National Institute for Health and Care Excellence

Draft for consultation

Integrated health and social care for people experiencing homelessness

NICE guideline: Methods

NICE guideline <number>

Methods

October 2021

Draft for Consultation

Evidence reviews were developed by the National Guideline Alliance hosted by the Royal College of Obstetricians and Gynaecologists

Disclaimer

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

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

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 All NICE guidance is subject to regular review and may be updated or withdrawn.

Copyright

ISBN:

Contents

Development of the guideline

Remit

The National Institute for Health and Care Excellence (NICE) commissioned the National Guideline Alliance (NGA) to develop a guideline about integrated health and social care for people experiencing homelessness.

To see “What this guideline covers” and “What this guideline does not cover” please see the final scope of the guideline on the NICE website.

Methods

This guideline was developed using the methods described in the Developing NICE guidelines: the manual.

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

Developing the review questions and outcomes

The review questions developed for this guideline were based on the key areas identified in the guideline scope. They were drafted by the NGA technical team, and refined and validated by the guideline committee.

The review questions were based on the following frameworks:

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

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

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

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

Evidence review	Review question	Type of review
[A & B] Effectiveness of approaches to improve access to and engagement with health and social care and joined up approaches[1]	[A] What approaches are effective in improving access to and/or engagement with health and social care for people who experience homelessness? AND [B] What joined up approaches are effective in responding to the health, social care and housing needs of people experiencing homelessness?	Intervention
[C] Views and experiences of health and social care for people experiencing homelessness	What works well and what could be improved about access to, engagement with and delivery of health and social care for people experiencing homelessness?	Qualitative

[1] *Original health economic analysis conducted*

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

- Supplement 1 (Methods; this document)
- Supplement 2 (Economic literature)
- Supplement 3 (NGA staff list).

Searching for evidence

Scoping search

During the scoping phase, searches were conducted for previous guidelines, economic evaluations, health technology assessments, systematic reviews, randomised controlled trials, observational studies and qualitative research. Searches of websites of organisations were also undertaken for relevant policies and related documents.

Systematic literature search

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

For Review A and Review B evidence published up to March 2020 was identified from an Evidence and Gap Map (EGM) developed by the Centre for Homelessness Impact and the Campbell Collaboration (White 2020a, White 2020b). Further details on the resources used to populate the EGM are given at the end of this section. For evidence published from March 2020 onwards, a de-novo top up search was conducted. The top up search used a combined search to cover Review A and Review B.

For Review C searches were conducted that were limited to 1999 onwards.

For Review C and the top up search for Review A and Review B databases were searched using subject headings, free-text terms and, where appropriate, study type filters. Where possible, searches were limited to retrieve studies published in English. The searches for Review C and the top up search for Review A and Review B were conducted in the following databases: Medline, Medline-in-Process, Cochrane Central Register of Controlled Trials (CCTR), Cochrane Database of Systematic Reviews (CDSR), Database of Abstracts of Reviews of Effects (DARE), Embase, Health Management Information Consortium (HMIC), Social Policy and Practice, PsycInfo, Emcare, Applied Social Sciences Index & Abstracts (ASSIA), Social Services Abstracts, Sociological Abstracts, Cumulative Index to Nursing and Allied Health Literature (CINAHL), Social Sciences Citation Index (SSCI) and Social Care Online. For the top up search for Review A and Review B the International Health Technology Assessment (IHTA) database was also searched. The webpages of the following organisations were also checked for evidence relevant to Review C and the top up search for Review A and Review B: Shelter, Groundswell, Crisis, St Mungos, Salvation Army, Centrepoint, Centre for Homelessness Impact, FEANTSA, Revolving Door, Centre for Housing Policy, Homeless Link, Kings Fund, Gov.uk, Campbell Collaboration, OpenGrey. For the top up search for Review A and Review B the websites of the following organisations were also searched: Homeless Hub, United State Interagency Council on Homelessness, Homelessness Australia and Housing First Europe Hub.

Due to the short development time for the guideline searches were run once for all reviews during development.

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

For reference, the resources used to populate the Evidence Gap Map (EGM) developed by the Centre for Homelessness Impact and the Campbell Collaboration (White 2020a) mentioned in the text above are listed below:

Systematic review: Munthe-Kaas, H.M., Berg, R.C. and Blaasvær, N. (2018), Effectiveness of interventions to reduce homelessness: a systematic review and meta-analysis. Campbell Systematic Reviews, 14: 1-281.

