National Institute for Health and Care Excellence

Final

End of life care for adults: service delivery

NICE guideline: methods

NICE guideline NG142

Methods

October 2019

Final

Developed by the National Guideline Centre, hosted by the Royal College of Physicians

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, where appropriate, their carer or guardian.

Local commissioners and 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 All NICE guidance is subject to regular review and may be updated or withdrawn.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 978-1-4731-3560-4

Contents

1 Introduction

End of life care is defined by NHS England as care that is provided in the 'last year of life'. After the Liverpool Care Pathway was withdrawn in 2014, a number of national reports and policy documents began to describe the changes needed for a new approach to end of life care services. They identified that high-quality, timely, compassionate and individualised care should be accessible to all those who need it. To progress this intention the models of care and the service delivery arrangements that need to be put in place for people as they approach the end of their life need to be defined.

End of life care may be delivered by disease-specific specialists and their associated teams; by generalists such as primary care teams or hospital-based generalists (for example, elderly care); or by palliative care specialists in hospices, hospitals and community settings. Care that is given alongside, and to enhance, disease-modifying and potentially life-prolonging therapies, often for years, is called 'supportive care'.

Giving this type of care can ensure that people live well until they die. Care that is aimed primarily at giving comfort and maintaining quality of life in the last months of life is commonly referred to as palliative care. Palliative care particularly aims to provide relief from pain and other distressing symptoms, integrate the psychological, social and spiritual aspects of the person's care, and continue to offer a support system to help people to live as actively as possible until their death.

This guideline describes end of life care services for providing palliative and supportive care to adults approaching the end of their life with any conditions and diseases. It advises on service models for care in acute settings by disease-specific specialists and their supportive services, or by primary care or specialists in palliative care in the community (for example, hospices).

2 Development of the guideline

2.1 What is a NICE guideline?

NICE guidelines are recommendations for the care of individuals in specific clinical conditions or circumstances within the NHS – from prevention and self-care through primary and secondary care to more specialised services. These may also include elements of social care or public health measures. We base our guidelines on the best available research evidence, with the aim of improving the quality of healthcare. We use predetermined and systematic methods to identify and evaluate the evidence relating to specific review questions.

NICE guidelines can:

- provide recommendations for the treatment and care of people by health professionals
- be used to develop standards to assess the clinical practice of individual health professionals
- be used in the education and training of health professionals
- help patients to make informed decisions
- improve communication between patient and health professional.

While guidelines assist the practice of healthcare professionals, they do not replace their knowledge and skills.

We produce our guidelines using the following steps:

- A guideline topic is referred to NICE from NHS England.
- Stakeholders register an interest in the guideline and are consulted throughout the development process.
- The scope is prepared by the National Guideline Centre (NGC).
- The NGC establishes a guideline committee.
- A draft guideline is produced after the group assesses the available evidence and makes recommendations.
- There is a consultation on the draft guideline.
- The final guideline is produced.

The guideline is made up of a collection of documents including this Methods report and a number of evidence reports covering each of the review questions included in the guideline. These can all be downloaded from NICE at www.nice.org.uk.

NICE also publishes a summary of the recommendation in this guideline, known as ‘the NICE guideline’.

NICE Pathways brings together all connected NICE guidance.

2.2 Remit

NICE received the remit for this guideline from NHS England. NICE commissioned the NGC to produce the guideline.

The remit for this guideline is:

- Service organisation that supports the identification of people thought to be entering the last year of life.

- Planning, coordinating and integrating the delivery of services, including sharing information between multidisciplinary teams.
- Service delivery models for end of life care, including both acute, community and third-sector settings, covering:
    - types of services (supportive and palliative care) provided by generalists and specialists during the course of the last year of life
    - who delivers the services and how
    - multidisciplinary team composition
    - timing and review of service provision
    - location of services, for example, place of care
    - out-of-hours, weekend and 24/7 availability of services.
- Service models that provide support for carers or those important to people accessing end of life services.
- Adaptations to adult palliative and end of life services for young adults thought to be entering the last year of life.

2.3 Who developed this guideline?

A multidisciplinary guideline committee comprising health professionals and researchers as well as lay members developed this guideline (see the list of guideline committee members and the acknowledgements).

The National Institute for Health and Care Excellence (NICE) funds the National Guideline Centre (NGC) and thus supported the development of this guideline. The committee was convened by the NGC and chaired by Mark Thomas in accordance with guidance from NICE.

The group met approximately every 6 weeks during the development of the guideline. At the start of the guideline development process all committee members declared interests including consultancies, fee-paid work, shareholdings, fellowships and support from the healthcare industry. At all subsequent committee meetings, members declared arising conflicts of interest.

Members were either required to withdraw completely or for part of the discussion if their declared interest made it appropriate. The details of declared interests and the actions taken are shown in the declaration of interest register for this guideline published on the NICE website.

Staff from the NGC provided methodological support and guidance for the development process. The team working on the guideline included a project manager, systematic reviewers (research fellows), health economists and information specialists. They undertook systematic searches of the literature, appraised the evidence, conducted meta-analysis and cost-effectiveness analysis where appropriate and drafted the guideline in collaboration with the committee.

2.3.1 What this guideline covers

Groups that will be covered:

- Adults (aged over 18 or over) with progressive life-limiting conditions thought to be entering the last year of life.
- Health and social care professionals delivering end of life care services to NHS patients.
- Carers of (or those important to) adults (aged over 18 or over) with progressive life-limiting conditions thought to be entering the last year of life. Includes young carers (<18 years).

Issues that will be covered:

- Identifying adults thought to be entering their last year of life
- Timing of referral
- Barriers to accessing end of life care
- The role and impact of a care coordinator/lead health professional
- Composition of a multiprofessional team
- Advance care planning
- Involving carers
- Carer support services
- Information sharing
- Review of service provision and identification of additional services
- Out of hours services
- Additional community services, available on a routine and emergency basis
- Optimal transition and facilitated discharge

For further details please refer to the scope for this guideline (published on the NICE website) and the review questions in section 3.1.

2.3.2 What this guideline does not cover

Groups that will not be covered:

- People under 18 years expected to be in their last year of life.
- People not expected to die within the next 12 months.

Issues that will not be covered:

- Clinical management at the end of life.

