

Vaccine uptake in the general population

[L] NICE guideline: methods

NICE guideline <number>

Methods

November 2021

Draft for Consultation

*Evidence reviews were developed by the
Guideline Updates Team*

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 All NICE guidance is subject to regular review and may be updated or withdrawn.

Copyright

© NICE 2021. All rights reserved. Subject to [Notice of rights](#).

ISBN:

Contents

| | |
|---|----------|
| Development of the guideline..... | 5 |
| Remit..... | 5 |
| What this guideline covers..... | 5 |
| Vaccines of interest..... | 5 |
| Groups of interest..... | 5 |
| Settings of interest..... | 5 |
| Activities, services or aspects of care..... | 5 |
| What this guideline does not cover..... | 5 |
| Methods..... | 7 |
| Developing the review questions and outcomes..... | 7 |
| Call for evidence..... | 7 |
| Reviewing research evidence..... | 8 |
| Type of studies and inclusion/exclusion criteria..... | 8 |
| Methods of combining evidence..... | 10 |
| Data synthesis for intervention studies..... | 10 |
| Data synthesis for qualitative studies..... | 11 |
| Data synthesis for mixed methods studies and reviews..... | 12 |
| Appraising the quality of evidence..... | 13 |
| Intervention studies..... | 13 |
| Qualitative studies..... | 15 |
| Mixed methods studies..... | 16 |
| Reviewing economic evidence..... | 17 |
| Inclusion and exclusion of economic studies..... | 17 |
| Appraising the quality of economic evidence..... | 17 |
| Costing exercises..... | 18 |
| References..... | 19 |

1 Development of the guideline

2 Remit

3 The Department of Health and Social Care in England has asked NICE to produce a
4 guideline on vaccine uptake in the general population.

5 This guideline will update and replace the [NICE guideline on immunisations: reducing
6 differences in uptake in under 19s \(PH21\)](#).

7 This guideline will also be used to develop the NICE quality standard for vaccine
8 uptake in the general population.

9 What this guideline covers

10 Vaccines of interest

11 Routine vaccines refers to vaccines available on the UK immunisation schedule as
12 mentioned in [chapter 11: the UK immunisation schedule of the Green Book](#). For this
13 guideline it excludes seasonal flu vaccine.

14 Groups of interest

15 All people who are eligible for vaccines on the routine UK immunisation schedule.
16 Specific consideration will be given to the groups listed in the [equality impact
17 assessment](#).

18 Settings of interest

- 19 • All settings where routine UK immunisation schedule vaccines are offered or
20 delivered.
21 • Occupational health services.
22 • Education settings, including early years settings, schools, pupil referral units and
23 universities.
24 • Private health clinics and vaccination centres where NHS-funded care is
25 delivered.
26 • Secure settings, including prisons and immigration removal centres.

27 Activities, services or aspects of care

28 We will look at evidence in the areas below when developing the guideline, but it may
29 not be possible to make recommendations in all the areas.

- 30 1. Identifying and recording a person's vaccination eligibility and status.
31 2. Increasing the uptake of routine vaccines.

32 What this guideline does not cover

- 33 • Areas covered by [NICE's guideline on tuberculosis](#).
34 • Areas covered by [NICE's guideline on flu vaccination: increasing uptake](#).

- 1 • Travel vaccines.
- 2 • Selective immunisation programmes, as defined in the Green Book.
- 3 • Seasonal vaccinations, for example flu vaccination.
- 4 • Catch-up campaigns alongside the introduction of a new vaccine.
- 5

1 **Methods**

2 This guideline was developed using the methods described in the [2018 NICE](#)
3 [guidelines manual](#).

4 Declarations of interest were recorded according to the NICE conflicts of interest
5 policy.

6 **Developing the review questions and outcomes**

7 The 11 review questions developed for this guideline were based on the key areas
8 identified in the guideline [scope](#). They were drafted by the Guideline Updates team
9 and refined and validated by the guideline committee.

10 The review questions were based on the following frameworks:

- 11 • population, intervention, comparator and outcome (PICO) for reviews of
12 interventions
- 13 • Sample, Phenomenon of Interest, Design, Evaluation, Research type (SPIDER)
14 for qualitative reviews.

