National Institute for Health and Care Excellence

Final

Reducing sexually transmitted infections (STIs)

NICE guideline: methods

NICE guideline NG221

Methods

June 2022

FINAL

National Institute for Health and Care Excellence

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 All NICE guidance is subject to regular review and may be updated or withdrawn.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: 978-1-4731-4612-9

Contents

Development of the guideline

Remit

To see "What this guideline covers" and "What this guideline does not cover" please see the guideline scope Reducing sexually transmitted infections.

Methods

This guideline was developed in accordance with the process set out in 'Developing NICE guidelines: the manual (2018)'. Where the guidelines manual does not provide advice, additional methods are described below.

Developing the review questions and outcomes

The 11 review questions developed for this guideline were based on the key areas identified in the guideline scope. Review questions were developed by the NICE Public Health Internal Guideline Development (PHIGD) team and refined, validated and signed off by the Public Health Advisory Committee (PHAC) and NICE quality assurance team.

The review questions were based on the PICO[S] framework - Population, Intervention, Comparator and Outcome [and Study type].

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

Details of these elements are found in the review protocols for each review (see Appendix A of each relevant review). Where protocol deviations have been made, these will be reported in the 'Methods' section of the individual review.

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

Evidence review	Review questions	Type of review
A	1.1 What interventions designed to reduce or prevent the acquisition and transmission of STIs, including HIV, are effective and cost effective at preventing STIs in higher risk groups: • Gay, bisexual and other men who have sex with men • Young people aged 16 to 24 years • People from a black African or Caribbean family background • Trans people • Migrant communities • People who are homeless • Asylum seekers	**Quantitative component** of a mixed methods review with review B. Qualitative, mixed methods and committee discussion of the evidence sections are all included in review B
B	1.2. What is the acceptability of interventions for reducing or preventing the acquisition and transmission of STIs in:	**Qualitative component** of a mixed methods review with review A. Including Mixed methods analysis and

Evidence review	Review questions	Type of review
	• Gay, bisexual and other men who have sex with men • Young people aged 16 to 24 years • People from a black African or Caribbean family background • Trans people • Migrant communities • People who are homeless • Asylum seekers	overall committee discussion of the evidence.
C	2.1 What strategies to improve the uptake of STI testing (excluding HIV testing) are effective and cost-effective? 2.2 What factors influence the acceptability of the strategies used to improve the uptake of STI testing?	Mixed methods review
D	2.3 What interventions are effective and cost effective at increasing frequent STI testing in very high risk groups?	Effectiveness review
E	3.1 What partner notification methods for STIs are effective and cost effective? 3.2 What is the acceptability of partner notification methods for STIs?	Mixed methods review
F	1.3a What interventions are effective and cost effective at increasing uptake of hepatitis A and hepatitis B vaccination in MSM? 1.3b What interventions are effective and cost effective at increasing uptake of HPV vaccination in MSM? 1.4a What are the barriers to, and facilitators for, uptake of hepatitis A, or hepatitis B vaccination in MSM? 1.4b What are the barriers to, and facilitators for, uptake of HPV vaccination in MSM?	Mixed methods review
G	1.5 What is the effectiveness, cost effectiveness and unintended consequences of pre-exposure prophylaxis (PrEP) for HIV? 1.6 What is the acceptability of pre-exposure prophylaxis (PrEP) for HIV, and what other factors influence its use?'	

Review protocols

Review protocols were developed with the guideline committee to outline the inclusion and exclusion criteria used to select studies for each evidence review. Where possible, review protocols were prospectively registered in the PROSPERO register of systematic reviews. Protocols are reproduced in each evidence review along with the PROPSERO registration number.

Searching for evidence

Evidence was searched for each review question using the methods specified in the 2018 NICE guidelines manual. Full details of search strategies, databases searched and numbers of studies identified can be found in the appendices of each individual review.

Selecting studies for inclusion

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

None of the evidence reviews made use of the priority screening functionality within the EPPI-reviewer software.

The full text of potentially eligible studies was retrieved and assessed according to the criteria specified in the review protocol. A standardised form was used to extract data from included studies into the EPPI reviewer software and for each intervention a Template for Intervention Description and Replication (TIDieR) checklist was completed.

Incorporating published evidence syntheses

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

If published evidence syntheses were identified sufficiently early in the review process (for example, from the surveillance review or early in the database search), they were considered for use as the primary source of data, rather than extracting information from primary studies. Syntheses considered for inclusion in this way were quality assessed to assess their suitability using the appropriate checklist, as outlined in

Table 2. Note that this quality assessment was solely used to assess the quality of the synthesis in order to decide whether it could be used as a source of data, as

outlined in Table 3, not the quality of evidence contained within it, which was assessed in the usual way as outlined in the section on 'Appraising the quality of evidence'.