2.3.3 Relationships between the guideline and other NICE guidance

Related NICE guidelines:

- Dementia. NICE quality standard QS184 (2019)
- Emergency and acute medical care in over 16s: service delivery and organisation. NICE guideline NG94 (2018)
- People's experience in adult social care services: improving the experience of care and support for people using adult social care services. NICE guideline NG86 (2018)
- Decision-making and mental capacity. NICE guideline NG108 (2018)
- End of life care for adults. NICE quality standard QS13 (2017)
- Care of the dying adult. NICE quality standard QS144 (2017)
- Breast cancer. NICE quality standard QS12 (2016).
- Chronic obstructive pulmonary disease in adults. NICE quality standard QS10 (2016)
- Transition from children's to adults' services for young people using health or social care services. NICE guideline NG43 (2016)
- Motor neurone disease: assessment and management. NICE guideline NG42 (2016)
- Major trauma: assessment and initial management. NICE guideline NG39 (2016)
- Older people with social care needs and multiple long-term conditions. NICE quality standard QS132 (2016)

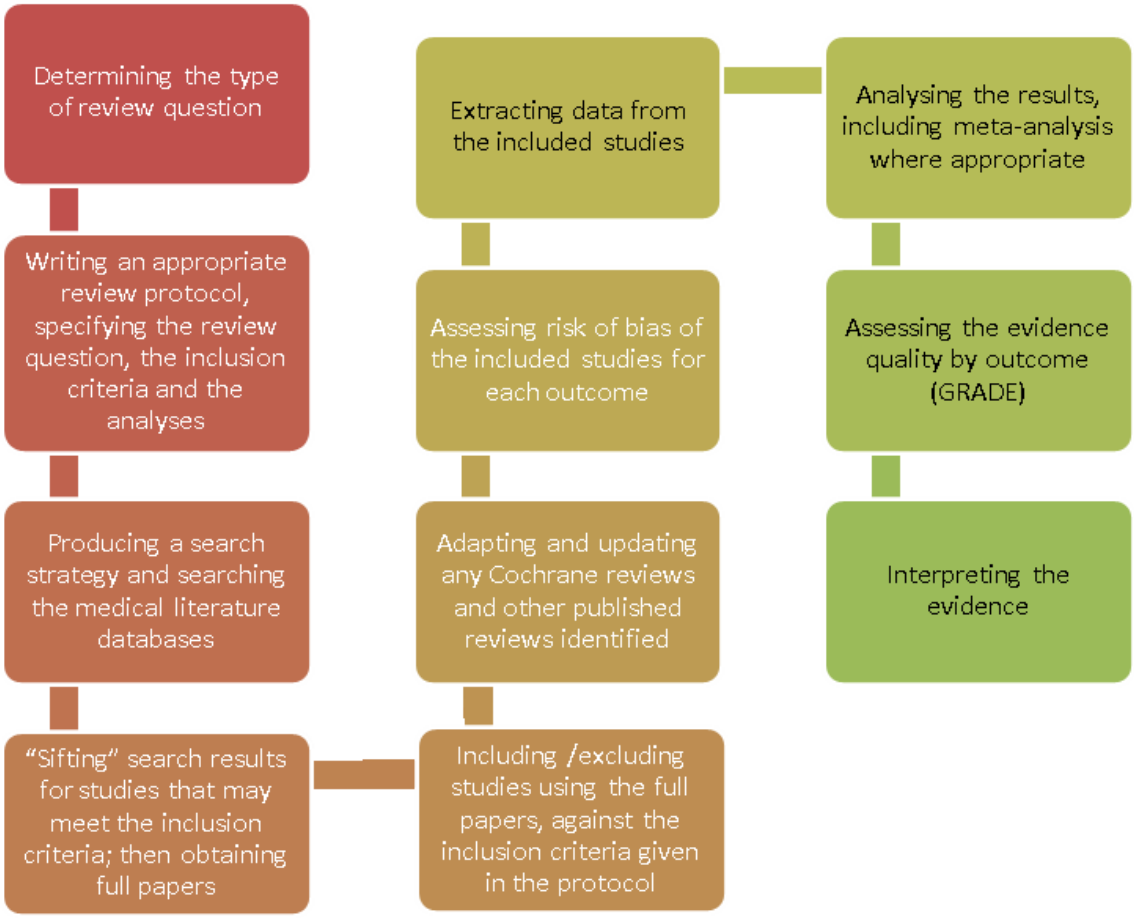
- Transition between inpatient mental health settings and community and care home settings for people with social care needs. NICE guideline NG53 (2016).
- End of life care for infants, children and young people. NICE guideline NG61 (2016)
- Transition between children's and adults' services. NICE quality standard QS140 (2016)
- Care of dying adults in the last days of life. NICE guideline NG31 (2015)
- Palliative care for adults: strong opioids for pain relief. NICE guideline CG140 (2012)
- Patient experience in adult NHS services. NICE guideline CG138 (2012)
- Service user experience in adult mental health. NICE guideline CG136 (2011)
- Medicines adherence. NICE guideline CG76 (2009)

3 Methods

This report sets out in detail the methods used to review the evidence and to develop the recommendations that are presented in each of the evidence reviews for this guideline. This guidance was developed in accordance with the methods outlined in the NICE guidelines manual, 2014 version[15] and the Interim methods guide for developing service guidance 2014 https://www.nice.org.uk/process/pmg8/chapter/introduction

Sections 3.1 to 3.3 describe the process used to identify and review clinical evidence (summarised in Figure 1), sections 3.2 and 3.4 describe the process used to identify and review the health economic evidence, and section 3.5 describes the process used to develop recommendations.

Figure 1: Step-by-step process of review of evidence in the guideline



3.1 Developing the review questions and outcomes

Review questions were developed using a PICO framework (population, intervention, comparison and outcome) for intervention reviews; using a framework of population, index tests, reference standard and target condition for reviews of diagnostic test accuracy; using population, presence or absence of factors under investigation (for example prognostic factors) and outcomes for prognostic reviews; and using a framework of population, setting and context for qualitative reviews.

This use of a framework guided the literature searching process, critical appraisal and synthesis of evidence, and facilitated the development of recommendations by the guideline committee. The review questions were drafted by the NGC technical team and refined and

validated by the committee. The questions were based on the key service delivery areas identified in the scope.

A total of 17 review questions were identified.

