National Institute for Health and Care Excellence

Final

Cardiovascular disease: risk assessment and reduction, including lipid modification

NICE guideline NG238

Methods for evidence review D

December 2023

NICE guideline: methods

Final

Developed by 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 All NICE guidance is subject to regular review and may be updated or withdrawn.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: 978-1-4731-5641-8

Contents

1 Development of the guideline

1.1 Remit

NICE received the remit for this guideline from NHS England.

The remit for this guideline is:

To update the area of escalation of treatment (secondary prevention) in the NICE guideline cardiovascular disease: risk assessment and reduction, including lipid modification (CG181).

1.2 What this guideline covers

The methods outlined in this document relate to lipid modification therapy for the secondary prevention of CVD. This guideline update will include looking at the evidence for and consider making new recommendations or updating existing recommendations on:

- Follow-up of people started on statin treatment.
- Secondary prevention: escalation of lipid modification therapy.

Specifically, this guideline update includes consideration of treatment target(s) for escalation of lipid modification therapy beyond statins, for people with CVD.

1.3 What this guideline does not cover

This guideline update does not include the following sections:

- Identifying and assessing cardiovascular disease (CVD) risk
  - Identifying people for full formal risk assessment
  - Full formal risk assessment
  - Communication about risk assessment and treatment
- Aspirin for primary prevention of CVD
- Lifestyle modifications for the primary and secondary prevention of CVD
  - Cardioprotective diet
  - Physical activity
  - Combined interventions
  - Weight management
  - Alcohol consumption
  - Smoking cessation
  - Plant stanols and sterols
- Lipid modification therapy for the primary and secondary prevention of CVD
  - Initial lipid measurement and referral for specialist review
  - Statins for the prevention of CVD (except follow-up of people started on stain treatment where it relates to secondary prevention)
  - Intolerance of statins
  - Adherence to statin therapy
  - Fibrates for preventing CVD

- Nicotinic acid for preventing CVD
- Bile acid sequestrants (anion exchange resins) for preventing CVD
- Omega 3 fatty acid compounds for preventing CVD
- Combination therapy for preventing CVD.

The evidence for these sections relates to the methods from previous versions of the guideline. For details of the methods used for these reviews please refer to the following sources:

- For 2023 recommendations developed in accordance with the methods outlined in the NICE Guidelines Manual 2020: the methods chapter document.
- For 2014 recommendations developed in accordance with the methods outlined in the NICE Guidelines Manual 2012: chapter 3 of the full guideline document
- For 2008 recommendations developed in accordance with the methods outlined in the NICE Guidelines Technical Manual 2007: sections Q10-Q19 of the guideline appendices document.

2 Methods

This guideline was developed using the methods described in the NICE guidelines manual update 2022.[12]

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

Sections 2.1 to 2.3 describe the process used to identify and review evidence. Sections 2.2 and 2.6 describe the process used to identify and review the health economic evidence.

2.1 Developing the review questions and outcomes

The review question developed for this guideline was based on the key area identified in the guideline scope. It was drafted by the technical team, refined and validated by the guideline committee and signed off by NICE.

The review question was based on the following framework:

- population, intervention, comparator and outcome (PICO) for reviews of interventions.

This framework informed a more detailed protocol that guided the literature searching process, critical appraisal and synthesis of evidence, and facilitated the development of recommendations by the guideline committee. A full literature search, critical appraisal and evidence review was completed for the specified review question:

- In adults with CVD requiring escalation of therapy beyond statins, what is the effectiveness of lipid lowering therapy?

This included the effectiveness of the following interventions: ezetimibe, inclisiran or PCSK9 inhibitors; alirocumab and evolocumab.

2.2 Searching for evidence

2.2.1 Clinical and health economics literature searches

The full strategy including population terms, intervention terms, study types applied, the databases searched, and the years covered can be found in Appendix B of the evidence review.

Systematic literature searches were undertaken to identify published clinical and health economic evidence relevant to the review question. These were run according to the parameters as stipulated within the NICE guideline's manual, https://www.nice.org.uk/process/pmg20/chapter/identifying-the-evidence-literature-searching-and-evidence-submission.

Databases were searched using relevant medical subject headings, free-text terms and where appropriate study-type filters. Studies published in languages other than English were not reviewed, and where possible, searches were restricted to English language. Searches were run on 16 November 2022. Papers published or added to databases after this date were not considered. Where new evidence was identified, for example in consultation comments received from stakeholders, the impact on the

guideline was considered, and the action agreed between the technical team and NICE staff with a quality assurance role.

Searches were quality assured using different approaches prior to being run. Medline search strategies were peer reviewed by a second information specialist using a QA process based on the PRESS checklist.[11] Key (seed) papers if provided, were checked if retrieved by the search.

Searching for unpublished literature was not undertaken. NICE do not have access to drug manufacturers' unpublished clinical trial results, so the clinical evidence considered by the committee for pharmaceutical interventions may be different from that considered by the MHRA and European Medicines Agency for the purposes of licensing and safety regulation.

Additional studies were added to the evidence base these consisted of references included in relevant systematic reviews, and those highlighted by committee members.