15 Full literature searches, critical appraisals and evidence reviews were completed for
16 all review questions.

17 **Call for evidence**

18 A call for evidence was used to try to identify evidence that answered the following
19 question:

- 20 • Is there any learning from the COVID-19 vaccination program that could be used
21 to increase uptake of routine vaccines?

22 We requested the following information:

- 23 • Qualitative evidence focusing the barriers and facilitators to COVID- 19
24 vaccination uptake.
- 25 • Quantitative evidence about the effectiveness of interventions aimed at
26 increasing COVID-19 vaccine uptake.

27 This approach was taken rather than carrying out a systematic review of evidence for
28 the 2 areas above because we expected that there would be limited published
29 evidence available at the time the call for evidence was made due to how recently
30 the COVID-19 vaccination programme had been started.

31 We requested published or unpublished information meeting the above criteria,
32 including any ongoing research. We did not accept promotional material, non-
33 evidence-based assertions of effectiveness or opinion pieces.

34 The results of the call for evidence are presented in document K. This contains the
35 methods used to select the studies of interest and how they were analysed.

1 Reviewing research evidence

2 Type of studies and inclusion/exclusion criteria

3 Searching for evidence

4 Evidence was searched for each review question using the methods specified in the
5 [2018 NICE guidelines manual](#).

6 Selecting studies for inclusion

7 All references identified by the literature searches and from other sources (for
8 example, previous versions of the guideline or studies identified by committee
9 members) were uploaded into EPPI reviewer software (version 5) and de-duplicated.
10 Titles and abstracts were assessed for possible inclusion using the criteria specified
11 in the review protocol. 10% of the abstracts were reviewed by two reviewers, with
12 any disagreements resolved by discussion or, if necessary, a third independent
13 reviewer.

14 The reviews undertaken for this guideline all made use of the priority screening
15 functionality with the EPPI-reviewer systematic reviewing software. This uses a
16 machine learning algorithm (specifically, an SGD classifier) to take information on
17 features (1, 2 and 3 word blocks) in the titles and abstract of papers marked as being
18 'includes' or 'excludes' during the title and abstract screening process, and re-orders
19 the remaining records from most likely to least likely to be an include, based on that
20 algorithm. This re-ordering of the remaining records occurs every time 25 additional
21 records have been screened. Priority screening was used for each review, but no
22 criteria was set for stopping abstract screening. Consequently, the whole abstract
23 database was searched.

24 The full text of potentially eligible studies was retrieved and assessed according to
25 the criteria specified in the review protocol. A standardised form was used to extract
26 data from included studies.

27 Incorporating published evidence syntheses

28 For all review questions where a literature search was undertaken looking for a
29 particular study design, published evidence syntheses (quantitative systematic
30 reviews or qualitative evidence syntheses) containing studies of that design were
31 also included. All included studies from those syntheses were screened to identify
32 any additional relevant primary studies not found as part of the initial search.
33 Evidence syntheses that were used solely as a source of primary studies were not
34 formally included in the evidence review (as they did not provide additional data) and
35 were not quality assessed.

36 If published evidence syntheses were identified sufficiently early in the review
37 process (for example, from the surveillance review or early in the database search),
38 they were considered for use as the primary source of data, rather than extracting
39 information from primary studies. Syntheses considered for inclusion in this way were
40 quality assessed to assess their suitability using the appropriate checklist, as outlined
41 in [Table 1](#). Note that this quality assessment was solely used to assess the quality of
42 the synthesis in order to decide whether it could be used as a source of data, as

1 outlined in [Table 2](#), not the quality of evidence contained within it, which was
 2 assessed in the usual way as outlined in the section on 'Appraising the quality of
 3 evidence'.