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

Table 1: Review questions

Chapter	Review questions	Outcomes
A	What are the best service models to support the identification of people who may be entering the last year of life?	Critical outcomes: • Quality of life • Preferred and actual place of death • Preferred and actual place of care • Longevity of carer Important outcomes: • Length of survival • Length of stay • Hospitalisation • Number of hospital visits • Number of visits to accident and emergency • Number of unscheduled admissions • Use of community services • Avoidable/inappropriate admissions to ICU • Inappropriate attempts at cardiopulmonary resuscitation • Staff satisfaction • Patient/carer reported outcomes (satisfaction) • Carer health (for example:GP visits, mental health, school/work attendance)
B	What is the best timing of referral to (or provision of) palliative care services in people thought to be entering their last year of life?	Critical outcomes: • Quality of life • Preferred and actual place of death • Preferred and actual place of care Important outcomes: • Length of survival • Length of stay • Hospitalisation • Number of hospital visits • Number of visits to accident and emergency • Number of unscheduled admissions • Use of community services • Avoidable/inappropriate admissions to ICU • Inappropriate attempts at cardiopulmonary resuscitation • Staff satisfaction

Chapter	Review questions	Outcomes
		• Patient/carer reported outcomes (satisfaction)
C	What are the barriers and facilitators the initial access to, and planning of end of life care services?	Initial access: • Any type of barriers and facilitators to the initial access of people in their last year of life to end of life care services. For example: • Communication around end of life issues such as for example, awareness of availability of end of life care services • Timing or setting of involvement in initial planning decision making (for example, ACP) • Facilitators (for example:coordinators, leaflets, information) Planning, choices, discharge: • What works well (and what doesn't) when facilitating discharge • Service features/elements that patients/carers considered as important for effective discharge process. • How and when to best incorporate patient's choice in the last year of life care pathway • Process for effective advance care planning • What process should be in place for allowing patients to change their minds/choices throughout their last year of life (after the initial advance care planning).
D	Is a lead health professional clinically and cost-effective to facilitate the continuity and coordination of care for people who are in their last year of life? Is a care facilitator/key worker/coordinator/case manager clinically and cost-effective to facilitate the continuity and coordination of care for people who are in their last year of life?	Critical outcomes: • Quality of life • Preferred and actual place of death • Preferred and actual place of care Important outcomes: • Length of survival • Length of stay • Hospitalisation • Number of hospital visits • Number of visits to accident and emergency • Number of unscheduled admissions • Use of community services • Avoidable/inappropriate admissions to ICU • Inappropriate attempts at cardiopulmonary resuscitation • Staff satisfaction • Patient/carer reported outcomes (satisfaction)

Chapter	Review questions	Outcomes
E	What is the best composition of a multidisciplinary team to facilitate the continuity and coordination of care for people who are in their last year of life?	Critical outcomes: • Quality of life • Preferred and actual place of death • Preferred and actual place of care Important outcomes: • Length of survival • Length of stay • Hospitalisation • Number of hospital visits • Number of visits to accident and emergency • Number of unscheduled admissions • Use of community services • Avoidable/inappropriate admissions to ICU • Inappropriate attempts at cardiopulmonary resuscitation • Staff satisfaction • Patient/carer reported outcomes (satisfaction)
F	What are the best service models to support advance care planning in people who may be entering the last year of life (including when it should be facilitated and by whom)?	Critical outcomes: • Quality of life • Preferred and actual place of death • Preferred and actual place of care Important outcomes: • Length of survival • Length of stay • Hospitalisation • Number of hospital visits • Number of visits to accident and emergency • Number of unscheduled admissions • Use of community services • Avoidable/inappropriate admissions to ICU • Inappropriate attempts at cardiopulmonary resuscitation • Staff satisfaction • Patient/carer reported outcomes (satisfaction)
G	What are the barriers and facilitators to the involvement of carers of (or people important to) those in their last year of life in planning and decision making?	Any type of barriers and facilitators to the involvement of carers (or people important to) people in their last year of life in planning and decision making described by studies (for example regarding discharge, transition in settings or advance care planning). For example: • Level of involvement • Timing or setting of involvement • Facilitators

Chapter	Review questions	Outcomes
		• Barriers • Financial and benefits support • Transportation geographical separation from patients/services
H	What are the most clinically and cost-effective support services for carers of (or those important to) people in their last year of life by health and social care professionals?	Critical outcomes: • Quality of life • Preferred and actual place of death • Preferred and actual place of care • Longevity of carer Important outcomes: • Carer health (for example:GP visits, mental health, school/work attendance) • Length of hospital stay • Use of community services • Staff (providing care to the person in their last year of life) satisfaction • Patient/carer reported outcomes (satisfaction)
I	What are the best ways to share information within multidisciplinary teams, between multidisciplinary teams and between multidisciplinary teams and services to ensure continuity of care for people who are in their last year of life?	Critical outcomes: • Quality of life • Preferred and actual place of death • Preferred and actual place of care Important outcomes: • Length of survival • Length of stay • Hospitalisation • Number of hospital visits • Number of visits to accident and emergency • Number of unscheduled admissions • Use of community services • Avoidable/inappropriate admissions to ICU • Inappropriate attempts at cardiopulmonary resuscitation • Staff satisfaction • Patient/carer reported outcomes (satisfaction)
J	When and how frequent should service need provision be reviewed in people thought to be entering their last year of life? What is the best method/service to review service provision and identify when additional services may be required in people thought to be entering their last year of life?	Critical outcomes: • Quality of life • Preferred and actual place of death • Preferred and actual place of care Important outcomes: • Length of survival • Length of stay • Hospitalisation • Number of hospital visits

Chapter	Review questions	Outcomes
		• Number of visits to accident and emergency • Number of unscheduled admissions • Use of community services • Avoidable/inappropriate admissions to ICU • Inappropriate attempts at cardiopulmonary resuscitation • Staff satisfaction • Patient/carer reported outcomes (satisfaction)
K	What are the best out of hours services, models and policies to support people in their last year of life to stay in their preferred place of care?	Critical outcomes: • Quality of life • Preferred and actual place of death • Preferred and actual place of care Important outcomes: • Length of survival • Length of stay • Hospitalisation • Number of hospital visits • Number of visits to accident and emergency • Number of unscheduled admissions • Use of community services • Avoidable/inappropriate admissions to ICU • Inappropriate attempts at cardiopulmonary resuscitation • Staff satisfaction • Patient/carer reported outcomes (satisfaction)
L	What additional community services are needed to support people in their last year of life to stay in their preferred place of care? What provision of additional community services should be available to reduce inappropriate/avoidable admissions in people in their last year of life?	Critical outcomes: • Quality of life • Preferred and actual place of death • Preferred and actual place of care Important outcomes: • Length of survival • Length of stay • Hospitalisation • Number of hospital visits • Number of visits to accident and emergency • Number of unscheduled admissions • Use of community services • Avoidable/inappropriate admissions to ICU • Inappropriate attempts at cardiopulmonary resuscitation • Staff satisfaction

Chapter	Review questions	Outcomes
		• Patient/carer reported outcomes (satisfaction)
M	What service models (or service components) enable an optimal transition between care settings in people in their last year of life? What is the best way to facilitate discharge of a person in their last year of life back to the community from another setting (for example, the hospital)?	Critical outcomes: • Quality of life • Preferred and actual place of death • Preferred and actual place of care Important outcomes: • Length of survival • Length of stay • Hospitalisation • Number of hospital visits • Number of visits to accident and emergency • Number of unscheduled admissions • Use of community services • Avoidable/inappropriate admissions to ICU • Inappropriate attempts at cardiopulmonary resuscitation • Staff satisfaction • Patient/carer reported outcomes (satisfaction)

3.2 Searching for evidence

Clinical and health economics literature searches

The full search 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 report.