2.3 Reviewing evidence

The evidence was reviewed using the following process:

- Potentially relevant studies were identified from the search results by reviewing titles and abstracts. The full papers were then obtained.
- Full papers were evaluated against the pre-specified inclusion and exclusion criteria set out in the protocol to identify studies that addressed the review question.
- Relevant studies were critically appraised using the preferred study design checklist as specified in the NICE guidelines manual.[12] and reported in the review protocol.
- Key information was extracted about interventional study methods and results into EPPI reviewer version 5. Summary evidence tables were produced from data entered into EPPI Reviewer, including critical appraisal ratings.
- Summaries of the evidence were generated by outcome. Outcome data from the randomised trials were meta-analysed where appropriate and reported in GRADE evidence profiles.
- A minimum of 10% of the abstracts were reviewed by two reviewers, with any disagreements resolved by discussion or, if necessary, a third independent reviewer.
- The evidence review was quality assured by a senior member of the systematic reviewing team. This included checking:
  - papers were included or excluded appropriately
  - a sample of the data extractions
  - a sample of the risk of bias assessments
  - correct methods were used to synthesise data.

  Discrepancies were identified and resolved through discussion (with a third reviewer where necessary).

2.3.1 Types of studies and inclusion and exclusion criteria

The inclusion and exclusion of studies was based on the criteria defined in the review protocol, which can be found in an appendix to the evidence report. Excluded studies

(with the reasons for their exclusion) are also listed in an appendix to the evidence report. The committee was consulted about any uncertainty regarding inclusion or exclusion.

Conference abstracts were not considered for inclusion. Literature reviews, posters, letters, editorials, comment articles, unpublished studies and studies not published in English language were excluded.

2.3.1.1 Type of studies

Randomised controlled trials (RCTs) were included in the evidence review where identified because they are considered the most robust type of study design that can produce an unbiased estimate of the intervention effects.

Systematic reviews and meta-analyses conducted to the same methodological standards as the NICE reviews were included within the evidence reviews in preference to primary studies, where they were available and applicable to the review questions and updated or added to where appropriate to the guideline review question. Individual patient data (IPD) meta-analyses were preferentially included if meeting the protocol and methodological criteria.

2.4 Methods of combining evidence

2.4.1 Data synthesis for intervention reviews

Meta-analyses were conducted using Cochrane Review Manager (RevMan5)[17] software.

2.4.1.1 Pairwise meta-analysis

Dichotomous outcomes

Fixed-effects (Mantel–Haenszel) techniques were used to calculate risk ratios (relative risk, RR) for the binary outcomes. The absolute risk difference was also calculated using GRADEpro[8] GDT software, using the median event rate in the control arm of the pooled results.

For binary variables where there were zero events in either arm or a less than 1% event rate, Peto odds ratios, rather than risk ratios, were calculated as they are more appropriate for data with a low number of events. Where there are zero events in both arms, the risk difference was calculated and reported instead. However, in cases where meta-analysis of an outcome included small studies with zero events, and a large study that had recorded events, risk ratios were used to avoid diluting the effect of the studies that were powered to report the outcome.

Time to event data

Where sufficient information was provided, hazard ratios were reported in addition to risk ratios for the outcome of major adverse cardiovascular events, where the time to the event occurring was important for decision-making. Both hazard ratios and risk ratios were presented, but only one measure was considered for decision making. As the majority of studies reported data to calculate the risk ratio rather than hazard ratio, the committee used the risk ratio for decision making in order to maximise the available pooled data. If there were differences in effect estimates between the two

measures, potential reasons for this were considered in the interpretation of the evidence.

Continuous outcomes

Continuous outcomes were analysed using an inverse variance method for pooling weighted mean differences.

Where the studies within a single meta-analysis reported either change from baseline or final values these were pooled, with a preference for the change scores if a single study reported both.

The means and standard deviations of continuous outcomes are required for meta-analysis. However, in cases where standard deviations were not reported, the standard error was calculated if the p-values or 95% confidence intervals (95% CI) were reported, and meta-analysis was undertaken with the mean and standard error using the generic inverse variance method in RevMan5.[17]

Generic inverse variance

If a study reported only the summary statistic and 95% CI the generic-inverse variance method was used to enter data into RevMan5.[17] If the control event rate was reported this was used to generate the absolute risk difference in GRADEpro.[8] If multivariate analysis was used to derive the summary statistic but no adjusted control event rate was reported no absolute risk difference was calculated.

2.4.1.2 Network meta-analysis

Network meta-analysis (NMA) was conducted by the NICE Technical Support Unit for the continuous outcomes of percentage and absolute change in both LDL cholesterol and non-HDL cholesterol.

The aim of the NMA was to include all relevant evidence on these outcomes in order to;

- estimate the clinical effectiveness of all interventions compared to each other (including for treatment comparisons which have not yet been directly compared in a trial) in terms of reducing serum cholesterol; and
- rank of treatments in terms of clinical effectiveness in reducing serum cholesterol.

A hierarchical Bayesian NMA was performed using the software WinBUGS version 1.4.[18] Models were based on the methods of the NICE Decision Support Unit.[1-7]

NMA models were fitted for LDL cholesterol (LDL-C; mmol/L) and for non-HDL cholesterol (non-HDL-C; mmol/L) data. Studies varied in whether they reported the outcomes as a percentage change from baseline or as an absolute outcome change from baseline (CFB), and some studies reported both. It is not possible to combine these two different outcome measure formats, and so a NMA was conducted for both outcome formats separately. Some studies did not report enough information to estimate a standard error of the outcome measure. To include the results from these studies, a standard deviation was imputed and then the standard error was formatted using the sample sizes from the study. Standard deviations were imputed by fitting a hierarchical model to the logged standard deviations from studies that did report them, and then using the mean of that distribution (unlogged).

