National Institute for Health and Care Excellence

FINAL

Babies, children and young people's experience of healthcare

Supplement 2: Methods

NICE guideline NG204

Development of the guideline and methods

August 2021

Final

These supplementary materials were developed by the National Guideline Alliance which is a part of the Royal College of Obstetricians and Gynaecologists

Disclaimer

The recommendations in this guideline represent the view of NICE, arrived at after careful consideration of the evidence available. When exercising their judgement, professionals are expected to take this guideline fully into account, alongside the individual needs, preferences and values of their patients or service users. The recommendations in this guideline are not mandatory and the guideline does not override the responsibility of healthcare professionals to make decisions appropriate to the circumstances of the individual patient, in consultation with the patient and/or their carer or guardian.

Local commissioners and/or providers have a responsibility to enable the guideline to be applied when individual health professionals and their patients or service users wish to use it. They should do so in the context of local and national priorities for funding and developing services, and in light of their duties to have due regard to the need to eliminate unlawful discrimination, to advance equality of opportunity and to reduce health inequalities. Nothing in this guideline should be interpreted in a way that would be inconsistent with compliance with those duties.

NICE guidelines cover health and care in England. Decisions on how they apply in other UK countries are made by ministers in the Welsh Government, Scottish Government, and Northern Ireland Executive. All NICE guidance is subject to regular review and may be updated or withdrawn.

ISBN: 978-1-4731-4231-2

Contents

Development of the guideline

Remit

The National Institute for Health and Care Excellence (NICE) commissioned the National Guideline Alliance (NGA) to develop a guideline on babies, children and young people's experience of healthcare.

What this guideline covers

Groups that are covered

- Babies, children and young people (aged 17 and under) accessing NHS physical or mental health services, or local authority-commissioned healthcare services.

Key areas that are covered

- Communication with babies, children and young people and their parents or carers.
- Access to healthcare information for children, young people and the parents or carers of babies or young children.
- Advocacy and support for babies, and advocacy and support for and by children and young people.
- Improving babies, children and young people's experience of their healthcare.
- Involving children, young people and the parents or carers of babies in improving experience of healthcare.
- Healthcare environment (including facilities and equipment) appropriate to the age and needs of babies, children and young people, and across all settings in which healthcare is provided.
- Family and peer relationships, including continuing with social activities and schooling.
- Accessibility, continuity and coordination throughout healthcare pathways.

For further details of what the guideline does and does not cover see the guideline scope on the NICE website.

Methods

Introduction

This section summarises methods used to identify and review the evidence, to consider cost effectiveness, and to develop guideline recommendations. This guideline was developed in accordance with methods described in Developing NICE guidelines: the manual.

Declarations of interest were recorded and managed in accordance with NICE's 2018 Policy on declaring and managing interests for NICE advisory committees (NICE 2018).

Developing the review questions and outcomes

The 17 review questions considered in this guideline were based on the key areas identified in the guideline scope .They were drafted by the NGA technical team, and refined and validated by the guideline committee. See Table 1 for a summary of the review questions and index to evidence reports.

The review questions were based on the following frameworks:

- intervention reviews – using population, intervention, comparison and outcome (PICO)
- qualitative reviews – using population, phenomenon of interest and context (PICo)

These frameworks guided the development of review protocols, the literature searching process, and critical appraisal and synthesis of evidence. They also facilitated development of recommendations by the committee.

Literature searches, critical appraisal and evidence reviews were completed for all review questions.

The review questions and evidence reviews corresponding to each question (or group of questions) are summarised in Table 1.

Table 1: Summary of review questions and index to evidence reviews

Evidence review and key area	Subtopic in scope	Review question	Type of review
[A] Communication	Planning healthcare and making shared decisions	How do children and young people, and the parents and carers of babies and young children, prefer to be involved and supported in planning their healthcare and making informed, shared decisions about their health?	Qualitative

Evidence review and key area	Subtopic in scope	Review question	Type of review
[B] Communication	Communication by healthcare staff	How should healthcare staff communicate with babies, children, young people and the parents or carers of babies and young children?	Qualitative
[C] Communication	Consent, privacy and confidentiality	How should issues about consent, privacy and confidentiality be addressed with babies, children and young people?	Qualitative
[D] Information	Providing information	How do children and young people, and the parents or carers of babies and young children prefer to access healthcare information?	Qualitative
[E] Information	Understanding the risks and benefits of healthcare decisions	What are the best ways to help children and young people and the parents and carers of babies and young children understand the risks and benefits of healthcare decisions?	Intervention
[F] Advocacy and support	Involving parents or carers in healthcare and healthcare decisions	How do children and young people want their parents or carers to be involved in their care and decisions about their care?	Qualitative
[G] Advocacy and support	Support from healthcare staff	How do children and young people want healthcare staff to support them?	Qualitative
[H] Advocacy and support	Empowering children and young people to advocate for themselves	How can children and young people be empowered to advocate for themselves?	Qualitative
[I] Advocacy and support	Independent advocacy in healthcare for children and young people	How can the views of babies, children and young people be best represented by independent advocates?	Qualitative