Systematic literature searches were undertaken to identify all published clinical and health economics evidence relevant to the review questions. Searches were undertaken according to the parameters stipulated within the NICE guidelines manual 2014 (see https://www.nice.org.uk/process/pmg20/). Databases were searched using relevant medical subject headings, free-text terms and study-type filters where appropriate. Studies published in languages other than English were not reviewed, where possible, searches were restricted to English Language. All searches were updated on January 4^{h} 2019. Papers published or added to databases after this date were not considered. If new evidence falls outside of the timeframe for the guideline searches e.g. from stakeholder comments, the impact on the guideline will be considered, and any further action agreed between the developer and NICE staff with a quality assurance role.

Prior to running, searches were quality assured using different approaches. Medline search strategies were checked by a second information specialist before being run. Searches were cross-checked with reference lists of highly relevant papers, searches in other systematic reviews analysed, and committee members requested to highlight additional studies.

During the scoping stage, a search was conducted for guidelines and reports on the websites listed below. Web sites searched include:

- Guidelines International Network database (www.g-i-n.net)
- National Guideline Clearing House (www.guideline.gov)

- NHS Evidence Search (www.evidence.nhs.uk).
- TRIP database (www.tripdatabase.com)

Searching for unpublished literature was not undertaken.

Call for evidence

This is initiated when the developer or committee believes there is relevant evidence in addition to that identified by the searches in some topic areas or for some review questions. This process is outlined in section 5.5 of Developing NICE guidelines: the manual [https://www.nice.org.uk/process/pmg20/]. In this guideline there was a call for evidence on the following topics:

- identify people who may be entering the last year of their life
- address the clinical and cost effectiveness of out-of-hours, weekend and 24/7 availability of services
- support people to stay in their preferred place of care (for example out of hours services)
- facilitate smooth transitions between care settings (for example discharge planning teams)
- facilitate continuity and coordination of care (for example multidisciplinary team working)
- reduce inappropriate/avoidable hospital admissions (for example community health services and telehealth)
- facilitate discharge back to the community from other settings (for example rapid discharge pathways)

Information was requested on service delivery models that reported measurable outcomes, for example the number of people who die in their preferred place of death, quality of life, and the use of hospital and community services (including staff time or any other information on resource use), and the costs associated with providing or implementing the service delivery model providing these.

239 items of research were received and reviewed, from which 21 items were ordered in full for review, 2 studies were included in the Barriers to accessing end of life care services [10, 13]. Table 2 lists the reasons for exclusion for the other 19 studies .

Table 2: Studies excluded from the call for evidence

Reference	Relevant review	Reason for exclusion
Aoun 2015 [1]	Barriers to accessing end of life care services	No relevant outcome
Aoun 2015 [2]	Barriers to accessing end of life care services	No relevant outcomes
Bajwah 2015 [3]	Advanced care planning	No relevant outcomes
Candy 2011[4]	Multiprofessional team	Not relevant to PICO , Inappropriate study design
Chapman 2016 [5]	Economic modelling for out of hours services	No relevant data
Gomes 2014[6]	Multiprofessional team	Inappropriate study design . Incorrect interventions
Grande 2017 [8]	Carer support services	Inappropriate study design
Harding 2012 [9]	Barriers to accessing end of life care services	Not review population

Reference	Relevant review	Reason for exclusion
Lamont, 2016 [11]	Multiprofessional team	Inappropriate study design. Incorrect interventions
Lucas 2008 [12]	Barriers to accessing end of life care services. Carers perspective	No relevant outcomes
Nakajima, 2015 [14]	Advanced care planning	No relevant outcomes
Perkins 2016 [19]	Advanced care planning	No relevant outcomes
Petrova, 2016 [20]	Information sharing	Inappropriate study design
Sandsdalen 2016 [22]	Barriers to accessing end of life care services	Inappropriate study design
Smith, 2012 [23]	Identifying adults	Inappropriate study design
Wilkinson, 2016 [24]	Barriers to accessing end of life care services	No relevant themes
Wilkinson, 2014 [25]	Barriers to accessing end of life care services	No relevant themes
Wye 2014 [26]	Barriers to accessing end of life care services	Inappropriate study design

3.3 Identifying and analysing evidence of effectiveness

Research fellows conducted the tasks listed below, which are described in further detail in the rest of this section:

- Identified potentially relevant studies for each review question from the relevant search results by reviewing titles and abstracts. Full papers were then obtained.
- Reviewed full papers against pre-specified inclusion and exclusion criteria to identify studies that addressed the review question in the appropriate population, and reported on outcomes of interest (review protocols are included in an appendix to each of the evidence reports).
- Critically appraised relevant studies using the appropriate study design checklist as specified in the NICE guidelines manual.[15] Prognostic studies were critically appraised using NGC checklists. Qualitative studies were critically appraised using the GRADE CERQual approach for rating confidence in the body of evidence as a whole and using an NGC checklist for the methodological limitations section of the quality assessment.
- Extracted key information about interventional study methods and results using 'Evibase', NGC's purpose-built software. Evibase produces summary evidence tables, including critical appraisal ratings. Key information about non-interventional study methods and results was manually extracted onto standard evidence tables and critically appraised separately (evidence tables are included in an appendix to each of the evidence reports).
- Generated summaries of the evidence by outcome. Outcome data were combined, analysed and reported according to study design:
  - Randomised data were meta-analysed where appropriate and reported in GRADE profile tables.
  - Data from non-randomised studies were presented as a range of values in GRADE profile tables or meta-analysed if appropriate.

  - Qualitative data were synthesised across studies and presented as summary statements with accompanying GRADE CERQual ratings for each review finding.
- A sample of a minimum of 10% of the abstract lists of the first 3 sifts by new reviewers and those for complex review questions (for example, prognostic reviews) were double-sifted by a senior research fellow and any discrepancies were rectified. All of the evidence reviews were quality assured by a senior research fellow. This included checking:
  - papers were included or excluded appropriately
  - a sample of the data extractions
  - correct methods were used to synthesise data
  - a sample of the risk of bias assessments.

3.3.1 Inclusion and exclusion criteria

The inclusion and exclusion of studies was based on the criteria defined in the review protocols, which can be found in an appendix to each of the evidence reports. Excluded studies (with the reasons for their exclusion) are listed in another appendix to each of the evidence reports. The committee was consulted about any uncertainty regarding inclusion or exclusion.