4 **Table 1 Checklist for published evidence syntheses**

| Type of synthesis | Checklist for quality appraisal |
|--|--|
| Systematic review of quantitative evidence | ROBIS |
| Qualitative evidence synthesis | <p>ENTREQ reporting standard for published evidence synthesis is the generic reporting standard for QES, however specific reporting standards exist for meta-ethnography (eMERGe) and for realist synthesis (RAMESES II). If these reporting standards are not appropriate to the QES then an adapted PRISMA framework is used (see Flemming K, Booth A, Hannes K, Cargo M, Noyes J. Cochrane Qualitative and Implementation Methods Group guidance series-paper 6: reporting guidelines for qualitative, implementation, and process evaluation evidence syntheses. <i>Journal of Clinical Epidemiology</i> 2018; 97: 79-85).</p> |

5 Each published evidence synthesis was classified into one of the following three
 6 groups:

- 7 • High quality – It is unlikely that additional relevant and important data would be
 8 identified from primary studies compared to that reported in the review, and
 9 unlikely that any relevant and important studies have been missed by the review.
- 10 • Moderate quality – It is possible that additional relevant and important data would
 11 be identified from primary studies compared to that reported in the review, but
 12 unlikely that any relevant and important studies have been missed by the review.
- 13 • Low quality – It is possible that relevant and important studies have been missed
 14 by the review.

15 Each published evidence synthesis was also classified into one of three groups for its
 16 applicability as a source of data, based on how closely the review matches the
 17 specified review protocol in the guideline. Studies were rated as follows:

- 18 • Fully applicable – The identified review fully covers the review protocol in the
 19 guideline.
- 20 • Partially applicable – The identified review fully covers a discrete subsection of
 21 the review protocol in the guideline (for example, some of the factors in the
 22 protocol only).
- 23 • Not applicable – The identified review, despite including studies relevant to the
 24 review question, does not fully cover any discrete subsection of the review
 25 protocol in the guideline.

26 The way that a published evidence synthesis was used in the evidence review
 27 depended on its quality and applicability, as defined in [Table 2](#). When published
 28 evidence syntheses were used as a source of primary data, data from these
 29 evidence syntheses were quality assessed and presented in GRADE/CERQual
 30 tables in the same way as if data had been extracted from primary studies. In
 31 questions where data was extracted from both systematic reviews and primary
 32 studies, these were checked to ensure none of the data had been double counted
 33 through this process.

1 Table 2 Criteria for using published evidence syntheses as a source of data

| Quality | Applicability | Use of published evidence synthesis |
|----------|----------------------|--|
| High | Fully applicable | Data from the published evidence synthesis were used instead of undertaking a new literature search or data analysis. Searches were only done to cover the period of time since the search date of the review. If the review was considered up to date (following discussion with the guideline committee and NICE lead for quality assurance), no additional search was conducted. |
| High | Partially applicable | Data from the published evidence synthesis were used instead of undertaking a new literature search and data analysis for the relevant subsection of the protocol. For this section, searches were only done to cover the period of time since the search date of the review. If the review was considered up to date (following discussion with the guideline committee and NICE lead for quality assurance), no additional search was conducted. For other sections not covered by the evidence synthesis, searches were undertaken as normal. |
| Moderate | Fully applicable | Details of included studies were used instead of undertaking a new literature search. Full-text papers of included studies were still retrieved for the purposes of data analysis. Searches were only done to cover the period of time since the search date of the review. |
| Moderate | Partially applicable | Details of included studies were used instead of undertaking a new literature search for the relevant subsection of the protocol. For this section, searches were only done to cover the period of time since the search date of the review. For other sections not covered by the evidence synthesis, searches were undertaken as normal. |

2 Methods of combining evidence

3 Data synthesis for intervention studies

4 Where possible, meta-analyses were conducted to combine the results of
5 quantitative studies for each outcome. Network meta-analyses was considered in
6 situations where there were at least 3 treatment alternatives. When there were 2
7 treatment alternatives, pairwise meta-analysis was used to compare interventions.
8 No network meta-analyses were carried out as part of the guideline development
9 process.

10 Pairwise meta-analysis

11 Pairwise meta-analyses were performed in Cochrane Review Manager V5.3. A
12 pooled relative risk was calculated for dichotomous outcomes (using the Mantel–
13 Haenszel method) reporting numbers of people having an event. Both relative and
14 absolute risks were presented, with absolute risks calculated by applying the relative
15 risk to the risk in the comparator arm of the meta-analysis (calculated as the total
16 number events in the comparator arms of studies in the meta-analysis divided by the
17 total number of participants in the comparator arms of studies in the meta-analysis).