Academic databases: Econlit; The National Bureau of Economic Research (NBER); Social Science Research Network (SSRN); International Bibliography of Social Sciences (IBSS); Applied Social Sciences Index and Abstracts (ASSIA); Social Service Abstract; Embase; PubMed; PsycINFO; MEDLINE; WHO's Global Health Library; CABI's Global Health; ERIC; CINHAL; SCOPUS; Web of Science; EPPI Centre Evaluation Database of Education Research

Evidence and Gap Map databases: 3ie Evidence and gap map repository; Global Evidence Mapping Initiative; Evidence based Synthesis Program (Department of Veteran affairs)

Systematic review databases: Swedish Agency For Health Technology Assessment and Assessment of Social Services; Collaboration for Environmental Evidence; Cochrane; Campbell; 3ie Systematic Review Database; Research for Development; Epistemonikos

French & Norwegian Academic databases: Scholar.google.fr; Cairn.info; Persee.fr; Scholar.google.no

Websites: Homeless Hub (https://www.homelesshub.ca/); European observatory on homelessness (https://www.feantsaresearch.org/en/publications); United State interagency council on homelessness (http://www.usich.gov/); EThOS (http://ethos.bl.uk/Home.do); WHO ICTRP (http://apps.who.int/trialsearch/); Focus on Prevention (http://www.preventionfocus.net/); Social Policy and Practice (http://www.spandp.net/); 100,00000 home campaigns (https://en.wikipedia.org/wiki/100,0000_Homes_Campaign); Anti poverty committee (https://en.wikipedia.org/wiki/AntiPoverty_Committee); Back on my feet (https://en.wikipedia.org/wiki/Back_on_My_Feet_(nonprofit_organization)); Feantsa (https://www.feantsa.org/); National Coalition Homeless (https://nationalhomeless.org/); Homelessness Australia (https://www.homelessnessaustralia.org.au/); Mission Australia (https://www.missionaustralia.com.au/publications/positionstatements/homelessness) ; National Alliance to end homelessness (https://endhomelessness.org/); Institute of global homelessness (https://www.ighomelessness.org/); Homelessness link (https://www.homeless.org.uk/); Crisis (https://www.crisis.org.uk/aboutus/howwework/); Housing first (https://housingfirsteurope.eu/aboutthehub/); Canadian Alliance to end homelessness (https://housingfirsteurope.eu/aboutthehub/); Social work and policy institutes (http://www.socialworkpolicy.org/research/homelessness.html); Association of housing advice services (https://www.ahas.org.uk/); Centre point

(https://centrepoint.org.uk/); Homelessness trust funds (https://housingtrustfundproject.org/htfelements/homelesstrustfunds/); Meliville charitable trust (https://melvilletrust.org/category/resourcesreports/); Conrad H Hilton foundation (https://www.hiltonfoundation.org/priorities/homelessness#resources); Abt Associates (https://www.abtassociates.com/); Mathematica (https://www.mathematicampr.com/); American Institutes of Research (https://www.air.org/); Rand (https://www.rand.org/); MDRC (https://www.mdrc.org/)

Please note that the top up search for Review A and Review B used a narrower list of resources than was used by the Centre for Homelessness Impact and the Campbell Collaboration to populate the EGM as some resources were considered to contain material that was not relevant to the details set out in the protocols for Review A and Review B.

Economic systematic literature search

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

A single search, using the population search terms used in the evidence reviews, was conducted to identify economic evidence in the NHS Economic Evaluation Database (NHS EED) and International Health Technology Assessments (IHTA). Another single search, using the population search terms used in the evidence reviews combined with an economic evaluations search filter, was conducted in Medline, Medline-in-Process, Cochrane Central Register of Controlled Trials (CCTR), Cochrane Database of Systematic Reviews (CDSR), Database of Abstracts of Reviews of Effects (DARE), Embase, Health Management Information Consortium (HMIC), Social Policy and Practice, PsycInfo, Emcare, Applied Social Sciences Index & Abstracts (ASSIA), Social Services Abstracts, Sociological Abstracts, Cumulative Index to Nursing and Allied Health Literature (CINAHL), Social Sciences Citation Index (SSCI) and Social Care Online. Where possible, searches were limited to studies published in English. The webpages of the following organisations were also checked for economic evidence: Shelter, Groundswell, Crisis, St Mungos, Salvation Army, Centrepoint, Centre for Homelessness Impact, FEANTSA, Revolving Door, Centre for Housing Policy, Homeless Link, Kings Fund, Gov.uk, Campbell Collaboration, OpenGrey.