Some studies reported relative effects between treatments rather than arm-level data, so all study data were converted into the relative treatment effects format, and then pooled using the TSD2-7aRE_NormalDiff_id.odc model code from the NICE DSU Technical Support Document 2.[2] This code accounts for correlations between relative effects from studies with 3 or more arms. For the NMAs of absolute outcomes, some studies reported change from baseline, some reported baseline and follow-up measures, and some reported follow-up outcomes only. For studies which reported baseline and follow-up measures, the mean change from baseline was calculated, and the standard error of the mean change from baseline was estimated by assuming a correlation between baseline and follow-up measures which was estimated from studies which reported all of baseline, follow-up and change from baseline summaries. For studies reporting follow-up data only, the mean differences between treatments were combined with the studies providing treatment difference in mean change from baseline. This assumes that the studies reporting follow-up data only were balanced at baseline and randomisation was conducted adequately. Where studies reported lipid levels as mg/dL, these were converted to mmol/L using a conversion factor of 0.02586.[9, 10]

Fixed effect (FE) and random effects (RE) models were fitted to explore whether there was evidence of between study heterogeneity. Model choice was based on goodness of fit measures (posterior mean deviance and Deviance Information Criteria (DIC), preferring lower values), and inspection of the estimated between studies standard deviation.

NMA assumes that the included studies are similar in terms of factors that might interact with the intervention effects (effect modifiers). So, the relative effect of intervention B vs intervention A would be expected to be similar in all of the studies (if they had included A and B interventions). This assumption was assessed by measuring statistical heterogeneity, and also by checking if the direct and indirect estimates are in agreement when there are loops of evidence in the network. To determine if there is evidence of inconsistency, the selected consistency model (fixed or random effects) was compared to an “inconsistency”, or unrelated mean effects (UME), model.[2] The latter is equivalent to having separate, unrelated, meta-analyses for every pairwise contrast, with a common variance parameter assumed in the case of random effects models.

Convergence was satisfactory for all NMA models after a burn-in of 60,000 iterations and results are based on at least a further 24,000 samples from three chains.

2.5 Appraising the quality of evidence by outcomes

2.5.1 Intervention reviews – pairwise analyses

The evidence for outcomes from the included RCTs were evaluated and presented using the ‘Grading of Recommendations Assessment, Development and Evaluation (GRADE) toolbox’ developed by the international GRADE working group (http://www.gradeworkinggroup.org/). The software (GRADEpro[8]) developed by the GRADE working group was used to assess the quality of each outcome, taking into account individual study quality and the meta-analysis results.

Each outcome was first examined for each of the quality elements listed and defined in Table 1.

Table 1: Description of quality elements in GRADE for intervention studies

Quality element	Description
Risk of bias	Limitations in the study design and implementation may bias the estimates of the treatment effect. Major limitations in studies decrease the confidence in the estimate of the effect. Examples of such limitations are selection bias (often due to poor allocation concealment), performance and detection bias (often due to a lack of blinding of the patient, healthcare professional or assessor) and attrition bias (due to missing data causing systematic bias in the analysis).
Indirectness	Indirectness refers to differences in study population, intervention, comparator and outcomes between the available evidence and the review question.
Inconsistency	Inconsistency refers to an unexplained heterogeneity of effect estimates between studies in the same meta-analysis.
Imprecision	Results are imprecise when studies include relatively few patients and few events (or highly variable measures) and thus have wide confidence intervals around the estimate of the effect relative to clinically important thresholds. 95% confidence intervals denote the possible range of locations of the true population effect at a 95% probability, and so wide confidence intervals may denote a result that is consistent with conflicting interpretations (for example a result may be consistent with both clinical benefit AND clinical harm) and thus be imprecise.
Publication bias	Publication bias is a systematic underestimate or overestimate of the underlying beneficial or harmful effect due to the selective publication of studies. A closely related phenomenon is where some papers fail to report an outcome that is inconclusive, thus leading to an overestimate of the effectiveness of that outcome.
Other issues	Sometimes randomisation may not adequately lead to group equivalence of confounders, and if so this may lead to bias, which should be taken into account. Potential conflicts of interest, often caused by excessive pharmaceutical company involvement in the publication of a study, should also be noted.

Details of how the 4 main quality elements (risk of bias, indirectness, inconsistency and imprecision) were appraised for each outcome are given below. Publication bias was considered with the committee, but was not suspected to be present in any of the analyses.

2.5.1.1 Risk of bias

The main domains of bias for RCTs are listed in Table 2. Each outcome had its risk of bias assessed within each study first using the appropriate checklist for the study design (Cochrane RoB 2 for RCTs, or ROBIS for systematic reviews). For each study, if there was no risk of bias in any domain, the risk of bias was given a rating of 'low risk of bias'. An overall judgment of 'some concerns' was made if some concerns were present in at least one domain for this outcome, and no domain was judged to be at high risk of bias. An overall judgment of 'high risk of bias' was made if high risk of bias was present in at least one domain or there were 'some concerns' for multiple domains in a way that substantially lowers confidence in the result. An overall rating of; not serious, serious or very serious, is applied in GRADEpro across all studies combined in a meta-analysis by taking into account the weighting of studies according to study precision.

Table 2: Principle domains of bias in randomised controlled trials

Limitation	Explanation
Selection bias (sequence	If those enrolling participants are aware of the group to which the next enrolled patient will be allocated, either because of a non-random sequence that is

Limitation	Explanation
generation and allocation concealment)	predictable, or because a truly random sequence was not concealed from the researcher, this may translate into systematic selection bias. This may occur if the researcher chooses not to recruit a participant into that specific group because of: • knowledge of that participant's likely prognostic characteristics, and • a desire for one group to do better than the other.
Performance and detection bias (lack of blinding)	Patients, caregivers, those adjudicating or recording outcomes, and data analysts should not be aware of the arm to which the participants are allocated. Knowledge of the group can influence: • the experience of the placebo effect • performance in outcome measures • the level of care and attention received, and • the methods of measurement or analysis all of which can contribute to systematic bias.
Attrition bias	Attrition bias results from an unaccounted for loss of data beyond a certain level (a differential of at least 10% between groups). Loss of data can occur when participants are compulsorily withdrawn from a group by the researchers (for example, when a per-protocol approach is used) or when participants do not attend assessment sessions. If the missing data are likely to be different from the data of those remaining in the groups, and there is a differential rate of such missing data from groups, systematic attrition bias may result.
Selective outcome reporting	Reporting of some outcomes and not others on the basis of the results can also lead to bias, as this may distort the overall impression of efficacy.
Other limitations	For example: • Stopping early for benefit observed in randomised trials, in particular in the absence of adequate stopping rules. • Use of unvalidated patient-reported outcome measures. • Lack of washout periods to avoid carry-over effects in crossover trials. • Recruitment bias in cluster-randomised trials.