Evidence review and key area	Subtopic in scope	Review question	Type of review
[J] Improving healthcare experience	Improving healthcare experience	What factors are important to babies, children and young people to improve their experience of healthcare services?	Qualitative and quantitative (intervention)
[K] Involvement in improving healthcare experience	Design of healthcare services	How can and how should the perspective of children and young people, and of the parents or carers of babies inform the design of healthcare services?	Qualitative
[L] Involvement in improving healthcare experience	Measuring experience	How can the experience of babies, children and young people be measured so as to improve their experience of healthcare?	Intervention
[M] Healthcare environment	Healthcare environment	What features of environment in which healthcare is provided are important to babies, children and young people to improve their experience of care?	Qualitative
[N] Maintaining usual activities	Supporting participation in usual activities	How can health services support babies, children and young people to participate in usual activities (for example family relationships, schooling, peer friendships, social activities)?	Qualitative
[O] Accessibility, continuity and coordination	Accessing healthcare	What are the facilitators of, and barriers to, accessing healthcare services for babies, children and young people?	Qualitative
[P] Accessibility, continuity and coordination	Continuity of care	What factors promote, or present barriers to, continuity and coordination of care for babies, children and young people?	Qualitative

The COMET database was searched for core outcome sets relevant to this guideline. No core outcome sets were identified and therefore the outcomes were chosen based on committee discussions.

Additional information related to development of the guideline is contained in:

- Supplement 1 Glossary and abbreviations
- Supplement 2 Methods (this document)
- Supplement 3 NGA staff list
- Supplement 4 Reference and focus groups final report
- Supplement 5 Grey literature review
- Supplement 6 Economic study selection

Searching for evidence

Scoping search

During the scoping phase, using the population search terms searches were conducted in Medline, Medline in Process and Embase for previous guidelines, economic evaluations, health technology assessments, systematic reviews, randomised controlled trials, observational studies and qualitative research. Searches of websites of organisations, institutional repositories and internet search engines were also undertaken for relevant policies and related documents, including grey literature.

Systematic literature search

Systematic literature searches were undertaken to identify published evidence relevant to each review question.

Databases were searched using subject headings, free-text terms and, where appropriate, study type filters. Where possible, searches were limited to retrieve studies published in English. All the searches were conducted in the following databases: Medline, Medline-in-Process, Cochrane Central Register of Controlled Trials (CCTR), Cochrane Database of Systematic Reviews (CDSR), Embase and Psycinfo.

Searches were run once for all reviews during development. Searches for all the review questions were updated in July 2020, 18 weeks in advance of the final committee meeting. This time was necessary to allow the large-guideline wide qualitative update search to be screened.

Details of the search strategies, including the study-design filters used and databases searched, are provided in appendix B of each evidence review.

Grey evidence literature search

During scoping, a number of national surveys of children and young people's experience of healthcare were identified. These provided qualitative and quantitative data on large numbers of children's and young people's views and priorities. In addition, there were some surveys of parents' or carer's views on the healthcare experience of their babies or young children. However, although the findings of these surveys were publically available, they had not necessarily been published in peer-reviewed literature and so would unlikely be identified or included in the systematic literature search. However, in order not to miss this potentially rich source of relevant views it was agreed that a focused grey literature search would be carried out. This search would aim to identify:

- National surveys on, or including data on, babies, children and young people's experience of healthcare
- Surveys conducted in or after 2014

Search terms used were:

((child or children or adolescent or adolescents or teenage or youth or paediatric or baby or babies or neonate) and (health*) and (survey or surveys))

Sources searched were:

1) Catalogues and databases:
   OpenGrey
   Greylit
   EThOS
   Patient Experience Library
2) Web searching and other:
   NHS England Scotland Wales and Northern Ireland
   NHS Digital
   NHSsurveys.org
   King's Fund
   Youthhealthtalk.org
   CQC
   nationalvoices.org.uk
   Royal College of Paediatric and Child Health
   Picker Institute
   Joseph Rowntree
   National children's Bureau
   NHS youth forum
   Patient participation groups
3) Charities:
   NSPCC
   Together for short lives
   Rainbow trust
   Save the children
   Children with cancer

Children's society
Barnado's
Honey Pot
Action for Children
Starlight
Child Autism
National Deaf Children's Society
Family Fund
Young minds
Mermaids
The Children's Hospital Charity
Support Evelina
The grand appeal
GOSH
Bliss
Childline
Girl guides
Scouts

4) Google Search

Economic systematic literature search

Systematic literature searches were also undertaken to identify published economic evidence. Databases were searched using subject headings, free-text terms and, where appropriate, an economic evaluations search filter.

A single search, using the population search terms used in the evidence reviews combined with an economic evaluations search filter, was conducted in Medline, Medline in Process, Embase and Psychinfo. Where possible, searches were limited to studies published in English.

The economic literature searches were updated in Aug 2020, 17 weeks in advance of the final committee meeting before consultation on the draft guideline.

Details of the search strategies, including the study-design filter used and databases searched, are provided in Supplement 6 Economic study selection.

Quality assurance

Search strategies were quality assured by cross-checking reference lists of relevant studies, analysing search strategies from published systematic reviews and asking members of the committee to highlight key studies. The principal search strategies for each search were also quality assured by a second information scientist using an adaptation of the PRESS 2015 Guideline Evidence-Based Checklist (McGowan 2016). In addition, all publications highlighted by stakeholders at the time of the consultation on the draft scope were considered for inclusion.