Due to the short development time for the guideline the economic literature searches were run once for all reviews during development.

Details of the search strategies, including the study-design filter used and databases searched are provided in Supplement 2 (Economic literature).

Quality assurance

Search strategies were quality assured by cross-checking reference lists of relevant studies, analysing search strategies from published systematic reviews and asking members of the committee to highlight key studies. The principal search strategies for each search were also quality assured by a second information scientist using an adaptation of the PRESS 2015 Guideline Evidence-Based Checklist (McGowan 2016).

Reviewing research evidence

Systematic review process

The evidence was reviewed in accordance with the following approach.

- For the quantitative reviews, potentially relevant articles were identified by their title and abstract from the Centre for Homelessness Impact and Campbell Collaboration Evidence and Gap Map plus the NGA's top up searches (described above under 'Searching for evidence'). Full-text copies of the articles were then obtained.
- The review team were aware of two systematic reviews also derived from the Evidence and Gap Map, which were to be published imminently (Miller 2019 [protocol], Keenan 2021). They liaised with the review authors who offered to contribute their analysis to the quantitative reviews for this guideline. However there were some differences in the analytical approach used for those systematic reviews, which meant they would not have fully supported the committee in making recommendations for practice. The technical team therefore did not use the materials from the above reviews and conducted their own analysis of quantitative data, according to NICE methodology, as described in this supplement.
- For the qualitative review, potentially relevant articles were identified from the search results by screening titles and abstracts. Full-text copies of the articles were then obtained.
- Full-text articles were reviewed against pre-specified inclusion and exclusion criteria in the review protocol (see Appendix A of each evidence review).
- Key information was extracted from each article on study methods and results, in accordance with factors specified in the review protocol. The information was presented in a summary table in the corresponding evidence review and in more detailed evidence tables (see Appendix D of each evidence review).
- Included studies were critically appraised using an appropriate checklist as specified in Developing NICE guidelines: the manual. Further detail on appraisal of the evidence is provided below.
- Summaries of effectiveness evidence by outcome and qualitative evidence by theme were presented in the corresponding evidence review and discussed by the committee.

Review questions were subject to dual screening and study selection through a 10% random sample of articles, as described in the review protocols. Any discrepancies were resolved by discussion between the first and second reviewers or by reference to a third (senior) reviewer.

Drafts of all evidence reviews were quality assured by a senior reviewer.

Type of studies and inclusion/exclusion criteria

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

Systematic reviews with meta-analyses or meta-syntheses were considered to be the highest quality evidence that could be selected for inclusion.

For the intervention reviews, randomised controlled trials (RCTs) were prioritised for inclusion because they are considered to be the most robust type of study design that could produce an unbiased estimate of intervention effects. Non-randomised studies (NRS) were also considered for inclusion as long as they were designed with matched comparisons or another method of controlling for confounding variables. In the absence of experimental studies (randomised or non-randomised assignment) about one of the interventions of interest, UK observational studies were also considered, providing that confounding factors were controlled for.

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

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

Narrative reviews, posters, letters, editorials, comment articles, unpublished studies and studies published in languages other than English were excluded. Conference abstracts were not considered for inclusion because conference abstracts typically do not have sufficient information to allow for full critical appraisal.

Methods of combining evidence

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

Data synthesis for intervention studies

Pairwise meta-analysis

Meta-analysis to pool results from comparative intervention studies was conducted where possible using Cochrane Review Manager (RevMan5) software.

For dichotomous outcomes, such as mortality, the Mantel–Haenszel method with a fixed effect model was used to calculate risk ratios (RRs). For all outcomes with zero events in both arms the risk difference was presented. For outcomes in which the majority of studies had low event rates (<1%), Peto odds ratios (ORs) were calculated as this method performs well when events are rare (Bradburn 2007).

For continuous outcomes, measures of central tendency (mean) and variation (standard deviation; SD) are required for meta-analysis. Data for continuous outcomes, such as quality of life, were meta-analysed using an inverse-variance method for pooling weighted mean differences (WMDs). Where SDs were not reported for each intervention group, the standard error (SE) of the mean difference was calculated from other reported statistics (p values or 95% confidence intervals; CIs), assuming the same SD for both groups and then meta-analysis was conducted as described above.