2.5.1.2 Indirectness

Indirectness refers to the extent to which the populations, interventions, comparisons and outcome measures are dissimilar to those defined in the inclusion criteria for the reviews. Indirectness is important when these differences are expected to contribute to a difference in effect size, or may affect the balance of harms and benefits considered for an intervention. As for the risk of bias, each outcome had its indirectness assessed within each study first. For each study, if there were no sources of indirectness, indirectness was given a rating of 'directly applicable'. If there was indirectness in just 1 source (for example in terms of population), indirectness was given a rating of partially applicable, but if there was indirectness in 2 or more sources (for example, in terms of population and treatment) the indirectness was given an 'indirectly applicable' rating. An overall rating of; not serious, serious or very serious, was applied in GRADEpro across all studies combined in a meta-analysis by taking into account the weighting of studies according to study precision.

2.5.1.3 Inconsistency

Inconsistency refers to an unexplained heterogeneity of results for an outcome across different studies. When estimates of the treatment effect across studies differ

widely, this suggests true differences in the underlying treatment effect, which may be due to differences in populations, settings or doses. Statistical heterogeneity was assessed for each meta-analysis estimate by an I-squared (I^2) inconsistency statistic.

Heterogeneity or inconsistency amongst studies was also visually inspected. Where statistical heterogeneity as defined above was present or there was clear visual heterogeneity not captured in the I^2 value, predefined subgrouping of studies was carried out according to the protocol. See the review protocol for the subgrouping strategy.

When heterogeneity existed within an outcome (I^2>50%), but no plausible explanation could be found, the quality of evidence for that outcome was downgraded. Inconsistency for that outcome was given a 'serious' rating if the I^2 was 50–74%, and a 'very serious' rating if the I^2 was 75% or more.

If inconsistency could be explained based on pre-specified subgroup analysis (that is, each subgroup had an I^2<50%) then each of the derived subgroups were presented separately for that analysis for both the forest plot and GRADE (providing at least 2 studies remained in each subgroup). The committee took this into account and considered whether to make separate recommendations based on the variation in effect across subgroups within the same outcome. In such a situation the quality of evidence was not downgraded.

If all predefined strategies of subgrouping were unable to explain statistical heterogeneity, then a random effects (DerSimonian and Laird) model was employed to the entire group of studies in the meta-analysis. A random-effects model assumes a distribution of populations, rather than a single population. This leads to a widening of the confidence interval around the overall estimate.

2.5.1.4 Imprecision

The criteria applied for imprecision were based on the 95% CIs for the pooled estimate of effect, and the minimal important differences (MID) for the outcome. The MIDs are the threshold for appreciable benefits and harms, separated by a zone either side of the line of no effect where there is assumed to be no clinically important effect. If either end of the 95% CI of the overall estimate of effect crossed 1 of the MID lines, imprecision was regarded as serious in the GRADEpro rating. This was because the overall result, as represented by the span of the confidence interval, was consistent with 2 interpretations as defined by the MID (for example, both no clinically important effect and clinical benefit were possible interpretations). If both MID lines were crossed by either or both ends of the 95% CI then imprecision was regarded as very serious. This was because the overall result was consistent with all 3 interpretations defined by the MID (no clinically important effect, clinical benefit and clinical harm). This is illustrated in Figure 1.

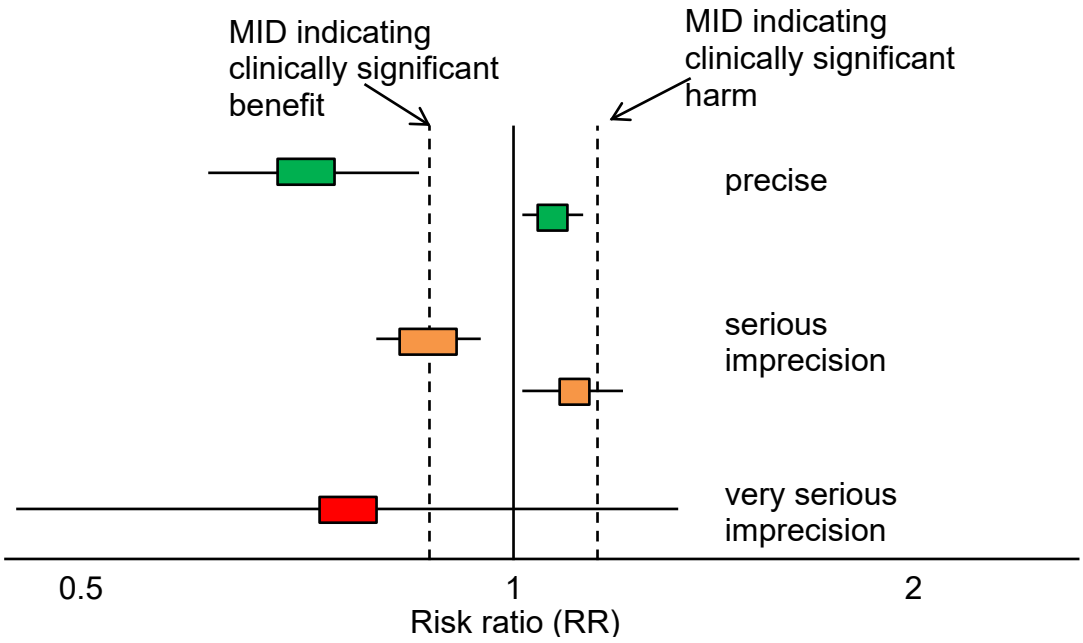
The value / position of the MID lines is ideally determined by values reported in the literature. 'Anchor-based' methods aim to establish clinically meaningful changes in a continuous outcome variable by relating or 'anchoring' them to patient-centred measures of clinical effectiveness that could be regarded as gold standards with a high level of face validity. For example, a MID for an outcome could be defined by the minimum amount of change in that outcome necessary to make patients feel their quality of life had 'significantly improved'. MIDs in the literature may also be based on expert clinician or consensus opinion concerning the minimum amount of change in a variable deemed to affect quality of life or health.