Reviewing evidence

Systematic review process

The evidence was reviewed in accordance with the following approach.

- Potentially relevant articles were identified from the search results for each review question by screening titles and abstracts. Full-text copies of the articles were then obtained.
- Full-text articles were reviewed against pre-specified inclusion and exclusion criteria in the review protocol (see appendix A of each evidence review).
- Key information was extracted from each article on study methods and results, in accordance with factors specified in the review protocol. The information was presented in a summary table in the corresponding evidence review and in a more detailed evidence table (see appendix D of each evidence review).
- Included studies were critically appraised using an appropriate checklist as specified in Developing NICE guidelines: the manual. Further detail on appraisal of the evidence is provided below.
- Summaries of evidence by outcome were presented in the corresponding evidence review and discussed by the committee.

For all review questions, internal (NGA) quality assurance processes included consideration of the outcomes of screening, study selection and data extraction and the committee reviewed the results of study selection and data extraction. The review protocol for each question specifies whether dual screening and study selection was undertaken for that particular question.

Drafts of all evidence reviews were checked by a senior reviewer.

Type of studies and inclusion/exclusion criteria

Inclusion and exclusion of studies was based on criteria specified in the corresponding review protocol.

Systematic reviews were considered to be the highest quality evidence that could be selected for inclusion.

For intervention reviews, randomised controlled trials (RCTs) were prioritised for inclusion because they are considered to be the most robust type of study design that could produce an unbiased estimate of intervention effects. Where there was limited evidence from RCTs, non-randomised controlled trials or observational studies were considered for inclusion.

For qualitative reviews, studies using focus groups, structured interviews or semi-structured interviews were considered for inclusion. Where qualitative evidence was sought, data from surveys or other types of questionnaire were considered for inclusion only if they provided data from open-ended questions, but not if they

reported only quantitative data. Mixed methods studies were considered for inclusion, but only qualitative data was extracted.

Studies that used the views of parents or carers as proxies were included only if they were responding on behalf of their child or charge, and i) the child or charge of the parent or carer was under 5 years-old, or ii) there was a clear rationale provided as to why the study was using parents' or carers' views on and experiences of healthcare as proxies for the child. Studies where part of the population is <18 years-old and part of the population was ≥18 years-old were only included if >66% of the population was in the former group or if the themes which used data for <18 years-old could be ascertained.

UK studies from 2009 onwards were prioritised for decision-making as those conducted in other countries may not be representative of current expectations about either services or current attitudes and behaviours of healthcare professionals. Studies conducted from 2009 were included to ensure that the views and experiences of babies, children and young people reflected the current healthcare system.

Systematic reviews that included evidence from countries other than UK were excluded if the sources of the data and evidence from high-income countries could not be clearly established. Evidence from individual studies conducted in the following high-income countries was included only if no relevant systematic review evidence from high-income countries was identified: Australia, Austria, Belgium, Canada Denmark, Finland, France, Germany, Greece, Iceland, Ireland, Italy, Liechtenstein, Luxembourg, Malta, Monaco, Netherlands, New Zealand, Norway, Portugal, Spain, Sweden, Switzerland, and US.

The committee was consulted about any uncertainty regarding inclusion or exclusion of studies. A list of excluded studies for each review question, including reasons for exclusion is presented in appendix D of the corresponding evidence review.

Narrative reviews, posters, letters, editorials, comment articles, and studies published in languages other than English were excluded. Conference abstracts were not considered for inclusion.

Methods of combining evidence

When planning reviews (through preparation of protocols), the following approaches for data synthesis were discussed and agreed with the committee.

Data synthesis for intervention reviews

Pairwise meta-analysis

Meta-analysis to pool results from RCTs and non-randomised evidence was conducted where possible using Cochrane Review Manager (RevMan5) software.

For continuous outcomes, standardised mean difference (SMD) was used as a summary statistic to pool multiple studies assessing the same outcome, but measured using a variety of scales.

When evidence was based on studies that reported descriptive data or medians with interquartile ranges or p values, this information was included in the corresponding GRADE tables (see below), but it was typically not possible or appropriate to estimate symmetrical 95% CI. Therefore, certain aspects of quality assessment, such as imprecision of the effect estimate could not be assessed as per standard methods for this type of evidence and subjective ratings were considered instead.

Subgroups for stratified analyses were agreed for some review questions as part of protocol development.

When meta-analysis was undertaken, the results were presented visually using forest plots generated using RevMan5 (see appendix E of the relevant evidence review)

When data from surveys were included, descriptive data from the studies were included and no further analysis was performed.

Data synthesis for qualitative reviews

Where possible, a meta-synthesis was conducted to combine evidence from qualitative studies. Whenever studies identified a qualitative theme relevant to the protocol, this was extracted and the main characteristics were summarised. When all themes had been extracted from studies, common concepts were categorised and tabulated. This included information on how many studies had contributed to each theme identified by the NGA technical team.

In qualitative synthesis, a theme being reported more than other themes across included studies does not necessarily mean that the theme is more important than other themes. The aim of qualitative research is to identify new perspectives on a particular topic. Study types and populations in qualitative research can differ widely, meaning that themes identified by just one or a few studies can provide important new information on a given topic.