1 A pooled mean difference was calculated for continuous outcomes (using the inverse
2 variance method) when the same scale was used to measure an outcome across
3 different studies. Where different studies presented continuous data measuring the
4 same outcome but using different numerical scales (e.g. a 0-10 and a 0-100 visual
5 analogue scale), these outcomes were all converted to the same scale before meta-
6 analysis was conducted on the mean differences. Where outcomes measured the
7 same underlying construct but used different instruments/metrics, data were
8 analysed using standardised mean differences (SMDs, Hedges' g).

9 For continuous outcomes analysed as mean differences, change from baseline
10 values were used in the meta-analysis if they were accompanied by a measure of
11 spread (for example standard deviation). Where change from baseline (accompanied
12 by a measure of spread) were not reported, the corresponding values at the
13 timepoint of interest were used. If only a subset of trials reported change from
14 baseline data, final timepoint values were combined with change from baseline
15 values to produce summary estimates of effect. For continuous outcomes analysed
16 as standardised mean differences this was not possible. In this case, if all studies
17 reported final timepoint data, this was used in the analysis. If some studies only
18 reported data as a change from baseline, analysis was done on these data, and for
19 studies where only baseline and final time point values were available, change from
20 baseline standard deviations were estimated, assuming a correlation coefficient
21 derived from studies reporting both baseline and endpoint data, or if no such studies
22 were available, assuming a correlation of 0.5 as a conservative estimate (Follman et
23 al., 1992; Fu et al., 2013). In cases where SMDs were used they were back
24 converted to a single scale to aid interpretation by the committee where possible.

25 Random effects models were fitted when there was significant between-study
26 heterogeneity in methodology, population, intervention or comparator was identified
27 by the reviewer in advance of data analysis. This decision was made and recorded
28 before any data analysis was undertaken.

29 For all other syntheses, fixed- and random-effects models were fitted, with the
30 presented analysis dependent on the degree of heterogeneity in the assembled
31 evidence. Fixed-effects models were the preferred choice to report, but in situations
32 where the assumption of a shared mean for fixed-effects model were clearly not met,
33 even after appropriate pre-specified subgroup analyses were conducted, random-
34 effects results are presented. Fixed-effects models were deemed to be inappropriate
35 if there was significant statistical heterogeneity in the meta-analysis, defined as
36 $I^2 \geq 50\%$.

37 However, in cases where the results from individual pre-specified subgroup analyses
38 were less heterogeneous (with $I^2 < 50\%$) the results from these subgroups were
39 reported using fixed effects models. This may have led to situations where pooled
40 results were reported from random-effects models and subgroup results were
41 reported from fixed-effects models.

42 Data synthesis for qualitative studies

43 Where multiple qualitative studies were identified for a single question, information
44 from the studies was combined using a thematic synthesis. The thematic synthesis
45 was based partly on a priori categories describing phenomena the committee was
46 interested in (for example, themes relating to information and misinformation) and
47 partly on themes that emerged from the coding of the included studies. Papers were

1 uploaded to NVivo 11 software where the relevant data from the papers were coded.
2 Once all of the included studies had been examined and coded, the resulting sets of
3 codes were aggregated into themes and sub-themes. The aggregated themes were
4 used to develop interpretive 'review findings'. These review findings were evaluated
5 using CERQual to examine their relevance to the review question, the importance
6 given to each theme, and the extent to which each theme recurred across the
7 different studies. The findings were reproduced in a summary of qualitative findings
8 table along with example quotes and details of the CERQual assessment of each
9 review finding.

10 **Data synthesis for mixed methods studies and reviews**

11 Where mixed methods studies were included in the reviews, the data was extracted
12 and analysed separately for the quantitative and qualitative components using the
13 relevant methods for each type of data (see above). If the data could not be analysed
14 separately then no data synthesis was attempted, and the results were presented to
15 the committee for discussion as individual studies. Any correlations or discrepancies
16 between the findings of the mixed methods studies and the syntheses of the
17 quantitative and qualitative findings of the above analyses will be noted.

18 **Data synthesis for mixed methods sections of reviews**

19 The quantitative and qualitative results were presented as a concept diagram with
20 quantitative findings mapped onto the qualitative ones.