If a study reported only the summary statistic and 95% CI the generic-inverse variance method was used to enter data into RevMan5. If the control event rate was

reported this was used to generate the absolute risk difference in GRADEpro. If multivariable analysis was used to derive the summary statistic but no adjusted control event rate was reported, no absolute risk difference was calculated.

When meta-analysis was undertaken, the results were presented visually using forest plots generated using RevMan5 (see Appendix E of evidence review A and B).

Data synthesis for qualitative reviews

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

The technical team were guided in their data extraction, synthesis and formulation of review findings, or themes, by a framework of phenomena developed by the guideline committee. This framework consisted of the themes that the committee anticipated would be covered by the included studies and these were set out a priori in the corresponding review protocol. As well as guiding the data extraction and synthesis, the framework also underpinned the approach referred to in the protocol as ‘thematic saturation’. Essentially, data or themes from included studies would not be extracted if they contributed to review findings which were judged to be ‘adequate’ and ‘coherent’ following assessment using the GRADE-CERQual approach; that is, they were not downgraded for either domain. Themes identified from the included studies, which were not set out in the protocol but which were considered relevant to answering the review question, were also extracted and the same approach to ‘thematic saturation’ would have been applied. In this qualitative review, ‘thematic saturation’ was reached for 2 themes, resulting in the exclusion of 1 study because no other relevant data were reported in that paper. This is described in the excluded studies list in appendix J of the qualitative review.

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

Emerging themes were placed into a series of thematic maps representing the relationship between themes and overarching categories. The purpose of these maps is to show relationships between overarching categories and associated themes.

Appraising the quality of evidence

Intervention studies

Pairwise meta-analysis

GRADE methodology for intervention reviews

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

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

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

The initial quality rating was based on the study design: RCTs start as 'high' quality evidence as do NRS assessed by ROBINS-I, other non-randomised studies start as 'low' quality evidence. The rating was then modified according to the assessment of each quality element (Table 2). Each quality element considered to have a 'serious' or 'very serious' quality issue was downgraded by 1 or 2 levels respectively (for example, evidence starting as 'high' quality was downgraded to 'moderate' or 'low' quality). In addition, there was a possibility to upgrade evidence from non-randomised studies (provided the evidence for that outcome had not previously been downgraded) if there was a large magnitude of effect, a dose–response gradient, or if all plausible confounding would reduce a demonstrated effect or suggest a spurious effect when results showed no effect.

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

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

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

Table 3: GRADE quality ratings (by quality element)

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

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

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

Assessing risk of bias in intervention reviews

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

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

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

- selection bias
- performance bias
- attrition bias
- detection bias
- reporting bias.

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

More details about the Cochrane risk of bias tool can be found in Section 8 of the Cochrane Handbook for Systematic Reviews of Interventions (Higgins 2011).

For systematic reviews of RCTs the AMSTAR checklist was used and for systematic reviews of other study types the ROBIS checklist was used (see Appendix H in Developing NICE guidelines: the manual).

For non-randomised studies the ROBINS-I checklist was used (see Appendix H in Developing NICE guidelines: the manual).

Assessing inconsistency in intervention reviews

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

For the one instance of pooled data in the quantitative review, inconsistency was assessed visually by inspecting the forest plot and observing whether there was considerable heterogeneity in the results of the meta-analysis (for example if the point estimates of the individual studies consistently showed benefits or harms). This was supported by calculating the I-squared statistic for the meta-analysis with an I-squared value of more than 50% indicating serious heterogeneity, and more than 80% indicating very serious heterogeneity. In the case of the one meta-analysis performed for this guideline, no heterogeneity was detected so no exploration of heterogeneity and no subgroup analyses were required to identify possible explanations.

Assessing indirectness in intervention reviews

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

Assessing imprecision and importance in intervention reviews

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

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

Imprecision was assessed in the guideline evidence reviews by considering whether the width of the 95% CI of the effect estimate was relevant to decision making, considering each outcome independently. This is illustrated in Figure 1, which considers a positive outcome for the comparison of two treatments. Three decision-making zones can be differentiated, bounded by the thresholds for minimal importance (minimally important differences; MIDs) for benefit and harm.