Themes from individual studies were integrated into a wider context and, when possible, overarching categories of themes with sub-themes were identified. Themes were derived from data presented in individual studies. When themes were extracted from 1 primary study only, theme names used in the guideline mirrored those in the source study. However, when themes were based on evidence from multiple studies, the theme names were assigned by the NGA technical team. The names of overarching categories of themes were also assigned by the NGA technical team.

Emerging themes were placed into a thematic map representing the relationship between themes and overarching categories in cases where these were interrelated. The purpose of such a map is to show relationships between overarching categories and associated themes.

Appraising the quality of evidence

Intervention studies

GRADE methodology for intervention reviews

For intervention reviews, the evidence for outcomes from included RCTs and comparative non-randomised studies was evaluated and presented using the Grading of Recommendations Assessment, Development and Evaluation (GRADE) methodology developed by the international GRADE working group.

GRADE was not used for evidence from surveys or closed-question questionnaires; instead quality of surveys or closed-question questionnaires evidence was assessed using the Center for Evidence-Based Management (CEBMa) critical appraisal tool. More information about this tool can be found on the developer's website.

When GRADE was applied, software developed by the GRADE working group (GRADEpro) was used to assess the quality of each outcome, taking account of individual study quality factors and any meta-analysis results. Results were presented in GRADE profiles (GRADE tables).

The selection of outcomes for each review question was agreed during development of the associated review protocol in discussion with the committee. The evidence for each outcome was examined separately for the quality elements summarised in Table 2. Criteria considered in the rating of these elements are discussed below. Each element was graded using the quality ratings summarised in Table 3. Footnotes to GRADE tables were used to record reasons for grading a particular quality element as having a 'serious' or 'very serious' quality issue. The ratings for each component were combined to obtain an overall assessment of quality for each outcome as described in Table 4.

The initial quality rating was based on the study design: RCTs start as 'high' quality evidence. The rating was then modified according to the assessment of each quality element (Table 2). Each quality element considered to have a 'serious' or 'very serious' quality issue was downgraded by 1 or 2 levels respectively (for example, evidence starting as 'high' quality was downgraded to 'moderate' or 'low' quality).

Table 2: Summary of quality elements in GRADE for intervention reviews

Quality element	Description
Risk of bias ('Study limitations')	This refers to limitations in study design or implementation that reduce the internal validity of the evidence
Inconsistency	This refers to unexplained heterogeneity in the results
Indirectness	This refers to differences in study populations, interventions, comparators or outcomes between the available evidence and inclusion criteria specified in the review protocol

Quality element	Description
Imprecision	This occurs when a study has few participants or few events of interest, resulting in wide confidence intervals that cross minimally important thresholds
Publication bias	This refers to systematic under- or over-estimation of the underlying benefit or harm resulting from selective publication of study results

Table 3: GRADE quality ratings (by quality element)

Quality issues	Description
None or not serious	No serious issues with the evidence for the quality element under consideration
Serious	Issues with the evidence sufficient to downgrade by 1 level for the quality element under consideration
Very serious	Issues with the evidence sufficient to downgrade by 2 levels for the quality element under consideration

Table 4: Overall quality of the evidence in GRADE (by outcome)

Overall quality grading	Description
High	Further research is very unlikely to change the level of confidence in the estimate of effect
Moderate	Further research is likely to have an important impact on the level of confidence in the estimate of effect and may change the estimate
Low	Further research is very likely to have an important impact on the level of confidence in the estimate of effect and is likely to change the estimate
Very low	The estimate of effect is very uncertain

Assessing risk of bias in intervention reviews

Bias is a systematic error, or consistent deviation from the truth in results obtained. When a risk of bias is present the true effect can be either under- or over-estimated.

Risk of bias in RCTs was assessed using the revised Cochrane risk-of-bias tool for randomized trials (RoB 2) (see Appendix H in Developing NICE guidelines: the manual.

The Cochrane risk of bias tool assesses the following possible sources of bias:

- Bias arising from the randomisation process
- Bias due to deviations from intended interventions
- Bias due to missing outcome data
- Bias in measurement of the outcome
- Bias in selection of the reported results

A study with a poor methodological design does not automatically imply high risk of bias; the bias is considered individually for each outcome and it is assessed whether the chosen design and methodology will impact on the estimation of the intervention effect.

Risk of bias in cluster RCTs was assessed using the Cochrane risk of bias tool for cluster randomized trials (see Appendix H in Developing NICE guidelines: the manual.

More details about the Cochrane risk of bias tool and about the Cochrane cluster risk of bias tool can be found in Section 8 of the Cochrane Handbook for Systematic Reviews of Interventions, version 6.1 (Higgins 2011, updated 2019).

For systematic reviews of observational studies the ROBIS checklist was used (see Appendix H in Developing NICE guidelines: the manual.

Assessing inconsistency in intervention reviews

Inconsistency refers to unexplained heterogeneity in results of meta-analysis. When estimates of treatment effect vary widely across studies (that is, there is heterogeneity or variability in results), this suggests true differences in underlying effects. Inconsistency is, thus, only truly applicable when statistical meta-analysis is conducted (that is, results from different studies are pooled). When outcomes were derived from a single study the rating 'no serious inconsistency' was used when assessing this domain, as per GRADE methodology (Santesso 2016).