21 To do this the following approach was taken using the education interventions review
22 as an example:

- 23 • A mixed methods summary diagram was produced which combined the main
24 education-related findings from the qualitative barriers and facilitators review
25 (evidence review B) with the relevant quantitative results from this review.
- 26 • Findings relating to infrastructure, were identified from review B and the ones that
27 were considered to be most important were summarised in the summary of the
28 evidence section within the education review chapter.
- 29 • These findings spanned the age groups and life stages and were further
30 summarised to produce a diagram with key barriers and facilitators to vaccine
31 uptake that related to education.
- 32 • Where possible links were made between barriers and corresponding facilitators
33 that had been raised in the findings themselves or that were logically linked. So,
34 for example, if a barrier concerned literacy problems and there was quantitative
35 evidence from a study using video information about vaccines then the results of
36 this study were summarised and placed in a box linked to the relevant barrier or
37 facilitator.
- 38 • The quantitative evidence was then mapped onto the qualitative evidence.
- 39 • If a study could not be linked to a barrier or facilitator then it was shown in
40 separate box at the side of the diagram.

1 Appraising the quality of evidence

2 Intervention studies

3 Intervention studies (relative effect estimates)

4 RCTs and quasi-randomised controlled trials were quality assessed using the
5 Cochrane Risk of Bias Tool. Non-randomised controlled trials and cohort studies
6 were quality assessed using the ROBINS-I tool. Other study types (for example
7 controlled or uncontrolled before and after studies) were assessed using the
8 preferred option specified in the NICE guidelines manual 2018 (appendix H).
9 Evidence on each outcome for each individual study was classified into one of the
10 following groups:

- 11 • Low risk of bias – The true effect size for the study is likely to be close to the
12 estimated effect size.
- 13 • Moderate risk of bias – There is a possibility the true effect size for the study is
14 substantially different to the estimated effect size.
- 15 • High risk of bias – It is likely the true effect size for the study is substantially
16 different to the estimated effect size.
- 17 • Critical risk of bias (ROBINS-I only) - It is very likely the true effect size for the
18 study is substantially different to the estimated effect size.

19 Each individual study was also classified into one of three groups for directness,
20 based on if there were concerns about the population, intervention, comparator
21 and/or outcomes in the study and how directly these variables could address the
22 specified review question. Studies were rated as follows:

- 23 • Direct – No important deviations from the protocol in population, intervention,
24 comparator and/or outcomes.
- 25 • Partially indirect – Important deviations from the protocol in one of the following
26 areas: population, intervention, comparator and/or outcomes.
- 27 • Indirect – Important deviations from the protocol in at least two of the following
28 areas: population, intervention, comparator and/or outcomes.

29 Minimally important differences (MIDs) and clinical decision thresholds

30 The Core Outcome Measures in Effectiveness Trials (COMET) database was
31 searched to identify published minimal clinically important difference thresholds
32 relevant to this guideline that might aid the committee in identifying clinical decision
33 thresholds for the purpose of GRADE. No relevant MIDs were identified from the
34 database. In addition, the Guideline Committee were asked to prospectively specify
35 any outcomes where they felt a consensus clinical decision threshold could be
36 defined from their experience. They were unable to define MIDs for the reviews in
37 this guideline because they thought the clinically meaningful change in the outcomes
38 (e.g., offers of vaccination, uptake of vaccines) may differ between vaccinations.
39 Therefore, the line of no effect was used to downgrade for imprecision using GRADE.

40 GRADE for intervention studies analysed using pairwise analysis

41 GRADE was used to assess the quality of evidence for the outcomes specified in the
42 review protocol. Data from randomised controlled trials, non-randomised controlled

1 trials and cohort studies (which were quality assessed using the Cochrane risk of
 2 bias tool or ROBINS-I) were initially rated as high quality while data from other study
 3 types were initially rated as low quality. The quality of the evidence for each outcome
 4 was downgraded or not from this initial point, based on the criteria given in [Table 3](#).