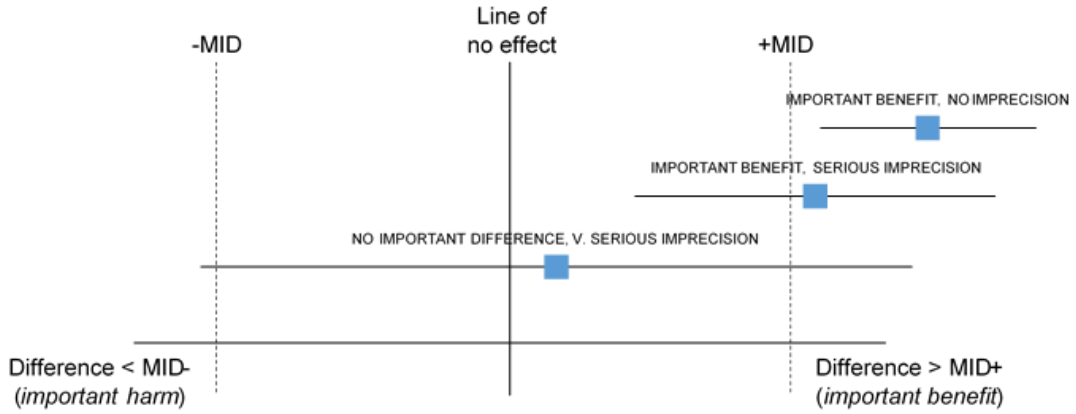
When the CI of the effect estimate is wholly contained in 1 of the 3 zones there is no uncertainty about the size and direction of effect, therefore, the effect estimate is considered precise; that is, there is no imprecision.

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

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

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

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



MID, minimally important difference

Defining minimally important differences for intervention reviews

The committee was not aware of any recognised or acceptable MIDs relevant to the guideline but in the case of the EQ-5D, SF-36 and the SF-12 the review team identified MIDs in the published literature and extrapolated these. For the EQ-5D, 3.7 (-3.7 and +3.7) was used based on a paper (McClure 2017) which reported that estimates of the EQ-5D-5L index score were generally between 0.037 and 0.069, apparently similar to the MID estimates of other preference based health related quality of life measures. The paper was not based on a homeless population so we

extrapolated a conservative estimate. For the SF-36, the MID was reported in one of the included papers (Tinland 2019) and this was also used as a basis for estimating a MID for the SF-12. This (the MID for SF-12) was set at 5 (-5 and +5) for both physical and mental components on the basis that in Tinland the MID for SF-36 (Physical Component Score) was 6.5 and for the Mental Component Score, 4.5.

In the absence of published or accepted MIDs, the committee agreed to use the GRADE default MIDs to assess imprecision. For dichotomous outcomes minimally important thresholds for a RR of 0.8 and 1.25 respectively were used as default MIDs in the guideline. The committee also chose to use 0.8 and 1.25 as the MIDs for ORs in the absence of published or accepted MIDs. ORs were predominantly used in the guideline when Peto OR were indicated due to low event rates, at low event rates OR are mathematically similar to RR making the extrapolation appropriate.

Where risk difference was used, for example because the study on which the data were based had zero events in both arms, imprecision was assessed based on sample size using 200 and 400 as cut-offs for very serious and serious imprecision respectively. The committee used these numbers based on commonly used optimal information size thresholds. The committee used the same approach to rating imprecision where medians were extracted.

The same thresholds were used as default MIDs in the guideline for all dichotomous outcomes considered in the intervention reviews. For continuous outcomes default MIDs are equal to half the median SD of the control groups at baseline (or at follow-up if the SD is not available a baseline).

Assessing publication bias in intervention reviews

The meta-analysis conducted in the quantitative review included fewer than 10 studies. The committee therefore subjectively assessed the likelihood of publication bias based on factors such as the proportion of trials funded by industry and the propensity for publication bias in the topic area.

Qualitative studies

GRADE-CERQual methodology for qualitative reviews

For the qualitative review an adapted GRADE Confidence in the Evidence from Reviews of Qualitative research (GRADE-CERQual) approach (Lewin 2015) was used. In this approach the quality of evidence is considered according to themes in the evidence. The themes may have been identified in the primary studies or they may have been identified by considering the reports of a number of studies. Quality elements assessed using GRADE-CERQual are listed and defined in Table 5. Each element was graded using the levels of concern summarised in Table 6.