In the absence of values identified in the literature, the alternative approach to deciding on MID levels is to use the modified GRADE 'default' values, as follows:

- For dichotomous outcomes the MIDs were taken to be RRs of 0.8* and 1.25. For 'positive' outcomes such as 'patient satisfaction', the RR of 0.8 is taken as the line denoting the boundary between no clinically important effect and a clinically important harm, whilst the RR of 1.25 is taken as the line denoting the boundary between no clinically important effect and a clinically important benefit. For 'negative' outcomes such as 'bleeding', the opposite occurs, so the RR of 0.8 is taken as the line denoting the boundary between no clinically important effect and a clinically important benefit, whilst the RR of 1.25 is taken as the line denoting the boundary between no clinically important effect and a clinically important harm. There aren't established default values for ORs, and the same values (0.8 and 1.25) were applied here but were acknowledged as arbitrary thresholds agreed by the committee.
    - In cases where there were zero events in one arm of a single study, or some or all of the studies in one arm of a meta-analysis, the same process was followed as for dichotomous outcomes. However, if there were no events in either arm in a meta-analysis (or in a single unpooled study) the sample size was used to determine imprecision using the following rule of thumb:
        - No imprecision: sample size ≥350
        - Serious imprecision: sample size ≥70 but <350
        - Very serious imprecision: sample size <70.
    - When there was more than one study in an analysis and zero events occurred in both groups for some but not all of the studies across both arms, the optimum information size was calculated using the sample size in each group and was used to determine imprecision using the following guide:
        - No imprecision: >90% power
        - Serious imprecision: 80-90% power
        - Very serious imprecision: <80% power.
- For time to event data, there aren't established default values for HRs so the same values as dichotomous outcomes were applied here (0.8 and 1.25) but were acknowledged as arbitrary thresholds agreed by the committee.
- For continuous outcome variables the MID was taken as half the median baseline standard deviation of that variable, across all studies in the meta-analysis. Hence the MID denoting the minimum clinically important benefit was positive for a 'positive' outcome (for example, a quality of life measure where a higher score denotes better health), and negative for a 'negative' outcome (for example, a visual analogue scale [VAS] pain score). Clinically important harms were the converse of these. If baseline values were unavailable, then half the median comparator group standard deviation of that variable was taken as the MID. As these vary for each outcome per review, details of the values used are reported in the footnotes of the relevant GRADE summary table.

*NB GRADE report the default values as 0.75 and 1.25. These are consensus values. This guideline follows NICE process to use modified values of 0.8 and 1.25 as they are symmetrical on a relative risk scale. For this guideline, no appropriate MIDs for continuous or dichotomous outcomes were found in the literature, and so the default method was adopted.

Figure 1: Illustration of precise and imprecise outcomes based on the 95% CI of dichotomous outcomes in a forest plot (Note that all 3 results would be pooled estimates, and would not, in practice, be placed on the same forest plot)



2.5.1.5 Overall grading of the quality of clinical evidence

Once an outcome had been appraised for the main quality elements, as above, an overall quality, or certainty rating was determined for that outcome. The ratings from each of the main quality elements were summed to give an overall rating from high to very low. The evidence for each outcome started at High, and the overall quality (or confidence in the evidence) remained High if there were no reasons for downgrading, or became Moderate, Low or Very Low according to the number of independent reasons for downgrading. The significance of these overall ratings is explained in Table 3. The reasons for downgrading in each case are specified in the footnotes of the GRADE tables.

Table 3: Overall quality of outcome evidence in GRADE

Level	Description
High	Further research is very unlikely to change our confidence in the estimate of effect
Moderate	Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate
Low	Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate
Very low	Any estimate of effect is very uncertain

2.5.2 Intervention studies – NMA

Risk of bias and indirectness for the included studies was assessed as described above for the pairwise analyses and was presented in a summary table for the studies and outcomes included in the NMA.

Inconsistency in the network was assessed as described in section 2.4.1.2 on network meta-analysis above.

Imprecision was considered by the committee when interpreting the evidence by assessing whether the 95% credible intervals had an impact on their ability to draw conclusions from the results on the NMA.

2.5.2.1 Modified GRADE for intervention studies analysed using NMA

A modified version of the standard GRADE approach for pairwise interventions was used to assess the quality of evidence across the network meta-analyses. While most criteria for pairwise meta-analyses still apply, it is important to adapt some of the criteria to take into consideration additional factors, such as how each 'link' or pairwise comparison within the network applies to the others. As a result, the following was used when modifying the GRADE framework to a network meta-analysis. It is designed to provide a single overall quality rating for an NMA to judge the overall strength of evidence.