Inconsistency was assessed visually by inspecting forest plots and observing whether there was considerable heterogeneity in the results of the meta-analysis (for example if the point estimates of the individual studies consistently showed benefits or harms). This was supported by calculating the I-squared statistic for the meta-analysis with an I-squared value of more than 50% indicating considerable heterogeneity, and more than 80% indicating very serious heterogeneity.

When considerable heterogeneity was present, the meta-analysis was re-run using the Der-Simonian and Laird (DerSimonian 1986) method with a random effects model and this was used for the final analysis.

When no plausible explanation for the heterogeneity could be found, the quality of the evidence was downgraded in GRADE for inconsistency.

Assessing indirectness in intervention reviews

Directness refers to the extent to which populations, interventions, comparisons and outcomes reported in the evidence are similar to those defined in the inclusion criteria for the review and was assessed by comparing the PICO elements in the studies to the PICO defined in the review protocol. Indirectness is important when such differences are expected to contribute to a difference in effect size, or may affect the balance of benefits and harms considered for an intervention.

Studies which included the responses of parents or carers as proxies for their child were included only if they were responding on behalf of their child or charge, and the

baby or child of the parent or carer was under-5 years-old, or if there was a clear rationale provided as to why the study is using parents' or carers' views on healthcare as proxies for their child. These studies were not downgraded for indirectness.

Assessing imprecision and importance in intervention reviews

Imprecision in GRADE methodology refers to uncertainty around the effect estimate and whether or not there is an important difference between interventions (that is, whether the evidence clearly supports a particular recommendation or appears to be consistent with several candidate recommendations). Therefore, imprecision differs from other aspects of evidence quality because it is not concerned with whether the point estimate is accurate or correct (has internal or external validity). Instead, it is concerned with uncertainty about what the point estimate actually represents. This uncertainty is reflected in the width of the CI.

The 95% CI is defined as the range of values within which the population value will fall on 95% of repeated samples, were the procedure to be repeated. The larger the study, the smaller the 95% CI will be and the more certain the effect estimate.

Imprecision was assessed in the guideline evidence reviews by considering whether the width of the 95% CI of the effect estimate was relevant to decision making, considering each outcome independently. This is illustrated in Figure 1, which considers a positive outcome for the comparison of treatment 'A' versus treatment 'B'. Three decision-making zones can be differentiated, bounded by the thresholds for minimal importance (minimally important differences; MIDs) for benefit and harm. The MID for harm for a positive outcome means the threshold at which treatment A is less effective than treatment B by an amount that is important to people with the condition of interest (favours B).

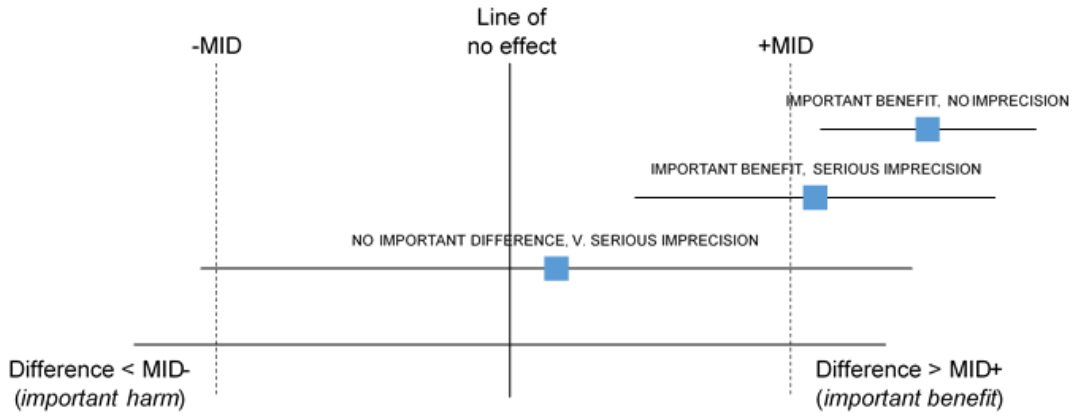
When the CI of the effect estimate is wholly contained in 1 of the 3 zones there is no uncertainty about the size and direction of effect, therefore, the effect estimate is considered precise; that is, there is no imprecision.

When the CI crosses 2 zones, it is uncertain in which zone the true value of the effect estimate lies and therefore there is uncertainty over which decision to make. The CI is consistent with 2 possible decisions, therefore, the effect estimate is considered to be imprecise in the GRADE analysis and the evidence is downgraded by 1 level ('serious imprecision').

When the CI crosses all 3 zones, the effect estimate is considered to be very imprecise because the CI is consistent with 3 possible decisions and there is therefore a considerable lack of confidence in the results. The evidence is therefore downgraded by 2 levels in the GRADE analysis ('very serious imprecision').

Implicitly, assessing whether a CI is in, or partially in, an important zone, requires the guideline committee to estimate an MID or to say whether they would make different decisions for the 2 confidence limits.

Figure 1: Assessment of imprecision and importance in intervention reviews using GRADE



MID: minimally important difference

Defining minimally important differences for intervention reviews

The committee was asked whether there were any recognised or acceptable MIDs in the published literature and community relevant to the review questions under consideration. The committee was not aware of any MIDs that could be used for the guideline.