The ratings for each component were combined (as with other types of evidence) to obtain an overall assessment of quality for each theme as described in Table 7. 'Confidence' in this context refers to the extent to which the review finding is a reasonable representation of the phenomenon of interest set out in the protocol. Similar to other types of evidence all review findings start off with 'high confidence' and are rated down by one or more levels if there are concerns about any of the individual CERQual components. In line with advice from the CERQual developers, the overall assessment does not involve numerical scoring for each component but in

order to ensure consistency across and between guidelines, the NGA has established some guiding principles for overall ratings. For example, a review finding would not be downgraded (and therefore would be assessed with 'high' confidence) if all 4 components had 'no or very minor' concerns or 3 'no or very minor' and 1 'minor'. At the other extreme, a review finding would be downgraded 3 times (to 'very low') if at least 2 components had serious concerns or at least 3 had moderate concerns. A basic principle was that if any components had serious concerns then overall confidence in the review finding would be downgraded at least once (potentially more depending on the other ratings). Transparency about overall judgements is provided in the CERQual tables, including a brief reference to components for which there were concerns in the 'overall confidence' cell.

Table 5: Adaptation of GRADE quality elements for qualitative reviews

Quality element	Description
Risk of bias ('Methodological limitations')	Limitations in study design and implementation may bias interpretation of qualitative themes identified. High risk of bias for the majority of the evidence reduces confidence in review findings. Qualitative studies are not usually randomised and therefore would not be downgraded for study design from the outset (they start as high quality)
Relevance (or applicability) of evidence	This refers to the extent to which the evidence supporting the review findings is applicable to the context specified in the review question
Coherence of findings	This refers to the extent to which review findings are well grounded in data from the contributing primary studies and provide a credible explanation for patterns identified in the evidence
Adequacy of data (theme saturation or sufficiency)	This corresponds to a similar concept in primary qualitative research, that is, whether a theoretical point of theme saturation was achieved, at which point no further citations or observations would provide more insight or suggest a different interpretation of the particular theme. Judgements are not based on the number of studies but do take account of the quantity and also richness of data underpinning a finding. The more complex the finding, the more detail the supporting data need to be. For simple findings, relatively superficial data would be considered adequate to explain and explore the phenomenon being described.

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

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

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

Overall confidence level	Definition
High	It is highly likely that the review finding is a reasonable representation of the phenomenon of interest
Moderate	It is likely that the review finding is a reasonable representation of the phenomenon of interest
Low	It is possible that the review finding is a reasonable representation of the phenomenon of interest
Very low	It is unclear whether the review finding is a reasonable representation of the phenomenon of interest

Assessing methodological limitations in qualitative reviews

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

Table 8: Methodological limitations in qualitative studies

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

Data analysis	This domain assesses whether sufficient detail was documented for the analytical process and whether it was in accordance with the theoretical approach. For example, if a thematic analysis was used, the assessment would focus on the description of the approach used to generate themes. Consideration of data saturation would also form part of this assessment (it could be reported directly or it might be inferred from the citations documented that more themes could be found)
Results	This domain assesses any reasoning accompanying reporting of results (for example, whether a theoretical proposal or framework is provided)

Assessing relevance of evidence in qualitative reviews

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

Assessing coherence of findings in qualitative reviews

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

Assessing adequacy of data in qualitative reviews

Adequacy of data (theme saturation or sufficiency) corresponds to a similar concept in primary qualitative research in which consideration is made of whether a theoretical point of theme saturation was achieved, meaning that no further citations or observations would provide more insight or suggest a different interpretation of the theme concerned. As noted above, it is not equivalent to the number of studies contributing to a theme, but it does take account of the quantity of data supporting a review finding (for instance whether sufficient quotations or observations were provided to underpin the findings) and in particular the degree of ‘richness’ of supporting data. Concerns about richness arise when insufficient details are provided by the data to enable an understanding of the phenomenon being described. Generally, if a review finding is fairly simple then relatively superficial data will be needed to understand it. Data underpinning a more complex finding would need to offer greater detail, allowing for interpretation and exploration of the phenomenon being described. Therefore in assessing adequacy our downgrading involved

weighing up the complexity of the review finding against the explanatory contribution of the supporting data.