5 **Table 3 Rationale for downgrading quality of evidence for intervention studies**

| GRADE criteria | Reasons for downgrading quality |
|------------------|---|
| Risk of bias | <p>Not serious: If less than 33.3% of the weight in a meta-analysis came from studies at moderate or high risk of bias, the overall outcome was not downgraded.</p> <p>Serious: If greater than 33.3% of the weight in a meta-analysis came from studies at moderate or high risk of bias, the outcome was downgraded one level.</p> <p>Very serious: If greater than 33.3% of the weight in a meta-analysis came from studies at high risk of bias, the outcome was downgraded two levels.</p> <p>Extremely serious: If greater than 33.3% of the weight in a meta-analysis came from studies at critical risk of bias, the outcome was downgraded three levels</p> |
| Indirectness | <p>Not serious: If less than 33.3% of the weight in a meta-analysis came from partially indirect or indirect studies, the overall outcome was not downgraded.</p> <p>Serious: If greater than 33.3% of the weight in a meta-analysis came from partially indirect or indirect studies, the outcome was downgraded one level.</p> <p>Very serious: If greater than 33.3% of the weight in a meta-analysis came from indirect studies, the outcome was downgraded two levels.</p> |
| Inconsistency | <p>Concerns about inconsistency of effects across studies, occurring when there is unexplained variability in the treatment effect demonstrated across studies (heterogeneity), after appropriate pre-specified subgroup analyses have been conducted. This was assessed using the I² statistic.</p> <p>N/A: Inconsistency was marked as not applicable if data on the outcome was only available from one study.</p> <p>Not serious: If the I² was less than 33.3%, the outcome was not downgraded.</p> <p>Serious: If the I² was between 33.3% and 66.7%, the outcome was downgraded one level.</p> <p>Very serious: If the I² was greater than 66.7%, the outcome was downgraded two levels.</p> |
| Imprecision | <p>Where the line of no effect was defined as an MID for the outcome, it was downgraded once if the 95% confidence interval for the effect size crossed the line of no effect (i.e. the outcome was not statistically significant), and twice if the sample size of the study was sufficiently small that it is not plausible any realistic effect size could have been detected.</p> |
| Publication bias | <p>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. When a funnel plot showed convincing evidence of publication bias, or the review team became aware of other evidence of publication bias (for example, evidence of unpublished trials where there was evidence that the effect estimate differed in published and unpublished data), the outcome was downgraded once. If no evidence of publication bias was found for any outcomes in a review (as was often the</p> |

| GRADE criteria | Reasons for downgrading quality |
|----------------|---|
| | case), this domain was excluded from GRADE profiles to improve readability. |

1 For outcomes that were originally assigned a quality rating of 'low' (when the data
2 was from observational studies that were not appraised using the ROBINS-I
3 checklist), the quality of evidence for each outcome was upgraded if any of the
4 following three conditions were met and the risk of bias for the outcome was rated as
5 'not serious':

- 6 • Data from studies showed an effect size sufficiently large that it could not be
7 explained by confounding alone.
- 8 • Data showed a dose-response gradient.
- 9 • Data where all plausible residual confounding was likely to increase our
10 confidence in the effect estimate.

11 Qualitative studies

12 Individual qualitative studies were quality assessed using the CASP qualitative
13 checklist. Each individual study was classified into one of the following three groups:

- 14 • Low risk of bias – The findings and themes identified in the study are likely to
15 accurately capture the true picture.
- 16 • Moderate risk of bias – There is a possibility the findings and themes identified in
17 the study are not a complete representation of the true picture.
- 18 • High risk of bias – It is likely the findings and themes identified in the study are
19 not a complete representation of the true picture

20 Each individual study was also classified into one of three groups for relevance,
21 based on if there were concerns about the perspective, population, phenomenon of
22 interest and/or setting in the included studies and how directly these variables could
23 address the specified review question. Studies were rated as follows:

- 24 • Highly relevant – No important deviations from the protocol in perspective,
25 population, phenomenon of interest and/or setting.
- 26 • Relevant – Important deviations from the protocol in one of perspective,
27 population, phenomenon of interest and/or setting.
- 28 • Partially relevant – Important deviations from the protocol in at least two of the
29 perspective, population, phenomenon of interest and/or setting.

30 CERQual was used to assess the confidence we have in each of the review findings.
31 Evidence from all qualitative study designs (interviews, focus groups etc.) was initially
32 rated as high confidence and the confidence in the evidence for each theme was
33 then downgraded from this initial point as detailed in [Table 4](#) below.