In the absence of published or accepted MIDs, the committee agreed to use the GRADE default MIDs to assess imprecision. For dichotomous outcomes minimally important thresholds for a RR of 0.8 and 1.25 respectively were used as default MIDs in the guideline. The same thresholds were used as default MIDs in the guideline for all dichotomous outcomes considered in intervention evidence reviews. For continuous outcomes default MIDs are equal to half the median SD of the control groups at baseline (or at follow-up if the SD is not available a baseline).

Assessing publication bias in intervention reviews

Where 10 or more studies were included as part of a single meta-analysis, a funnel plot was produced to graphically assess the potential for publication bias. Where fewer than 10 studies were included for an outcome, the committee subjectively assessed the likelihood of publication bias based on factors such as the proportion of trials funded by industry and the propensity for publication bias in the topic area.

Qualitative reviews

GRADE-CERQual methodology for qualitative reviews

For qualitative reviews an adapted GRADE Confidence in the Evidence from Reviews of Qualitative research (GRADE-CERQual) approach (Lewin 2015) was used. In this approach the quality of evidence is considered according to themes in the evidence. The themes may have been identified in the primary studies or they may have been identified by considering the reports of a number of studies. Quality

elements assessed using GRADE-CERQual are listed and defined in Table 5. Each element was graded using the levels of concern summarised in Table 6. The ratings for each component were combined (as with other types of evidence) to obtain an overall assessment of quality for each theme as described in Table 7.

Table 5: Adaptation of GRADE quality elements for qualitative reviews

Quality element	Description
Risk of bias ('Methodological limitations')	Limitations in study design and implementation may bias interpretation of qualitative themes identified. High risk of bias for the majority of the evidence reduces confidence in review findings. Qualitative studies are not usually randomised and therefore would not be downgraded for study design from the outset (they start as high quality)
Coherence of findings	This refers to the extent to which review findings are well grounded in data from the contributing primary studies and provide a credible explanation for patterns identified in the evidence
Relevance (or applicability) of evidence	This refers to the extent to which the evidence supporting the review findings is applicable to the context specified in the review question
Adequacy of data (theme saturation or sufficiency)	This corresponds to a similar concept in primary qualitative research, that is, whether a theoretical point of theme saturation was achieved, at which point no further citations or observations would provide more insight or suggest a different interpretation of the particular theme. Individual studies that may have contributed to a theme or sub-theme may have been conducted in a manner that by design would have not reached theoretical saturation at an individual study level

Table 6: CERQual levels of concern (by quality element)

Level of concern	Definition
None or very minor concerns	Unlikely to reduce confidence in the review finding
Minor concerns	May reduce confidence in the review finding
Moderate concerns	Will probably reduce confidence in the review finding
Serious concerns	Very likely to reduce confidence in the review finding

Table 7: Overall confidence in the evidence in CERQual (by review finding)

Overall confidence level	Definition
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

Overall confidence level	Definition
Very low	It is unclear whether the review finding is a reasonable representation of the phenomenon of interest

Assessing methodological limitations in qualitative reviews

Methodological limitations in qualitative studies were assessed using the Critical Appraisal Skills Programme (CASP) checklist for qualitative studies (see appendix H in Developing NICE guidelines: the manual; NICE 2014, updated 2018). Overall methodological limitations were derived by assessing the methodological limitations across the 6 domains summarised in Table 8.

Table 8: Methodological limitations in qualitative studies

Domain	Rationale
Aim and appropriateness of qualitative evidence	This domain assesses whether the aims and relevance of the study were described clearly and whether qualitative research methods were appropriate for investigating the research question
Rigour in study design or validity of theoretical approach	This domain assesses whether the study approach was documented clearly and whether it was based on a theoretical framework (such as ethnography or grounded theory). This does not necessarily mean that the framework has to be stated explicitly, but a detailed description ensuring transparency and reproducibility should be provided
Sample selection	This domain assesses the background, the procedure and reasons for the method of selecting participants. The assessment should include consideration of any relationship between the researcher and the participants, and how this might have influenced the findings
Data collection	This domain assesses the documentation of the method of data collection (in-depth interviews, semi-structured interviews, focus groups or observations). It also assesses who conducted any interviews, how long they lasted and where they took place
Data analysis	This domain assesses whether sufficient detail was documented for the analytical process and whether it was in accordance with the theoretical approach. For example, if a thematic analysis was used, the

Domain	Rationale
	assessment would focus on the description of the approach used to generate themes. Consideration of data saturation would also form part of this assessment (it could be reported directly or it might be inferred from the citations documented that more themes could be found)
Results	This domain assesses any reasoning accompanying reporting of results (for example, whether a theoretical proposal or framework is provided)

Assessing relevance of evidence in qualitative reviews

Relevance (applicability) of findings in qualitative research is the equivalent of indirectness for quantitative outcomes, and refers to how closely the aims and context of studies contributing to a theme reflect the objectives outlined in the guideline review protocol.

Studies which included the responses of parents or carers as proxies for their child were included only if they were responding on behalf of their child or charge, and the baby or child of the parent or carer was under-5 years-old, or if there was a clear rationale provided as to why the study is using parents' or carers' views on healthcare as proxies for their child. These studies were not downgraded for relevance.