The key population inclusion criterion was:

- Adults (aged over 18 or over) with progressive life-limiting conditions thought to be entering the last year of life.
- Carers of (or those important to) adults (aged over 18 or over) with progressive life-limiting conditions thought to be entering the last year of life. Includes young carers (<18 years).

The key population exclusion criterion was:

- People under 18 years expected to be in their last year of life.
- People not expected to die within the next 12 months.

3.3.1.1 Saturation of qualitative studies

Data extraction in qualitative reviews is a thorough process and may require more time compared to intervention reviews. It is common practice to stop extracting data once saturation has been reached. This is the point when no new information emerges from studies that match the review protocol. The remaining identified studies are, however, not directly excluded from the review as they nevertheless fit the criteria defined in the review protocol. Data saturation was not reached for the qualitative reviews in this guideline.

3.3.2 Type of studies

Randomised trials, non-randomised intervention studies, and other observational studies were included in the evidence reviews as appropriate.

For intervention reviews in this guideline, parallel randomised controlled trials (RCTs) were prioritised for inclusion because they are considered the most robust type of study design that can produce an unbiased estimate of the intervention effects. Non-randomised intervention studies were considered appropriate for inclusion in cases where little or no randomised evidence was available for critical outcomes. Please refer to the review protocols in each evidence report for full details on the study design of studies selected for each review question.

Where sufficient evidence was found from RCTs, non-randomised studies were not considered for inclusion. Where data from non-randomised studies were included, the results for each outcome were presented separately for each study.

3.3.3 Methods of combining clinical studies

3.3.3.1 Data synthesis for intervention reviews

Where possible, meta-analyses were conducted using Cochrane Review Manager (RevMan5)[21]software to combine the data given in all studies for each of the outcomes of interest for the review question. As the interventions were so complex and different from each other, it was deemed inappropriate to combine the studies in a meta-analysis. However where appropriate we have lumped similar interventions together when looking at results, depending on the intensity of the intervention.

3.3.3.1.1 Analysis of different types of data

Dichotomous outcomes

Fixed-effects (Mantel–Haenszel) techniques (using an inverse variance method for pooling) were used to calculate risk ratios (relative risk, RR) for the binary outcomes, which included:

- preferred and actual place of death
- preferred and actual place of care
- hospital admissions
- use of community services
- Inappropriate attempts at cardiopulmonary resuscitation

The absolute risk difference was also calculated using GRADEpro[7] 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. Peto odds ratios are more appropriate for data with a low number of events.

Continuous outcomes

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

- heath-related quality of life (HRQoL)
- length of survival
- length of stay in hospital
- satisfaction

Where the studies within a single meta-analysis had different scales of measurement, standardised mean differences were used (providing all studies reported either change from baseline or final values rather than a mixture of both); each different measure in each study was ‘normalised’ to the standard deviation value pooled between the intervention and comparator groups in that same study.

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 Cochrane Review Manager (RevMan5)[21] software. Where p values were reported as ‘less than’, a conservative approach was undertaken. For example, if a p value

was reported as 'p≤0.001', the calculations for standard deviations were based on a p value of 0.001. If these statistical measures were not available then the methods described in section 16.1.3 of the Cochrane Handbook (version 5.1.0, updated March 2011) were applied.

3.3.3.1.2 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.[21]If the control event rate was reported this was used to generate the absolute risk difference in GRADEpro.[7] If multivariate analysis was used to derive the summary statistic but no adjusted control event rate was reported no absolute risk difference was calculated.

3.3.3.1.3 Heterogeneity

Statistical heterogeneity was assessed for each meta-analysis estimate by considering the chi-squared test for significance at $p<0.1$ or an I-squared (I^2) inconsistency statistic (with an I-squared value of more than 50% indicating significant heterogeneity) as well as the distribution of effects. Where significant heterogeneity was present, predefined subgrouping of studies was carried out for:

- Younger adults (aged 18-25)
- Frail elderly
- People with dementia
- People with hearing loss
- People with advanced heart and lung disease
- People in prisons
- Socioeconomic inequalities (people from lower income brackets)
- Homeless people/vulnerably housed
- Travelers
- People with learning difficulties
- People with disabilities
- People with mental health problems
- Migrant workers
- LGBT
- People in whom life-prolonging therapies are still an active option

If the subgroup analysis resolved heterogeneity within all of the derived subgroups, then each of the derived subgroups were adopted as separate outcomes (providing at least 1 study remained in each subgroup. For example, instead of the single outcome of '*quality of life'*, this was separated into 2 outcomes '*quality of life in people aged under 25'* and '*quality of life in people aged 25 and over'*. Assessments of potential differences in effect between subgroups were based on the chi-squared tests for heterogeneity statistics between subgroups. Any subgroup differences were interpreted with caution as separating the groups breaks the study randomisation and as such is subject to uncontrolled confounding.

For some questions additional subgrouping was applied, and this is documented in the individual review question protocols. These additional subgrouping strategies were applied independently, so subunits of subgroups were not created, unlike the situation with strata. Other subgrouping strategies were only used if the age category subgroup was unable to explain heterogeneity, then these further subgrouping strategies were applied in order of priority. Again, once a subgrouping strategy was found to explain heterogeneity from all derived subgroups, further subgrouping strategies were not used.

If all predefined strategies of subgrouping were unable to explain statistical heterogeneity within each derived subgroup, 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, thus providing a more realistic interpretation of the true distribution of effects across more than 1 population. If, however, the committee considered the heterogeneity was so large that meta-analysis was inappropriate, then the results were described narratively.

3.3.3.2 Data synthesis for qualitative study reviews

The main findings for each included paper were identified and thematic analysis methods were used to synthesise this information into broad overarching themes which were summarised into the main review findings. The evidence was presented in the form of a narrative summary detailing the evidence from the relevant papers and how this informed the overall review finding plus a statement on the level of confidence for that review finding. Considerable limitations and issues around relevance were listed. A summary evidence table with the succinct summary statements for each review finding was produced including the associated quality assessment.

3.3.4 Appraising the quality of evidence by outcomes

3.3.4.1 Intervention reviews

The evidence for outcomes from the included RCTs and, where appropriate, non-randomised intervention studies, were evaluated and presented using an adaptation of the 'Grading of Recommendations Assessment, Development and Evaluation (GRADE) toolbox' developed by the international GRADE working group (http://www.gradeworkinggroup.org/). The software (GRADEpro[7]) 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 2.

Table 2: 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.

Quality element	Description
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 or other bias was only taken into consideration in the quality assessment if it was apparent.

