MHRA safety information.

Chapter 6 – reviewing the evidence

We now recommend GRADE as the first approach to quality assessment for all guidelines, including those covering public health and social care topics. We recommend GRADE-CERQual for qualitative evidence.

Results of the analysis and confidence in the evidence should now be presented as GRADE profiles. Evidence statements should be presented when the GRADE approach is not used.

We now have preferred 'checklists' for assessing the quality of the evidence (see appendix H). Use of any other checklist should be agreed in advance with NICE staff with a quality assurance role.

We now indicate that an agreed proportion of papers should be sifted by 2 analysts (not less than 10%) because duplicate sifting of all papers is time consuming and there are other ways of ensuring that relevant papers aren't missed. We have included details of using a machine learning algorithm for priority screening.

We include advice on the minimum outputs and reporting standards for network meta-analyses (see appendix K) and how these apply to developing NICE guidelines.

Chapter 7 – economic evaluation

For base-case analysis, we recommend a cost–utility analysis using a cost per QALY. This will allow more consistent decisions related to costs.

We have clarified that the same levels of evidence and considerations should be used for disinvestment and investment decisions.

We have added information on end of life criteria in line with technology appraisal methods.

Chapter 8 – linking to other guidance

We advise linking to technology appraisal recommendation in the NICE Pathway rather than incorporating TA recommendations verbatim in a guideline.

We have removed the details on updating technology appraisals within a guideline and have added a link to the policy from the Department of Health and Social Care.

We include advice for developers on what to do when similar review questions are covered in other guidelines. Options include linking to the recommendations in the other guideline, using the evidence review to make new recommendations or doing a new systematic review.

Chapter 9 – Interpreting the evidence and writing the guideline

We have simplified advice on writing guidelines, and a separate writing guide with more details and examples will be coming soon.

The section on supporting shared decision-making has been clarified, and includes

information on when a separate decision aid could be produced. The writing guide includes more detail on summarising evidence in the guideline to support a professional's discussion with the person making the decision.

We have added new advice on what to do when recommendations are made for the use of CE-marked devices outside their instructions for use. This includes standard footnote wording.

We have clarified advice on recommendations on the off-label use of medicines.

Chapter 10 – validation

We have defined the types of 'additional consultation' that can inform development. There is more information about the changes in appendix B.

Chapter 12 – implementation

Information on how we work and the tools we produce has been updated. A new section on how we work with other organisations, including endorsing resources, has also been added.

We highlight the role of the new Guideline Recommendations Implementation Panel.

Chapter 13 – surveillance

We have focused the process on event-driven checks of guidelines as well as a standard check every 5 years.

We plan themed surveillance of guidelines covering similar populations or settings to ensure that the process is efficient.

We have revised the process for considering whether to remove a guideline from the static list.

We have indicated that we may refresh some recommendations following an event-driven or standard check.

We have added information about the quality assurance of the surveillance process.

Chapter 14 – updating guidelines

We have added a new section on refreshing recommendations, to make minor changes to improve the usability of recommendations without the need for an evidence review or committee input.

Minor changes since publication

December 2024: We updated names of teams and roles, linked to the interim methods and processes statement for bringing together NICE guidance, and removed reference to the closed NICE accreditation programme.

April 2024: We added a sentence to the section on off-label use of licensed medicines to say that we clearly state when a medicine is being used for off-label use.

September 2023: We added details about the responsibilities of suite faculties to appendix M.

July 2023: We amended appendices M and N to separate, and clarify, surveillance tasks from topic prioritisation tasks.

June 2023: We amended the definition of quality assurance to remove mention of commissioning the developer.

May 2023: We published our NICE interim principles for methods and processes for supporting digital living guideline recommendations (appendix M) and surveillance decision framework and multi-criteria decision framework (appendix N) for deciding whether to develop or update recommendations and which methods to use.

March 2023: We added text in section 4.5 to say that the review questions are published on our website 6 weeks before consultation.

October 2022: We added information to the introduction in relation to the consultation on updating this manual.

September 2022: We added text to section 2.4 to clarify that some small updates will not have a consultation.

November 2020: We aligned the publication of declarations of interest for members of independent advisory panels with process and methods for standard guideline development.

January 2020: A statement was added to the introduction and overview section to reflect the publication of our list of NICE principles. They succeed our social value judgements document which was first published in 2005.

ISBN: 978-1-4731-5692-0