Assessing coherence of findings in qualitative reviews

For qualitative research, a similar concept to inconsistency is coherence, which refers to the way findings within themes are described and whether they make sense. This concept was used in the quality assessment across studies for individual themes. This does not mean that contradictory evidence was automatically downgraded, but that it was highlighted and presented, and that reasoning was provided. Provided the themes, or components of themes, from individual studies fit into a theoretical framework, they do not necessarily have to reflect the same perspective. It should, however, be possible to explain these by differences in context (for example, the views of children might not be the same as those of young people, but they could contribute to the same overarching themes).

Assessing adequacy of data in qualitative reviews

Adequacy of data (theme saturation or sufficiency) corresponds to a similar concept in primary qualitative research in which consideration is made of whether a theoretical point of theme saturation was achieved, meaning that no further citations or observations would provide more insight or suggest a different interpretation of the theme concerned. It is not equivalent to the number of studies contributing to a theme, but rather to the depth of evidence and whether sufficient quotations or observations were provided to underpin the findings.

Assessing importance in qualitative reviews

For themes stemming from qualitative findings, importance was agreed by the committee taking account of the generalisability of the context from which the theme was derived and whether it was sufficiently convincing to support or warrant a change in current practice, as well as the quality of the evidence.

Reviewing economic evidence

Inclusion and exclusion of economic studies

Titles and abstracts of articles identified through the economic literature searches were independently assessed for inclusion using the predefined eligibility criteria listed in Table 9.

Table 9: Inclusion and exclusion criteria for systematic reviews of economic evaluations

Inclusion criteria
Intervention or comparators in accordance with the guideline scope
Study population in accordance with the guideline scope
Full economic evaluations (cost-utility, cost effectiveness, cost-benefit or cost-consequence analyses) assessing both costs and outcomes associated with interventions of interest. Due to the anticipated lack of full economic evaluations cost analyses were also considered.
Exclusion criteria
Abstracts containing insufficient methodological details
Cost-of-illness type studies

Once the screening of titles and abstracts was completed, full-text copies of potentially relevant articles were requested for detailed assessment. Inclusion and exclusion criteria were applied to articles obtained as full-text copies.

Details of economic evidence study selection are provided in the Supplement 6 (Economic study selection). Lists of excluded studies, economic evidence tables, the results of quality assessment of economic evidence (see below) and health economic evidence profiles are presented in respective evidence chapters.

Appraising the quality of economic evidence

The quality of economic evidence was assessed using the economic evaluations checklist specified in Developing NICE guidelines: the manual (NICE 2014).

Economic modelling

The aims of the economic input to the guideline were to inform the guideline committee of potential economic issues to ensure that recommendations represented a cost effective use of healthcare resources. Economic evaluations aim to integrate

data on healthcare benefits (ideally in terms of quality-adjusted life-years; QALYs) with the costs of different options. In addition, the economic input aimed to identify areas of high resource impact; these are recommendations which (while cost effective) might have a large impact on for commissioners so need special attention.

The guideline committee prioritised the following review questions for economic modelling where it was thought that economic considerations would be particularly important in formulating recommendations.

- What are the best ways to help children and young people and the parents and carers of babies and young children understand the risks and benefits of healthcare decisions?
- What factors are important to babies, children and young people to improve their experience of healthcare services? (Quantitative part of the question)

Original economic modelling was not undertaken for the above review questions due to the lack of suitable effectiveness data to undertake useful new modelling. There was flexibility around economic priorities, however, since all other review questions were qualitative the committee did not identify any other area that would benefit from economic modelling.

When the new economic analysis was not prioritised and where relevant, the committee made a qualitative judgement regarding cost effectiveness by considering expected differences in resource and cost use between options, alongside clinical effectiveness evidence identified from the clinical evidence review.

Cost effectiveness criteria

As specified in Developing NICE guidelines: the manual an intervention was considered to be cost effective if any of the following criteria applied (provided that the estimate was considered plausible):

- the intervention dominated other relevant strategies (that is, it was both less costly in terms of resource use and more effective compared with all the other relevant alternative strategies)
- the intervention cost less than £20,000 per QALY gained compared with the next best strategy
- the intervention provided important benefits at an acceptable additional cost when compared with the next best strategy.

The committee's considerations of cost effectiveness are discussed explicitly under the heading 'The committee's discussion of the evidence' and 'Cost effectiveness and resource use' in the relevant evidence reviews.

Other sources of evidence

Reference and focus groups

In addition to the normal sources of systematic review evidence used in the development of this guideline, a series of reference and focus groups were convened to provide additional input into the guideline. These groups comprised children from age 4 to 14, from a range of geographical and socio-economic backgrounds in attempt to adequately represent different groups of children and young people, and to ensure generalisability across the guideline. Detailed information on the composition of the groups (including ethnicity, geographical location, and socio-economic background), the meeting schedules and the methods of obtaining the children and young people's input is contained in supplements 4 and 4b. The reference and focus groups were conducted independently of the committee and the developer by a third party (the National Children's Bureau, NCB) to minimise bias in the collection of the data.

The committee prioritised the review areas or specific questions on which they asked the groups to focus their discussions, with priority being given to questions where the systematic review of the literature or the grey literature review had failed to provide evidence on the review question, or certain aspects of the question. Asking the same questions using the same workshop tools to multiple groups of children representing diverse populations of children from across England was designed to ensure that the information obtained from these groups was representative of a wider population of children and young people.