Table 4: Rationale for downgrading quality of evidence for network meta-analysis

GRADE criteria	Reasons for downgrading quality
Risk of bias	Not serious: If fewer than 50% of the study participants in the network meta-analysis were judged to have some concerns or high risk of bias, the overall network was not downgraded. Serious: If greater than 50% of the study participants in the network meta-analysis were judged to have some concerns, the network was downgraded one level. Very serious: If greater than 50% of the study participants in the network meta-analysis were at high risk of bias, the network was downgraded two levels.
Indirectness	Not serious: If fewer than 50% of the studies in the network meta-analysis were partially indirect or indirect, the overall network was not downgraded. Serious: If greater than 50% of the studies in the network meta-analysis had serious indirectness, the network was downgraded one level. Very serious: If greater than 50% of the studies in the network meta-analysis had very serious indirectness, the network was downgraded two levels.
Inconsistency	N/A: Inconsistency was marked as not applicable if there were no links in the network where data from multiple studies (either direct or indirect) were synthesised. For network meta-analyses conducted under a Bayesian framework, the network was downgraded one level if the DIC for an inconsistency model was more than 3 points lower than the corresponding consistency model or, for a random effects model, the between studies standard deviation was meaningfully lower for the inconsistency model that the corresponding consistency model.
Imprecision	This was not included in the GRADE table, but was considered during committee discussions of the evidence, taking into account 95% confidence intervals around the point estimate of the effect, any relevant MIDs, committee expertise and the effect of a single intervention based on multiple outcomes.

2.6 Identifying and analysing evidence of cost effectiveness

The committee is required to make decisions based on the best available evidence of both clinical effectiveness and cost effectiveness. Guideline recommendations should be based on the expected costs of the different options in relation to their expected health benefits (that is, their 'cost effectiveness') rather than the total implementation cost. However, the committee will also need to be increasingly confident in the cost effectiveness of a recommendation as the cost of implementation increases. Therefore, the committee may require more robust evidence on the effectiveness and cost effectiveness of any recommendations that are expected to have a substantial impact on resources; any uncertainties must be offset by a compelling argument in favour of the recommendation. The cost impact or savings potential of a recommendation should not be the sole reason for the committee's decision.[12]

Health economic evidence was sought relating to the key clinical issues being addressed in the guideline. Health economists:

- Undertook a systematic review of the published economic evaluations of different cholesterol treatment thresholds or targets.
- Undertook new cost-utility analysis.

2.6.1 Literature review

The health economists would have:

- Identified potentially relevant studies from the health economic search results by reviewing titles and abstracts. Full papers were then obtained.
- Reviewed full papers against prespecified inclusion and exclusion criteria to identify relevant studies (see below for details).
- Critically appraised relevant studies using economic evaluations checklists as specified in the NICE guidelines manual.[12]
- Extracted key information about the studies' methods and results into health economic evidence tables.
- Generated summaries of the evidence in NICE health economic evidence profile tables.

However, no studies were identified.

2.6.1.1 Inclusion and exclusion criteria

Full economic evaluations (studies comparing costs and health consequences of alternative courses of action: cost–utility, cost-effectiveness, cost–benefit and cost–consequences analyses) and comparative costing studies that addressed the review question in the relevant population were considered potentially includable as health economic evidence.

Studies that only reported cost per hospital (not per patient), or only reported average cost effectiveness without disaggregated costs and effects were excluded. Literature reviews, abstracts, posters, letters, editorials, comment articles, unpublished studies and studies not in English were excluded. Studies published before 2007 and studies from non-OECD countries or the USA were also excluded, on the basis that the applicability of such studies to the present UK NHS context is likely to be too low for them to be helpful for decision-making.

For more details about the assessment of applicability and methodological quality see Table 5 below and the economic evaluation checklist (appendix H of the NICE guidelines manual[12]) and the health economics review protocol, which can be found in the evidence report.

However, no relevant studies were identified.

2.6.1.2 NICE health economic evidence profiles

NICE health economic evidence profile tables were used to summarise cost and cost-effectiveness estimates for the health economic model. The health economic evidence profile shows an assessment of applicability and methodological quality for each economic study, with footnotes indicating the reasons for the assessment. These assessments were made by the health economist using the economic evaluation checklist from the NICE guidelines manual.[12] It also shows the incremental costs, incremental effects (for example, quality-adjusted life years [QALYs]) and incremental cost-effectiveness ratio (ICER) for the base case analysis, as well as information about the assessment of uncertainty in the analysis. See Table 5 for more details.

Table 5: Content of NICE health economic evidence profile

Item	Description
Study	Surname of first author, date of study publication and country perspective with a reference to full information on the study.
Applicability	An assessment of applicability of the study to this guideline, the current NHS situation and NICE decision-making:[(a)] • Directly applicable – the study meets all applicability criteria, or fails to meet 1 or more applicability criteria but this is unlikely to change the conclusions about cost effectiveness. • Partially applicable – the study fails to meet 1 or more applicability criteria, and this could change the conclusions about cost effectiveness. • Not applicable – the study fails to meet 1 or more of the applicability criteria, and this is likely to change the conclusions about cost effectiveness. Such studies would usually be excluded from the review.
Limitations	An assessment of methodological quality of the study:[(a)] • Minor limitations – the study meets all quality criteria, or fails to meet 1 or more quality criteria, but this is unlikely to change the conclusions about cost effectiveness. • Potentially serious limitations – the study fails to meet 1 or more quality criteria, and this could change the conclusions about cost effectiveness. • Very serious limitations – the study fails to meet 1 or more quality criteria, and this is highly likely to change the conclusions about cost effectiveness. Such studies would usually be excluded from the review.
Other comments	Information about the design of the study and particular issues that should be considered when interpreting it.
Incremental cost	The mean cost associated with one strategy minus the mean cost of a comparator strategy.
Incremental effects	The mean QALYs (or other selected measure of health outcome) associated with one strategy minus the mean QALYs of a comparator strategy.
Cost effectiveness	Incremental cost-effectiveness ratio (ICER): the incremental cost divided by the incremental effects (usually in £ per QALY gained).