3.3.4.1.1 Risk of bias

The main domains of bias for RCTs are listed in Table 3. Each outcome had its risk of bias assessed within each study first. For each study, if there were no risks of bias in any domain, the risk of bias was given a rating of 0. If there was risk of bias in just 1 domain, the risk of bias was given a 'serious' rating of −1, but if there was risk of bias in 2 or more domains the risk of bias was given a 'very serious' rating of −2. A weighted average score was then calculated across all studies contributing to the outcome, by taking into account the weighting of studies according to study precision. For example if the most precise studies tended to each have a score of −1 for that outcome, the overall score for that outcome would tend towards −1.

Table 3: Principle domains of bias in randomised controlled trials

Limitation	Explanation
Selection bias (sequence generation and allocation concealment)	If those enrolling patients are aware of the group to which the next enrolled patient will be allocated, either because of a non-random sequence that is 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 of patients and healthcare professionals)	Patients, caregivers, those adjudicating or recording outcomes, and data analysts should not be aware of the arm to which patients 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 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:

Limitation	Explanation
	• 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.

The assessment of risk of bias differs for non-randomised intervention studies, as they are inherently at high risk of selection bias. For this reason, GRADE requires that non-randomised evidence is initially downgraded on the basis of study design, starting with a rating of −2. This accounts for selection bias and so non-randomised intervention studies are not downgraded any further on that domain. Non-randomised evidence was assessed against the remaining domains used for RCTs in Table 3, and downgraded further as appropriate.

3.3.4.1.2 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 0. If there was indirectness in just 1 source (for example in terms of population), indirectness was given a 'serious' rating of −1, but if there was indirectness in 2 or more sources (for example, in terms of population and treatment) the indirectness was given a 'very serious' rating of −2. A weighted average score was then calculated across all studies contributing to the outcome by taking into account study precision. For example, if the most precise studies tended to have an indirectness score of −1 each for that outcome, the overall score for that outcome would tend towards −1.

3.3.4.1.3 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. When heterogeneity existed within an outcome (chi-squared $p<0.1$, or $I^2>50\%$), but no plausible explanation could be found, the quality of evidence for that outcome was downgraded. Inconsistency for that outcome was given a 'serious' score of −1 if the I^2 was 50–74%, and a 'very serious' score of −2 if the I^2 was 75% or more.

If inconsistency could be explained based on prespecified subgroup analysis (that is, each subgroup had an $I^2<50\%$), the committee took this into account and considered whether to make separate recommendations on new outcomes based on the subgroups defined by the assumed explanatory factors. In such a situation the quality of evidence was not downgraded for those emergent outcomes.

Since the inconsistency score was based on the meta-analysis results, the score represented the whole outcome and so weighted averaging across studies was not necessary.

3.3.4.1.4 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 and a 'serious' score of −1 was given. 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 and a 'very serious' score of −2 was given. 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 2. As for inconsistency, since the imprecision score was based on the meta-analysis results, the score represented the whole outcome and so weighted averaging across studies was not necessary.

The 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. For binary variables, any MIDs reported in the literature will inevitably be based on expert consensus, as such MIDs relate to all-or-nothing population effects rather than measurable effects on an individual, and so are not amenable to patient-centred 'anchor' methods.

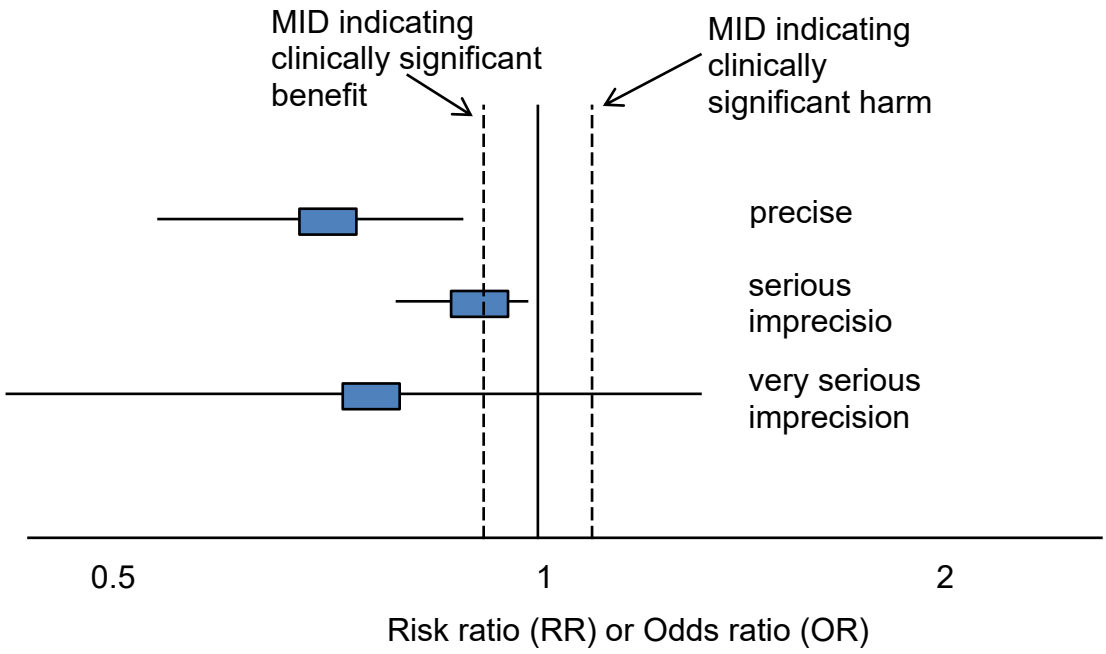
In the absence of values identified in the literature, the alternative approach to deciding on MID levels is the 'default' method, as follows:

- For categorical outcomes the MIDs were taken to be RRs, ORs, Peto Odds Ratio (POR) of 0.75 and 1.25. For 'positive' outcomes such as 'patient satisfaction', the RR, OR or POR of 0.75 is taken as the line denoting the boundary between no clinically important effect and a clinically significant harm, whilst the RR, OR or POR of 1.25 is taken as the line denoting the boundary between no clinically important effect and a clinically significant benefit. For 'negative' outcomes such as 'bleeding', the opposite occurs, so the RR, OR or POR of 0.75 is taken as the line denoting the boundary between no clinically important effect and a clinically significant benefit, whilst the RR, OR or POR of 1.25 is taken as the line denoting the boundary between no clinically important effect and a clinically significant harm.
- For mortality any change was considered to be clinically important and the imprecision was assessed on the basis of the whether the confidence intervals crossed the line of no effect, that is whether the result was consistent with both benefit and harm.
- 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 significant 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 significant harms will be the converse of these. If baseline values are unavailable, then half the median comparator group standard deviation of that variable will be taken as the MID.
- If standardised mean differences have been used, then the MID will be set at the absolute value of $\pm$0.5. This follows because standardised mean differences are mean differences normalised to the pooled standard deviation of the 2 groups, and are thus effectively expressed in units of 'numbers of standard deviations'. The 0.5 MID value in this context therefore indicates half a standard deviation, the same definition of MID as used for non-standardised mean differences.