34 Table 4 Rationale for downgrading confidence in evidence for qualitative 35 questions

| CERQual criteria | Reasons for downgrading confidence |
|----------------------------|---|
| Methodological limitations | Not serious: If the theme was identified in studies at low risk of bias, the outcome was not downgraded |

| CERQual criteria | Reasons for downgrading confidence |
|------------------|---|
| | Serious: If the theme was identified only in studies at moderate or high risk of bias, the outcome was downgraded one level. Very serious: If the theme was identified only in studies at high risk of bias, the outcome was downgraded two levels. |
| Relevance | High: If the theme was identified in highly relevant studies, the outcome was not downgraded Moderate: If the theme was identified only in in relevant and partially relevant studies, the outcome was downgraded one level. Low: If the theme was identified only in partially relevant studies, the outcome was downgraded two levels. |
| Coherence | Coherence was addressed based on two factors: Between study – does the theme represent the range of viewpoints covered by the included studies. Theoretical – does the theme provide a convincing theoretical explanation for the patterns found in the data. The outcome was downgraded once if there were concerns about one of these elements of coherence, and twice if there were concerns about both elements. |
| Adequacy of data | The outcome was downgraded if there was insufficient data to develop an understanding of the phenomenon of interest, either due to insufficient studies, participants or observations. |

1 Mixed methods studies

2 Where published mixed methods studies were included in this guideline the following
3 processes were followed:

- 4 • the quantitative and qualitative components had to both be relevant for inclusion
5 using the same criteria for these types of evidence as detailed above and in the
6 review protocols, otherwise only the relevant component was used if it could be
7 extracted separately.
- 8 • if the data could be extracted separately, then:
- 9 ○ the quantitative and qualitative components of the study were assessed
10 for risk of bias using the relevant checklists for these study types (as
11 detailed above) where possible.
- 12 ○ the quantitative and qualitative results were extracted separately and
13 analysed with other quantitative and qualitative data (as detailed above).
- 14 ○ the quality of the synthesised evidence was assessed using GRADE or
15 GRADE-CerQual as relevant (as detailed above above).
- 16 • if the quantitative and qualitative components were both relevant for inclusion but
17 the data could not be extracted separately, then the quality of the study was
18 assessed using the [mixed methods appraisal tool](#) (2018 version). The quality of
19 the results was not assessed separately but covered as part of committee
20 discussions about this evidence.

1 Reviewing economic evidence

2 Inclusion and exclusion of economic studies

3 Literature reviews seeking to identify published cost–utility analyses of relevance to
4 the issues under consideration were conducted for all questions. In each case, the
5 search undertaken for the clinical review was modified, retaining population and
6 intervention descriptors, but removing any study-design filter and adding a filter
7 designed to identify relevant health economic analyses. In assessing studies for
8 inclusion, population, intervention and comparator, criteria were always identical to
9 those used in the parallel clinical search. Initially only cost–utility analyses were
10 included, but due to a lack of evidence in the younger age groups the criteria was
11 expanded in those groups to include cost-effectiveness analyses. Economic
12 evidence profiles, including critical appraisal according to the Guidelines manual,
13 were completed for included studies.

14 Appraising the quality of economic evidence

15 Economic studies identified through a systematic search of the literature were
16 appraised using a methodology checklist designed for economic evaluations (NICE
17 guidelines manual; 2014). This checklist is not intended to judge the quality of a
18 study per se, but to determine whether an existing economic evaluation is useful to
19 inform the decision-making of the committee for a specific topic within the guideline.

20 There are 2 parts of the appraisal process. The first step is to assess applicability
21 (that is, the relevance of the study to the specific guideline topic and the NICE
22 reference case); evaluations are categorised according to the criteria in [Table 5](#).

23 Table 5 Applicability criteria

| Level | Explanation |
|----------------------|--|
| Directly applicable | The study meets all applicability criteria, or fails to meet one or more applicability criteria but this is unlikely to change the conclusions about cost effectiveness |
| Partially applicable | The study fails to meet one or more applicability criteria, and this could change the conclusions about cost effectiveness |
| Not applicable | The study fails to meet one or more applicability criteria, and this is likely to change the conclusions about cost effectiveness. These studies are excluded from further consideration |

24 In the second step, only those studies deemed directly or partially applicable are
25 further assessed for limitations (that is, methodological quality); see categorisation
26 criteria in [Table 6](#).