Reviewing economic evidence

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

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

Inclusion criteria
Intervention or comparators in accordance with the guideline scope.
Study population in accordance with the guideline scope.
Full economic evaluations (cost-utility, cost effectiveness, cost-benefit or cost-consequence analyses) assessing both costs and outcomes associated with interventions of interest. Cost analyses were also considered for inclusion due to the anticipated lack of economic evidence.
Only studies published from 2010 onwards were included in the review, as older costings were considered to be out of date and less/not relevant to the current practice.
In areas with sufficient modelling or RCT-based economic evaluations for the committee decision making, economic evaluations with costs and effectiveness from observational study designs were not considered.
Exclusion criteria
Abstracts containing insufficient methodological details.
Cost-of-illness type studies.

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

Details of economic evidence study selection and lists of excluded studies across all reviews are presented in Supplement 2 (Economic literature). Economic evidence tables, the results of quality assessment of economic evidence (see below) and economic evidence profiles are presented in each of the evidence reports.

Appraising the quality of economic evidence

The quality of economic evidence was assessed using the economic evaluations checklist specified in Developing NICE guidelines: the manual. See the economic evidence profiles in the Summary of included economic evidence section of the Evidence review A & B for further details.

Economic modelling

The aims of the economic input to the guideline were to inform the guideline committee of potential economic issues to ensure that recommendations represented a cost effective use of resources. Economic evaluations aim to integrate data on benefits with the costs of different options. In addition, the economic input aimed to identify areas of high resource impact; these are recommendations which (while cost effective) might have a large impact on commissioners and so need special attention.

The guideline committee prioritised the economic analysis that would explore the cost-effectiveness of a strategy utilising lower caseloads for a practitioner within, for example, multidisciplinary outreach teams. This topic is crosscutting across the reviews and could be relevant to the following two review questions (out of the three):

- [A] What approaches are effective in improving access to and/or engagement with health and social care for people who experience homelessness? (Lower caseloads would mean that a practitioner has more time to engage with an individual, be persistent, invest in building trust, and facilitate access to care.)
- [B] What joined up approaches are effective in responding to the health, social care and housing needs of people experiencing homelessness? (Lower caseloads would mean that a practitioner has more time to engage with other services involved in the care of people experiencing homelessness which would facilitate joined up ways of working.)

The methods and results of the de novo economic analyses are reported in Appendix I of the combined evidence report for the above two questions (A & B). When new economic analysis was not prioritised, the committee made a qualitative judgement regarding cost effectiveness by considering expected differences in resource use and costs between options, alongside effectiveness evidence identified from the effectiveness evidence review. For example, the committee considered the potential cost of an approach to facilitate access and engagement and what it was relative to the value of improvements in health outcomes and also any changes in the cost drivers, such as unplanned attendances, crisis care, temporary accommodation or other types of resource use. The approach was likely to be cost-effective if its benefits were likely to outweigh the costs or an approach had very small costs relative to benefits and a very small budget impact. The committee also considered published public sector costs (Pleace 2016) as a reference for decision making. For example, they have considered how much public sector costs would need to be reduced or how much outcomes would need to improve to offset any additional costs associated with a particular approach or result in an incremental cost-effectiveness ratio that's below the NICEs cost-effectiveness threshold (see below).

Cost effectiveness criteria

In general, an intervention was considered to be cost effective if any of the following criteria applied (provided that the estimate was considered plausible):

- the intervention dominated other relevant strategies (that is, it was both less costly in terms of resource use and more effective compared with all the other relevant alternative strategies)
- the intervention cost less than £20,000 per quality-adjusted life year (QALY) gained compared with the next best strategy, however, it was acknowledged that this threshold may not be suitable for interventions that go beyond NHS and Personal Social Services (PSS) perspective
- the intervention provided important benefits at an acceptable additional cost when compared with the next best strategy.

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

Other sources of evidence

External experts (expert witness)

In addition to the systematic review evidence, testimony from expert witnesses was also used as a basis for recommendations, namely as a means of addressing gaps in the evidence reviews. The committee agreed to invite expert witnesses to address the paucity of evidence in the quantitative reviews about the role of social work and adult safeguarding in supporting access to and engagement with health and social care for people experiencing homelessness. The expert witnesses responded to a brief drafted by the technical team, which set out the key evidence gaps and the committee then used the testimony to make recommendations about safeguarding adults, aimed at commissioners, providers and Safeguarding Adults Boards.