Item	Description
Uncertainty	A summary of the extent of uncertainty about the ICER reflecting the results of deterministic or probabilistic sensitivity analyses, or stochastic analyses of trial data, as appropriate.

(a) Applicability and limitations were assessed using the economic evaluation checklist in appendix H of the NICE guidelines manual[12]

2.6.2 Undertaking new health economic analysis

New health economic analysis was undertaken by the health economists.

- The model compares costs and QALYs for different LDL cholesterol (or non-HDL cholesterol) treatment thresholds for people with CVD and receiving a statin.
- The population has a baseline distribution of LDL cholesterol (or non-HDL cholesterol). At lower LDL (or non-HDL) treatment thresholds, more of the cohort will need to be escalated to ezetimibe or an injectable therapy.
- Cardiovascular events are being modelled based on LDL cholesterol (or non-HDL cholesterol) and age and sex.
- Costs include the cost of therapy for those people that are escalated and cost savings from subsequent cardiovascular events averted.

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

- Methods were consistent with the NICE reference case for interventions with health outcomes in NHS settings.[12, 15]
- The committee was involved in the design of the model, selection of inputs and interpretation of the results.
- Model inputs were based on the systematic review of the clinical literature supplemented with other published data sources where possible.
- 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.
- The model was peer-reviewed by another health economist.

Full methods and results of the cost-effectiveness analysis are described in a separate economic analysis report.

2.6.3 Population characteristics and event rates

For the model baseline descriptive statistics were calculated for people with CVD and prescribed with a statin, using the Clinical Practice Research Datalink (CPRD) linked to the Hospital Episode Statistics (HES) admitted patient care and Office of National Statistics (ONS) mortality datasets.

The following analyses were conducted:

- LDL cholesterol or non-HDL cholesterol distribution by age and sex
- Incidence of CVD events, by age and sex
- Incidence of CVD and non-CVD mortality, by age and sex.

2.6.4 Cost-effectiveness criteria

NICE sets out the principles that committees should consider when judging whether an intervention offers good value for money.[12, 14, 16] In general, an intervention was considered to be cost effective (given that the estimate was considered plausible) if either of the following criteria applied:

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

If the committee recommended an intervention that was estimated to cost more than £20,000 per QALY gained, or did not recommend one that was estimated to cost less than £20,000 per QALY gained, the reasons for this decision are discussed explicitly in 'The committee's discussion of the evidence' section of the relevant evidence report, with reference to issues regarding the plausibility of the estimate or to factors set out in NICE methods manuals.[12]

When QALYs or life years gained are not used in the analysis, results are difficult to interpret unless one strategy dominates the others with respect to every relevant health outcome and cost.

2.7 Developing recommendations

Over the course of the guideline development process, the committee was presented with:

- Summaries of clinical and health economic evidence and quality (as presented in evidence report A).
- Evidence tables of the clinical evidence reviewed from the literature. All evidence tables can be found in an appendix to the evidence report.
- Forest plots (in an appendix to the evidence report).
- A description of the methods and results of the cost-effectiveness analysis undertaken for the guideline (in a separate economic analysis report).

Decisions on whether a recommendation could be made, and if so in which direction, were made on the basis of the committee's interpretation of the available evidence, taking into account the balance of benefits, harms and costs between different courses of action. This was either done formally in an economic model, or informally. The net clinical benefit over harm (clinical effectiveness) was considered, focusing on the magnitude of the effect (or clinical importance), quality of evidence (including the uncertainty) and amount of evidence available. When this was done informally, the committee took into account the clinical benefits and harms when course of action was compared with another. The assessment of net clinical benefit was moderated by the importance placed on the outcomes (the committee's values and preferences), and the confidence the committee had in the evidence (evidence quality). Secondly, the committee assessed whether the net clinical benefit justified any differences in costs between the alternative approaches. When the clinical harms were judged by the committee to outweigh any clinical benefits, they considered making a recommendation not to offer an intervention. This was dependant on whether the intervention had any reasonable prospect of providing cost-effective benefits to people using services and whether stopping the intervention was likely to cause harm

for people already receiving it. When clinical and health economic evidence was of poor quality, conflicting or absent, the committee decided on whether a recommendation could be made based on its expert opinion. The considerations for making consensus-based recommendations include the balance between potential harms and benefits, the economic costs compared to the economic benefits, current practices, recommendations made in other relevant guidelines, patient preferences and equality issues. The consensus recommendations were agreed through discussions in the committee. The committee also considered whether the uncertainty was sufficient to justify delaying making a recommendation to await further research, taking into account the potential harm of failing to make a clear recommendation (see section 2.7.1 below).

The committee considered the appropriate 'strength' of each recommendation. This takes into account the quality of the evidence but is conceptually different. Some recommendations are 'strong' in that the committee believes that the vast majority of healthcare and other professionals and patients would choose a particular intervention if they considered the evidence in the same way that the committee has. This is generally the case if the benefits clearly outweigh the harms for most people and the intervention is likely to be cost effective. However, there is often a closer balance between benefits and harms, and some patients would not choose an intervention whereas others would. This may happen, for example, if some patients are particularly averse to some side effect and others are not. In these circumstances the recommendation is generally weaker, although it may be possible to make stronger recommendations about specific groups of patients.