Table 2: Checklists for published evidence syntheses

Type of synthesis	Checklist for quality appraisal
Systematic review of quantitative evidence	ROBIS
Network meta-analysis	Modified version of the PRISMA NMA tool (see appendix K of 'Developing NICE guidelines, the manual')
Qualitative evidence synthesis	ENTREQ reporting standard for published evidence synthesis (https://bmcmedresmethodol.biomedcentral.com/articles/10.1186/1471-2288-12-181) is the generic reporting standard for QES, however specific reporting standards exist for meta-ethnography (eMERGe [https://emergeproject.org/]) and for realist synthesis (RAMESES II [https://www.ramesesproject.org/]). 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. Journal of Clinical Epidemiology 2018; 97: 79-85).
Individual patient data meta-analysis	Checklist based on Tierney, Jayne F., et al. "Individual participant data (IPD) meta-analyses of randomised controlled trials: guidance on their use." PLoS Med 12.7 (2015): e1001855.

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

- **High quality** – It is unlikely that additional relevant and important data would be identified from primary studies compared to that reported in the review, and unlikely that any relevant and important studies have been missed by the review.
- **Moderate quality** – It is possible that additional relevant and important data would be identified from primary studies compared to that reported in the review, but unlikely that any relevant and important studies have been missed by the review.
- **Low quality** – It is possible that relevant and important studies have been missed by the review.

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

- **Fully applicable** – The identified review fully covers the review protocol in the guideline.
- **Partially applicable** – The identified review fully covers a discrete subsection of the review protocol in the guideline (for example, some of the factors in the protocol only).
- **Not applicable** – The identified review, despite including studies relevant to the review question, does not fully cover any discrete subsection of the review protocol in the guideline.

The way that a published evidence synthesis was used in the evidence review depended on its quality and applicability, as defined in Table 3. When published evidence syntheses were used as a source of primary data, data from these evidence syntheses were quality assessed and presented in GRADE/CERQual

tables in the same way as if data had been extracted from primary studies. In questions where data was extracted from both systematic reviews and primary studies, these were checked to ensure none of the data had been double counted through this process.

Table 3: 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.

Methods for combining evidence

Data synthesis for intervention studies

Where possible, meta-analyses were conducted to combine the results of quantitative studies for each outcome. Network meta-analyses was considered in situations where there were at least 3 alternatives. When there were 2 alternatives, pairwise meta-analysis was used to compare interventions. No network meta-analyses were undertaken for this guideline.

Pairwise meta-analysis

Pairwise meta-analyses were performed in Cochrane Review Manager V5.3 where possible. A pooled relative risk was calculated for dichotomous outcomes (using the Mantel–Haenszel method) reporting numbers of people having an event. Both relative and absolute risks were presented, with absolute risks calculated by applying the relative risk to the risk in the comparator arm of the meta-analysis (calculated as the total number events in the comparator arms of studies in the meta-analysis divided by the total number of participants in the comparator arms of studies in the meta-analysis). Where thee were zero events in both intervention and control arms,

studies were excluded from meta-analysis. Where there were zero events in one arm no adjustment was made (zero cell adjustment is not required with the Mantel–Haenszel method unless all studies have zero events in one arm in which case 0.5 is added to the arm).

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

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

An a priori decision was taken to use a random effects model for all meta-analyses to reflect the heterogeneity inherent in public health evidence.

Data synthesis for qualitative reviews

Where multiple qualitative studies were identified for a single question, information from the studies was combined using a thematic synthesis. The thematic synthesis was based partly on a priori categories describing phenomena the committee was interested in (for example, using an existing model [framework synthesis]) and partly on themes that emerged from the coding of the included studies. Papers were uploaded to NVivo 11 software where the relevant data from the papers were coded. Once all of the included studies had been examined and coded, the resulting sets of codes were aggregated into themes and sub-themes. The aggregated themes were used to develop interpretive 'review findings' that were evaluated using CERQual. These review findings were reproduced in a summary of qualitative findings table along with example quotes and details of the CERQual assessment of each review finding.

Data synthesis for mixed methods reviews

Data synthesis for mixed methods reviews was carried out in accordance with the Joanna Briggs Institute manual for evidence synthesis (https://wiki.jbi.global/display/MANUAL) chapter 8. Synthesis followed a convergent segregated approach where independent synthesis of quantitative data and

qualitative data was undertaken, followed by the integration of the two types of evidence.

The qualitative and quantitative reviews were presented separately in the reviews and an integration section was written that addressed the following questions:

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

Where appropriate, and data from quantitative and qualitative sections of the review were integrated into tables or logic models/conceptual frameworks to show possible interrelationships between them.