27 Table 6 Methodological criteria

| Level | Explanation |
|---------------------------------|---|
| Minor limitations | Meets all quality criteria, or fails to meet one or more quality criteria but this is unlikely to change the conclusions about cost effectiveness |
| Potentially serious limitations | Fails to meet one or more quality criteria and this could change the conclusions about cost effectiveness |

| Level | Explanation |
|--------------------------|---|
| Very serious limitations | Fails to meet one or more quality criteria and this is highly likely to change the conclusions about cost effectiveness. Such studies should usually be excluded from further consideration |

1 Where relevant, a summary of the main findings from the systematic search, review
 2 and appraisal of economic evidence is presented in an economic evidence profile
 3 alongside the clinical evidence.

4 Costing exercises

5 To support recommendations with a potential resource impact where there was
 6 clinical evidence, simple cost-effectiveness analyses were conducted where the cost
 7 per additional person vaccinated was estimated. Cost-utility analyses were not
 8 conducted because there was not sufficient published evidence to inform the
 9 downstream costs and consequences of both the uptake interventions and current
 10 practice. Specially, the costs of the vaccines themselves are not known (as they are
 11 based on confidentially agreed prices with the DHSC), and because at most
 12 vaccination timepoints, vaccinations against multiple diseases would be given, this
 13 means that to accurately estimate QALY gains all of these potentially preventable
 14 diseases would need to be modelled, which was deemed to be unrealistic.
 15 Additionally, the vaccinations that are routine in the UK have already been deemed
 16 cost-effective, so it was not felt appropriate to be reassessing those consultations as
 17 part of this guideline, as that is the remit of the JCVI, not NICE.

18 Because cost-utility analyses were not conducted, the usual willingness-to-pay
 19 threshold of £20,000 per QALY gained could not be used in these costing analyses.
 20 Additionally, it was noted that this usual NICE threshold is only relevant if the
 21 resources considered are an NHS opportunity cost (i.e. where resources could
 22 theoretically be moved around between different services/parts of the NHS). For the
 23 majority of this guideline, the recommendations do not represent such a situation, but
 24 rather one where vaccine providers are trying to make the most efficient use of a
 25 finite set of resources available to them (and therefore opportunity costs fall more
 26 directly on the vaccine providers themselves, affecting other potential uptake
 27 increasing interventions that could be done).

28 This different situation therefore led to the committee interpreting these cost-
 29 effectiveness results in the following way. First, they agreed it was appropriate to be
 30 comparing different possible ways of increasing vaccine uptake, to identify the most
 31 cost-effective (i.e. something with a lower cost per additional person vaccinated
 32 should be given a higher priority for implementation than something with a higher
 33 cost per person vaccinated). Second, the committee agreed that it was not possible
 34 to draw a fixed threshold above which things would not be cost-effective. Rather, the
 35 guideline should recommend a range of possible interventions to increase uptake,
 36 with providers then implementing as many of those as possible within their resource
 37 constraints (again, starting with the most cost-effective). Finally, the committee
 38 agreed these cost-effectiveness results needed to be considered in combination with
 39 the other evidence in the guideline (such as the qualitative findings) when drawing
 40 conclusions, as there may be reasons (for example, equality issues) where simply
 41 adopting the approach that maximises the number of people vaccinated is not the
 42 correct choice, if for example it leads to substantial variations in uptake, or particular
 43 groups being poorly served. It was also noted that this is a health as well as an

- 1 equality issue, as the existence of particular communities/areas with low uptake rates
- 2 can increase the risk of disease outbreaks, even if overall vaccination rates are high.

3 **References**

- 4 Follmann D, Elliott P, Suh I, Cutler J (1992) Variance imputation for overviews of
- 5 clinical trials with continuous response. *Journal of Clinical Epidemiology* 45:769–73

- 6 Fu R, Vandermeer BW, Shamliyan TA, et al. (2013) Handling Continuous Outcomes
- 7 in Quantitative Synthesis In: *Methods Guide for Effectiveness and Comparative*
- 8 *Effectiveness Reviews* [Internet]. Rockville (MD): Agency for Healthcare Research
- 9 and Quality (US); 2008. Available from
- 10 <http://www.ncbi.nlm.nih.gov/books/NBK154408/>

- 11 Norman G., Sloan JA., Wyrwich KW. (2003) Interpretation of changes in health-
- 12 related quality of life: the remarkable universality of half a standard deviation. *Med*
- 13 *Care* 41(5):582-92.