The committee focused on the following factors in agreeing the wording of the recommendations:

- The actions health professionals need to take.
- The information readers need to know.
- The strength of the recommendation (for example the word 'offer' was used for strong recommendations and 'consider' for weaker recommendations).
- The involvement of patients (and their carers if needed) in decisions on treatment and care.
- Consistency with NICE's standard advice on recommendations about drugs, waiting times and ineffective interventions (see section 9.2 in the NICE guidelines manual[13]).

The main considerations specific to each recommendation are outlined in 'The committee's discussion of the evidence' section within each evidence report.

2.7.1 Research recommendations

When areas were identified for which good evidence was lacking, the committee considered making recommendations for future research. Decisions about the inclusion of a research recommendation were based on factors such as:

- the importance to patients or the population
- national priorities
- potential impact on the NHS and future NICE guidance
- ethical and technical feasibility.

2.7.2 Validation process

This guidance is subject to a 2-week public consultation and feedback as part of the quality assurance and peer review of the document. All comments received from registered stakeholders are responded to in turn and posted on the NICE website.

2.7.3 Updating the guideline

Following publication, and in accordance with the NICE guidelines manual, NICE will undertake a review of whether the evidence base has progressed significantly to alter the guideline recommendations and warrant an update.

References

1. Ades AE, Caldwell DM, Reken S, Welton NJ, Sutton AJ, Dias S. Evidence synthesis for decision making 7: a reviewer's checklist. Medical Decision Making. 2013; 33(5):679-691

2. Dias S, Sutton AJ, Ades AE, Welton NJ. Evidence synthesis for decision making 2: a generalized linear modeling framework for pairwise and network meta-analysis of randomized controlled trials. Medical Decision Making. 2013; 33(5):607-617

3. Dias S, Sutton AJ, Welton NJ, Ades AE. Evidence synthesis for decision making 3: heterogeneity--subgroups, meta-regression, bias, and bias-adjustment. Medical Decision Making. 2013; 33(5):618-640

4. Dias S, Sutton AJ, Welton NJ, Ades AE. Evidence synthesis for decision making 6: embedding evidence synthesis in probabilistic cost-effectiveness analysis. Medical Decision Making. 2013; 33(5):671-678

5. Dias S, Welton NJ, Sutton AJ, Ades AE. Evidence synthesis for decision making 1: introduction. Medical Decision Making. 2013; 33(5):597-606

6. Dias S, Welton NJ, Sutton AJ, Ades AE. Evidence synthesis for decision making 5: the baseline natural history model. Medical Decision Making. 2013; 33(5):657-670

7. Dias S, Welton NJ, Sutton AJ, Caldwell DM, Lu G, Ades AE. Evidence synthesis for decision making 4: inconsistency in networks of evidence based on randomized controlled trials. Medical Decision Making. 2013; 33(5):641-656

8. GRADE Working Group. The Grading of Recommendations Assessment, Development and Evaluation (GRADE) Working Group website. 2011. Available from: http://www.gradeworkinggroup.org/ Last accessed:

9. Haney EM, Huffman LH, Bougatsos C, Freeman M, Fu R, Steiner RD et al. Screening for Lipid Disorders in Children and Adolescents. U.S. Preventive Services Task Force Evidence Syntheses, formerly Systematic Evidence Reviews. Rockville (MD). 2007.

10. Koperska M. Cholesterol Units Converter. Available from: https://www.omnicalculator.com/health/cholesterol-units

11. McGowan J, Sampson M, Salzwedel DM, Cogo E, Foerster V, Lefebvre C. PRESS Peer Review of Electronic Search Strategies: 2015 Guideline Explanation and Exploration (PRESS E&E). Ottawa. CADTH, 2016. Available from: https://www.cadth.ca/sites/default/files/pdf/CP0015_PRESS_Update_Report_2016.pdf

12. National Institute for Health and Care Excellence. Developing NICE guidelines: the manual. London. National Institute for Health and Care Excellence, 2014. Available from: http://www.nice.org.uk/article/PMG20/chapter/1%20Introduction%20and%20overview

13. National Institute for Health and Care Excellence. Developing NICE guidelines: the manual [updated October 2018]. London. National Institute for Health and Care Excellence, 2014. Available from: http://www.nice.org.uk/article/PMG20/chapter/1%20Introduction%20and%20overview

14. National Institute for Health and Care Excellence. The NICE Charter. 2020. Available from: https://www.nice.org.uk/about/who-we-are/our-charter Last accessed: 15/11/2021.

15. National Institute for Health and Clinical Excellence. Guide to the methods of technology appraisal 2013. London. National Institute for Health and Clinical Excellence, 2013. Available from: http://publications.nice.org.uk/pmg9

16. National Institute for Health and Clinical Excellence. Social value judgements: principles for the development of NICE guidance. London. National Institute for Health and Clinical Excellence, 2008. Available from: https://www.nice.org.uk/media/default/about/what-we-do/research-and-development/social-value-judgements-principles-for-the-development-of-nice-guidance.pdf

17. Review Manager (RevMan) [Computer program]. Version 5.3. Copenhagen. The Nordic Cochrane Centre and The Cochrane Collaboration, 2014. Available from: http://tech.cochrane.org/Revman

18. WinBUGS [Computer programme] version 1.4. Cambridge. MRC Biostatistics Unit University of Cambridge, 2015. Available from: http://www.mrc-bsu.cam.ac.uk/software/bugs/the-bugs-project-winbugs/