Evidence from the reference and focus groups (in the form of quotes or the results of priority setting activities) was mapped to individual review questions and shared with the committee as part of the discussion of the evidence for each review question. The relevant reference and focus group evidence is summarised in each evidence review and the full evidence is provided in Appendix M of each evidence review.

The final NCB report (Supplement 4b) was produced at the end of the development phase of the guideline and was not used by the committee to make recommendations. The evidence from the reference and focus groups used by the committee during development was the tables of quotes mapped across to each question (Supplement 4a).

Grey literature

In addition to the systematic review evidence and the input from the reference groups and focus groups described above, a third source of evidence was utilised in the development of recommendations. National surveys on, or including data on, babies, children and young people's experience of healthcare, and parents' views on the healthcare experience of their babies or young children, published since 2014 were included. The results of these surveys were reviewed and relevant quantitative or qualitative data were extracted in a narrative summary document.

For each identified form of grey literature the evidence (in the form of quantitative results, or quotes from participants) was extracted and mapped to the individual review questions. This information was shared with the committee as part of the discussion of the evidence for each review question. The relevant grey literature evidence is provided in Appendix N of each evidence review. More detail on the included surveys and their quality assessment are contained in supplement 5.

External experts (expert witness)

In addition to the systematic review evidence, the input from the reference groups and focus groups, and grey literature described above, a fourth source of evidence was utilised. This was in the development of recommendations for the review question on independent advocacy, as the systematic review did not identify any published evidence. An expert witness, who was a practising independent advocate, was invited to give testimony to the committee to provide additional evidence from their experience and specific expertise. The committee established that it needed evidence in this particular area from an expert witness because a number of stakeholders (namely NCB, Rethink, Young Minds, Voiceability, public involvement programme [PIP] expert database) were approached to see if they had done work in this particular area or had access to children and young people with experience of using an independent advocate, however none of them did. As a result, it was also established that a call for evidence would have been unsuccessful.

The expert witness was invited to submit a written testimony, addressing the phenomena of interest identified by the committee in the review protocol, and then presented this testimony at a committee meeting via videoconference, and answered questions from the committee. Expert witness are not members of the committee, they do not have voting rights and they are not involved in the final decisions or influence the wording of recommendations.

The written testimony was shared with the committee as part of the discussion of the evidence and is provided in appendix O of evidence review I.

Developing recommendations

Guideline recommendations

Recommendations were drafted on the basis of the committee's interpretation of the available evidence, taking account of the balance of benefits, harms and costs between different courses of action. When effectiveness and economic evidence was of poor quality, conflicting or absent, the committee drafted recommendations based on their expert opinion. The considerations for making consensus-based recommendations include the balance between potential benefits and harms, the economic costs or implications compared with the economic benefits, current practices, recommendations made in other relevant guidelines, person's preferences and equality issues.

The main considerations specific to each recommendation are outlined under the heading 'The committee's discussion of the evidence' within each evidence review.

For further details refer to Developing NICE guidelines: the manual.

Research recommendations

When areas were identified for which evidence was lacking, the committee considered making recommendations for future research. For further details refer to Developing NICE guidelines: the manual.

Validation process

This guideline was subject to a 6-week public consultation and feedback process. All comments received from registered stakeholders were responded to in writing and posted on the NICE website at publication. For further details refer to Developing NICE guidelines: the manual.

Updating the guideline

Following publication, NICE will undertake a surveillance review to determine whether the evidence base has progressed sufficiently to consider altering the guideline recommendations and warrant an update. For further details refer to Developing NICE guidelines: the manual.

Funding

The NGA was commissioned by NICE to develop this guideline.

References

DerSimonian 1986

DerSimonian, R, Laird, N. (1986). Meta-analysis in clinical trials. Controlled clinical trials, 7(3), 177-188

Dixon-Woods 2005

Dixon-Woods M, Agarwal S, Jones D et al. (2005) Synthesising qualitative and quantitative evidence: a review of possible methods. Journal of Health Services Research & Policy 10(1), 45–53

Higgins 2011, updated 2019

Higgins JPT, Green S (editors) (2011) Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 (updated 2019) The Cochrane Collaboration. Available from www.handbook.cochrane.org (accessed 4 March 2021)

Lewin 2015

Lewin S, Glenton C, Munthe-Kaas H et al. (2015) Using qualitative evidence in decision making for health and social interventions: an approach to assess confidence in findings from qualitative evidence syntheses (GRADE-CERQual). PLoS Med 12(10), e1001895

McGowan 2016

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

NICE

National Institute for Health and Care Excellence (NICE) (2014) Developing NICE guidelines: the manual (updated 2018 and 2020). Available from Developing NICE guidelines: the manual (accessed 4 March 2021)

NICE 2018

National Institute for Health and Care Excellence (NICE) (2014) NICE Policy on conflicts of interest (updated 2017). Available from Policy on declaring and managing interests for NICE advisory committees (accessed 4 March 2021)

Santesso 2016

Santesso N, Carrasco-Labra A, Langendam M et al. (2016) Improving GRADE evidence tables part 3: detailed guidance for explanatory footnotes supports creating and understanding GRADE certainty in the evidence judgments. Journal of clinical epidemiology 74, 28-39