Appraising the quality of evidence

Intervention studies (relative effect estimates)

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

- **Low risk of bias** – The true effect size for the study is likely to be close to the estimated effect size.
- **Moderate risk of bias** – There is a possibility the true effect size for the study is substantially different to the estimated effect size.
- **High risk of bias** – It is likely the true effect size for the study is substantially different to the estimated effect size.
- **Critical risk of bias** (ROBINS-I only) - It is very likely the true effect size for the study is substantially different to the estimated effect size.

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

- **Direct** – No important deviations from the protocol in population, intervention, comparator and/or outcomes.
- **Partially indirect** – Important deviations from the protocol in one of the following areas: population, intervention, comparator and/or outcomes.
- **Indirect** – Important deviations from the protocol in at least two of the following areas: population, intervention, comparator and/or outcomes.

Intervention description and replication

Each intervention included in an effectiveness review was assessed for the clarity of its description within the paper using the TIDieR checklist[1]. This assessment was included alongside the critical appraisal for each paper in the evidence tables.

Minimally important differences (MIDs) and decision thresholds

The Core Outcome Measures in Effectiveness Trials (COMET) database was searched to identify published minimal important difference thresholds relevant to this guideline that might aid the committee in identifying decision thresholds for the purpose of GRADE. Identified MIDs were assessed to ensure they had been developed and validated in a methodologically rigorous way, and were applicable to the populations, interventions and outcomes specified in this guideline. In addition, PHAC members were asked to prospectively specify any outcomes where they felt a consensus decision threshold could be defined from their experience.

Decision thresholds were used to assess imprecision using GRADE and aid interpretation of the size of effects for different outcomes.

No published MIDs were identified for the reviews underpinning this guideline so default MIDs were used where appropriate as described below.

For continuous outcomes expressed as a mean difference where no other decision threshold was available, a decision threshold of 0.5 of the median standard deviations of the comparison group arms was used (Norman et al. 2003). For continuous outcomes expressed as a standardised mean difference where no other decision threshold was available, a decision threshold of 0.5 standard deviations was used.. For relative risks and hazard ratios, the committee assessed the effects of the intervention by noting whether the effect estimate and 95% confidence intervals all lay to one side of the line of no effect. They agreed that when discussing interventions through a population level lens, any definite effect is a meaningful effect since even a very small effect multiplied across a large population will make a meaningful difference. The line of no effect was also used for downgrading outcomes for imprecision as detailed below.

GRADE for pairwise meta-analyses of interventional evidence

GRADE was used to assess the quality of evidence for the outcomes specified in the review protocol. Data from randomised controlled trials, non-randomised controlled trials and cohort studies (which were quality assessed using the Cochrane risk of bias tool or ROBINS-I) were initially rated as high quality while data from other study types were initially rated as low quality. The quality of the evidence for each outcome was downgraded or not from this initial point, based on the criteria given in Table 4.

Table 4: Rationale for downgrading quality of evidence for intervention studies

GRADE criteria	Reasons for downgrading quality
Risk of bias	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.

[1] Hoffmann T C, Glasziou P P, Boutron I, Milne R, Perera R, Moher D et al. Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide *BMJ* 2014; 348 :g1687 doi:10.1136/bmj.g1687

GRADE criteria	Reasons for downgrading quality
	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. 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. 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
Indirectness	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. 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. Very serious: If greater than 33.3% of the weight in a meta-analysis came from indirect studies, the outcome was downgraded two levels.
Inconsistency	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 I^2 statistic. N/A: Inconsistency was marked as not applicable if data on the outcome was only available from one study. Not serious: If the I^2 was less than 33.3%, the outcome was not downgraded. Serious: If the I^2 was between 33.3% and 66.7%, the outcome was downgraded one level. Very serious: If the I^2 was greater than 66.7%, the outcome was downgraded two levels.
Imprecision	The line of no effect was defined as a key indicator of imprecisionfor the outcome. Outcomes were 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 effect estimate crossed both default MIDs (0.8 and 1.25).
Publication bias	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 case), this domain was excluded from GRADE profiles to improve readability.

For outcomes that were originally assigned a quality rating of ‘low’ (when the data was from observational studies that were not appraised using the ROBINS-I checklist), the quality of evidence for each outcome was upgraded if any of the following three conditions were met and the risk of bias for the outcome was rated as ‘no serious’:

- Data from studies showed an effect size sufficiently large that it could not be explained by confounding alone.
- Data showed a dose-response gradient.

- Data where all plausible residual confounding was likely to increase our confidence in the effect estimate.

GRADE-CERQual for qualitative evidence synthesis findings

CERQual was used to assess the confidence we have in each of the review findings. Evidence from all qualitative study designs (interviews, focus groups etc.) was initially rated as high confidence and the confidence in the evidence for each theme was assessed from this initial point as detailed in Table 9 below. Confidence in each criterion was assessed as:

- No or very minor concerns
- Minor concerns
- Moderate concerns
- Serious concerns

And an overall confidence rating of High, Moderate, Low or Very Low was determined based on this.