Developing recommendations

Guideline recommendations

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

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

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

Research recommendations

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

Validation process

This guideline was subject to a 4-week public consultation and feedback process. All comments received from registered stakeholders were responded to in writing and posted on the NICE website at publication. For further details refer to Developing NICE guidelines: the manual.

Updating the guideline

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

Funding

The NGA was commissioned by NICE to develop this guideline.

References

Bradburn 2007

Bradburn, M. J., Deeks, J. J., Berlin, J. A., & Localio, A. R. Much ado about nothing: A comparison of the performance of meta-analytical methods with rare events. Statistics in Medicine, 26, 53–77, 2007.

Dixon-Woods 2005

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

Hayden 2013

Jill A. Hayden, Danielle A. van der Windt, Jennifer L. Cartwright, Pierre Côté, Claire Bombardier. Assessing Bias in Studies of Prognostic Factors. Ann Intern Med. 2013;158:280–286. doi: 10.7326/0003-4819-158-4-201302190-00009

Higgins 2011

Higgins JPT, Green S (editors) (2011) Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated 2019] The Cochrane Collaboration. Available from www.handbook.cochrane.org (accessed 20 July 2021)

Keenan 2021

Keenan, C, Miller, S, Hanratty, J, Piggot, T, Hamilton, J, Coughlan, C, Mackie, P, Fitzpatrick, S, Cowman, J (2021) Accommodation-based interventions for individuals experiencing, or at risk of experiencing, homelessness, Campbell Systematic Reviews, 17(2), e1165.

Lewin 2018

Lewin S, Booth A, Glenton C, Munthe-Kaas H et al. (2018) Applying GRADE-CERQual to qualitative evidence synthesis findings: introduction to the series. Implement Sci. 2018 Jan 25;13 (Suppl1):2

McClure 2017

McClure, N; Al Sayah, F; Xie, F; Luo, N; Johnson, J. Instrument-Defined Estimates of the Minimally Important Difference for EQ-5D-5L Index Scores. Value in Health 20 (4): 644-650

McGowan 2016

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

Miller 2019

Miller S, Keenan C, Hanratty J. (2019) Improving access to health and social services for individuals experiencing, or at risk of experiencing, homelessness: a systematic review. The Campbell Collaboration.

https://pure.qub.ac.uk/en/publications/improving-access-to-health-and-social-services-for-individuals-ex (accessed 11 August 2021)

NICE 2020

National Institute for Health and Care Excellence (NICE) (2020), Process and methods [PMG20], Developing NICE guidelines: the manual (first published 2014). Available from **https://www.nice.org.uk/process/pmg20/chapter/introduction** (accessed 02 September 2021)

NICE 2018

National Institute for Health and Care Excellence (NICE) (2014) NICE Policy on conflicts of interest (updated 2017). Available from https://www.nice.org.uk/Media/Default/About/Who-we-are/Policies-and-procedures/declaration-of-interests-policy.pdf (accessed 20 July 2021

Pleace 2016

Pleace, N., and Culhane, D., Better than Cure? : Testing the case for Enhancing Prevention of Single Homelessness in England. Research Report, Crisis, London, 2016

Santesso 2016

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

Tinland 2019

Tinland, A; Loubiere, S; Boucekine, M; Boyer, L; Fond, G; Girard, V; Auquier, P; Effectiveness of a Housing Support Team Intervention with a Recovery-Oriented Approach on Hospital and Emergency Department Use by Homeless People with Severe Mental Illness: A Randomized Controlled Trial; French Housing First Study, Effectiveness of a Housing Support Team Intervention with a Recovery-Oriented Approach on Hospital and Emergency Department Use by Homeless People with Severe Mental Illness: A Randomized Controlled Trial (July 11, 2019); 2019

White 2020a

White, H, Saran, A, Fowler, B, Portes, A, Fitzpatrick, S, Teixeira, L. PROTOCOL: Studies of the effectiveness of interventions to improve the welfare of those affected by, and at risk of, homelessness in high-income countries: An evidence and gap map. Campbell Systematic Reviews. 2020; 16:e1069.

White 2020b

White, H., Saran, A., Verma, A., & Verma, K. The effectiveness of interventions to improve the welfare of those experiencing and at risk of homelessness: An updated evidence and gap map. Global Evidence and Gap Map of Effectiveness, Third Edition. Centre for Homelessness Impact, Campbell Collaboration & Heriot Watt University. 2020.