The default MID value was subject to amendment after discussion with the committee. If the committee decided that the MID level should be altered, after consideration of absolute as well as relative effects, this was allowed, provided that any such decision was not influenced by any bias towards making stronger or weaker recommendations for specific outcomes.

For this guideline, no appropriate MIDs for continuous or dichotomous outcomes were found in the literature and the committee did not decide to alter the MID in any of the reviews, and so the default method was adopted throughout the reviews..

Figure 2: 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)



3.3.4.1.5 Overall grading of the quality of clinical evidence

Once an outcome had been appraised for the main quality elements, as above, an overall quality grade was calculated for that outcome. The scores (0, −1 or −2) from each of the main quality elements were summed to give a score that could be anything from 0 (the best possible) to −8 (the worst possible). However scores were capped at −3. This final score was then applied to the starting grade that had originally been applied to the outcome by default, based on study design. All RCTs started as High and the overall quality became Moderate, Low or Very Low if the overall score was −1, −2 or −3 points respectively. The significance of these overall ratings is explained in Table 4. The reasons for downgrading in each case were specified in the footnotes of the GRADE tables.

Non-randomised intervention studies started at Low, and so a score of −1 would be enough to take the grade to the lowest level of Very Low. Non-randomised intervention studies could, however, be upgraded if there was a large magnitude of effect or a dose-response gradient.

Table 4: 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

3.3.4.2 Qualitative reviews

Review findings from the included qualitative studies were evaluated and presented using the 'Confidence in the Evidence from Reviews of Qualitative Research' (CERQual) Approach developed by the GRADE-CERQual Project Group, a subgroup of the GRADE Working Group.

The CERQual Approach assesses the extent to which a review finding is a reasonable representation of the phenomenon of interest (the focus of the review question). Each review finding was assessed for each of the 4 quality elements listed and defined below in Table 5.

Table 5: Description of quality elements in GRADE-CERQual for qualitative studies

Quality element	Description
Methodological limitations	The extent of problems in the design or conduct of the included studies that could decrease the confidence that the review finding is a reasonable representation of the phenomenon of interest. Assessed at the study level using an NGC checklist.
Coherence	The extent to which the reviewer is able to identify a clear pattern across the studies included in the review.
Relevance	The extent to which the body of evidence from the included studies is applicable to the context (study population, phenomenon of interest, setting) specified in the protocol.
Adequacy	The degree of the confidence that the review finding is being supported by sufficient data. This is an overall determination of the richness (depth of analysis) and quantity of the evidence supporting a review finding or theme.

Details of how the 4 quality elements (methodological limitations, coherence, relevance and adequacy) were appraised for each review finding are given below.

3.3.4.2.1 Methodological limitations

Each review finding had its methodological limitations assessed within each study first using an NGC checklist. Based on the degree of methodological limitations studies were evaluated as having minor, moderate or severe limitations. The questions to be answered in the checklist below included:

- Was qualitative design an appropriate approach?
- Was the study approved by an ethics committee?
- Was the study clear in what it sought to do?
- Is the context clearly described?
- Is the role of the researcher clearly described?
- Are the research design and methods rigorous?
- Was the data collection rigorous?
- Was the data analysis rigorous?

- Are the data rich?
- Are the findings relevant to the aims of the study?
- Are the findings and conclusions convincing?

The overall assessment of the methodological limitations of the evidence was based on the primary studies contributing to the review finding. The relative contribution of each study to the overall review finding and of the type of methodological limitation(s) were taken into account when giving an overall rating.

3.3.4.2.2 Coherence

Coherence is the extent to which the reviewer is able to identify a clear pattern across the studies included in the review, and if there is variation present (contrasting or disconfirming data) whether this variation is explained by the contributing study authors. If a review finding in 1 study does not support the main finding and there is no plausible explanation for this variation, then the confidence that the main finding reasonably reflects the phenomenon of interest is decreased. Each review finding was given a rating of minor, moderate or major concerns about coherence.

3.3.4.2.3 Relevance

Relevance is the extent to which the body of evidence from the included studies is applicable to the context (study population, phenomenon of interest, setting) specified in the protocol. As such, relevance is dependent on the individual review and discussed with the guideline committee. Relevance is categorised in 3 ways: partial relevance, indirect relevance and no concerns about relevance.

3.3.4.2.4 Adequacy

The judgement of adequacy is based on the confidence of the finding being supported by sufficient data. This is an overall determination of the richness (depth of analysis) and quantity of the evidence supporting a review finding or theme. Rich data provide sufficient detail to gain an understanding of the theme or review finding, whereas thin data do not provide enough detail for an adequate understanding. Quantity of data is the second pillar of the assessment of adequacy. For review findings that are only supported by 1 study or data from only a small number of participants, the confidence that the review finding reasonable represents the phenomenon of interest might be decreased. As with richness of data, quantity of data is review dependent. Based on the overall judgement of adequacy, a rating of no concerns, minor concerns, or substantial concerns about adequacy was given.

3.3.4.2.5 Overall judgement of the level of confidence for a review finding

GRADE-CERQual is used to assess the body of evidence as a whole through a confidence rating representing the extent to which a review finding is a reasonable representation of the phenomenon of interest. The 4 components (methodological limitations, coherence, relevance and adequacy) are used in combination to form an overall judgement. GRADE-CERQual uses 4 levels of confidence: high, moderate, low and very low confidence. The significance of these overall ratings is explained in Table 6. Each review finding starts at a high level of confidence and is downgraded based on the concerns identified in any 1 or more of the 4 components. Quality assessment of qualitative reviews is a subjective judgement by the reviewer based on the concerns that have been noted. A detailed explanation of how such a judgement had been made was included in the narrative summary.

Table 6: Overall level of confidence for a review finding in GRADE-CERQual

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

3.3.5 Assessing clinical importance

The committee assessed the evidence by outcome in order to determine if there was, or potentially was, a clinically important benefit, a clinically important harm or no clinically important difference between interventions. To facilitate this, binary outcomes were converted into absolute risk differences (ARDs) using GRADEpro[7] software: the median control group risk across studies was used to calculate the ARD and its 95% CI from the pooled risk ratio.

The assessment of clinical benefit, harm, or no benefit or harm was based on the point estimate of absolute effect for intervention studies, which was standardised across the reviews.

This assessment was carried out by the committee for each critical outcome, and an evidence summary table was produced to compile the committee's assessments of clinical importance per outcome, alongside the evidence quality and the uncertainty in the effect estimate (imprecision).