Table 5 Overall confidence in qualitative outcome

Level	Definition
High confidence	It is highly likely that the review finding is a reasonable representation of the phenomenon of interest
Moderate confidence	It is likely that the review finding is a reasonable representation of the phenomenon of interest
Low confidence	It is possible that the review finding is a reasonable representation of the phenomenon of interest
Very low confidence	It is not clear whether the review finding is a reasonable representation of the phenomenon of interest

Table 9 Rationale for downgrading confidence in evidence for qualitative questions

CERQual criteria	Reasons for downgrading confidence
Methodological limitations	One or more studies contribute data to each review finding in a qualitative evidence synthesis, and these data make up the body of data for a review finding. The methodological limitations of the body of data supporting a review finding are assessed as a whole to identify whether or not any methodological weaknesses within individual studies impact our confidence in a review finding. The methodological limitations for each review finding must be assessed separately since different studies contribute varying amounts of data to each review finding, and methodological quality issues may have varying impacts on different review findings.
Relevance	Relevance is the extent to which the body of data from the primary studies supporting a review finding is applicable to the context specified in the review question. Relevance is the CERQual component that is anchored to the context specified in the review question. How the review question and objectives are expressed, how a priori subgroup analyses are specified and how theoretical considerations inform the

CERQual criteria	Reasons for downgrading confidence
	review design are therefore critical to making an assessment of relevance when applying CERQual.
Coherence	The coherence of a review finding is an assessment of how clear and cogent the fit is between the data from the primary studies and a review finding that synthesises that data. It includes consideration of the general 'fit' of data and whether any discrepancies can be explained.
Adequacy of data	Adequacy of data is an overall determination of the degree of richness as well as the quantity of data supporting a review finding. • Richness of the data is the extent to which the information that the individual study authors have provided is detailed enough to allow the review author to interpret the meaning and context of what is being researched. • Quantity of data relates to the number of studies and participants that this data comes from.

Mixed methods studies

Mixed methods studies were evaluated using the appropriate quality assessment tools for the component study types, see sections on intervention studies and qualitative studies. Other methods of assessing mixed methods studies were agreed with the NICE methods and economics team QA lead and reported in the individual reviews.

Reviewing economic evidence

Inclusion and exclusion of economic studies

Literature reviews seeking to identify published cost–utility analyses of relevance to the issues under consideration were conducted for all questions. In each case, the search undertaken for the public health review was modified, retaining population and intervention descriptors, but removing any study-design filter and adding a filter designed to identify relevant health economic analyses. In assessing studies for inclusion, population, intervention and comparator, criteria were always identical to those used in the parallel public health search; only cost–utility analyses were included. Economic evidence profiles, including critical appraisal according to the Guidelines manual, were completed for included studies.

Appraising the quality of economic evidence

Economic studies identified through a systematic search of the literature were appraised using a methodology checklist designed for economic evaluations (NICE guidelines manual; 2020). This checklist is not intended to judge the quality of a study per se, but to determine whether an existing economic evaluation is useful to inform the decision-making of the committee for a specific topic within the guideline.

There are 2 parts of the appraisal process. The first step is to assess applicability (that is, the relevance of the study to the specific guideline topic and the NICE reference case); evaluations are categorised according to the criteria in Table 10.

Table 10 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

In the second step, only those studies deemed directly or partially applicable are further assessed for limitations (that is, methodological quality); see categorisation criteria in Table 11.

Table 11 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
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

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

Health economic modelling

As well as reviewing the published economic literature for each review question, as described above, original economic analysis was undertaken in selected areas. Priority areas for new health economic analysis were agreed by the committee.

The following general principles were adhered to in developing the analysis:

- Methods were consistent with the NICE reference case.
- The design of the model, selection of inputs and interpretation of the results was discussed and agreed with the committee.
    - Where possible, model inputs were based on the systematic review of the public health literature, supplemented with other published data sources identified by the committee as required.
    - When published data were not available committee expert opinion was used to populate the model.
    - Model inputs and assumptions were reported fully and transparently.
    - The results were subject to sensitivity analysis and limitations were discussed.

Full methods for the original cost-effectiveness analyses are described as in the evidence reviews on pre-exposure prophylaxis for HIV, and interventions to increase uptake of STI testing, as these were the two topics where modelling was undertaken.

Resource impact assessment

The resource impact team used the methods outlined in the in Assessing resource impact process manual: guidelines

The resource impact team worked with the guideline committee from an early stage to identify recommendations that either individually or cumulatively would a substantial impact on resources. The aim was to ensure that a recommendation would not introduce a cost pressure into the health and social care system unless the committee was convinced of the benefits and cost effectiveness of the recommendation. The team gave advice to the committee on issues related to the workforce, capacity and demand, training, facilities and educational implications of the recommendations.