Some of the outcomes were difficult in distinguishing whether they represented a clinical benefit or harm as they were people's preferences which may not be reported or known. For the outcome 'preferred and actual place of death', many studies did not provide any details on the preferred place of death, therefore the Committee agreed to include the surrogate outcome of 'final place of death' and this was downgraded for indirectness.

The Committee discussed the direction of effect (benefit/harm) for this outcome. They acknowledged that people change their views about their preferred place of death, depending on the circumstances in which the question is asked (for example, depending on the stage of the disease or the availability of carers). However, the Committee considered it fair to assume that people would prefer to die in their usual place of care when they are confident they would be able to have a 'good death', the Committee agreed to consider dying at home as a clinically important benefit for people in their last year of life throughout the guideline. This would also apply to outcomes relating to deaths in the community (e.g. hospice, nursing homes), while dying in hospital would be considered as a clinically important harm. These assumptions were revised by the Committee every time they were presented evidence for these outcomes, and the evidence was interpreted accordingly.

The Committee felt that for outcomes that assessed the impact of an intervention on quality of life, the effectiveness may not necessarily be demonstrated by an improvement but in maintaining someone's (or their carer's) quality of life, The Committee also acknowledged the challenge of measuring satisfaction in end of life care, as people are often either very dissatisfied or very satisfied, which might polarise results. The measurement of satisfaction is further complicated by the fact that often papers do not report whether the care people received matched their expectations.

See tables 7 and 8 for more detail on the importance and interpretation of outcome measures.

Table 7: Critical outcome measures

Outcome	Why is important in EOLC? Why did we choose this particular outcome?	What is the interpretation of the outcome? For example, is an increase in the scale/n of events of the outcome a benefit or harm?	
		From a clinical point of view	From a service delivery point of view (i.e. costs, logistics)
Quality of life of person in their last year of life (Continuous)	This outcome was identified as critical because towards the end of life, people might place more importance on quality of life rather than 'quantity' (i.e. length of survival).	The Committee acknowledged the difficulties in measuring quality of life at the end of life. • For example, many of the quality of life scales used by studies are designed for RCTs and might not be relevant in the last year of life. Quality of life is often used as an 'umbrella' term, and studies might not use standardised tools. • Furthermore, different studies might measure different indicators of quality of life, therefore the Committee acknowledged that is safer to talk of aspects of quality of life (for example, the fatigue or pain experienced by patients could have an impact on quality of life). • An additional issue in the measurement of quality of life is that often it is measured at discrete time points. Quality of life perception might change with time and sometimes this is not reported in papers. Overall, the Committee stressed that regardless of the specific item or aspect measured, quality of life measures should always preferably be patient-rated.	The Committee commented that when quality of life is better, people are less likely to access services as they are more likely to self-manage. For the purpose of evaluating the evidence presented in the guideline, the Committee generally assumed that a slower deterioration in quality of life is a service delivery benefit. However, the Committee was aware that a better quality of life can often be associated to a higher increase in use of services. In fact, services delivered at the right place and time can lead to better quality of life. The Committee therefore decided not to generalise on the direction of effect for this outcome and to consider case by case in the context of the single studies.

Outcome	Why is important in EOLC? Why did we choose this particular outcome?	What is the interpretation of the outcome? For example, is an increase in the scale/n of events of the outcome a benefit or harm?	
		From a clinical point of view	From a service delivery point of view (i.e. costs, logistics)
		The Committee noted that it is difficult to observe an increase in quality of life at the end of life. The direction of change most likely observed is deterioration in quality of life. Observing a slower deterioration or a smaller degree of deterioration in quality of life was considered to be a clinically important benefit by the Committee.	
Quality of life of carer of (or person important to) the person in their last year of life (Continuous)	*As above*	The Committee agreed that an increase in quality of life of carer or people important to the person in their last year of life is a clinically important benefit. The Committee was aware that carers and patients' views may be in conflict and therefore an increase in quality of life for the carer might not be reflected in a better quality of life for the patient.	The Committee commented that when carers are able to cope with the situation, this is likely to decrease patients' admissions. The Committee therefore agreed that an improvement in quality of life of carers can be considered a service delivery benefit.
Preferred and actual place of death (Dichotomous)	The Committee recognised that a critical issue for patients is where they spend their time leading to death. The Committee acknowledged that the preferred and actual place of death outcome was a blunt measure to measure where people spend their time before death, as there is an element of unpredictability in death, therefore people might die in a place they did not prefer, and make choices different from their preferences.	The Committee found it difficult to establish whether dying at the preferred place of death would represent a clinically important benefit for patients. The Committee acknowledged that people change their views about their preferred place of death, depending on the circumstances in which the question is asked (for example: depending on the stage of the disease, the availability of carers). Furthermore, as there is an element of unpredictability in death,	The Committee noted that dying in the preferred place of death may be a benefit from a service delivery point of view. When preference on the place of death is not reported by papers, the Committee acknowledged that there is no clear cut data on whether dying at home would be a service delivery benefit. Dying at home might not be as cheap as might be expected, as evidence is

Outcome	Why is important in EOLC? Why did we choose this particular outcome?	What is the interpretation of the outcome? For example, is an increase in the scale/n of events of the outcome a benefit or harm?	
		From a clinical point of view	From a service delivery point of view (i.e. costs, logistics)
	However, this was identified as critical outcome as it is more often reported by papers. Where the preference was not reported by the paper, and only the actual place of death was reported, the outcome was extracted but downgraded for indirectness.	people might die in a place they did not prefer, and make choices different from their preferences. When preference is not reported by papers, the Committee found it equally difficult to interpret whether dying in a certain setting would be a clinically important benefit for patients. As being transported to hospital is relatively 'easier' (for example, by calling an ambulance) for people at the end of life, the Committee felt it was fair to assume that people that die at home also wished to die at home. For this reason, the Committee assumed that dying at home or in the community could generally be interpreted as a clinically important benefit. On the contrary, it was more difficult for the Committee to establish whether dying in hospital could be interpreted as a clinically important benefit or harm. In fact, the Committee noted it is more difficult for people admitted to hospital to get back home in their last days, even if that was their preference. However, some people think it would be best to die in hospital, e.g. as they would expect better treatment.	mixed and the available analyses might not have accounted for all costs potentially involved (for example:GP visits). The margin of savings from supporting people to die at home could be small. However it was noted that the health economics evidence from the community services reviews (Q9 and Q12) showed benefit of additional services to aid keeping people in their usual place of residence. The Committee therefore decided not to generalise on the direction of effect for this outcome and to consider case by case in the context of the single studies.
Preferred and actual place of care (Dichotomous)	The Committee recognised that a critical issue for patients is where they spend their time leading to death.	It was difficult for the Committee to say whether being cared for at home or in the community would represent a clinically important benefit for patients. People	The Committee made similar considerations to those recorded for the 'preferred and actual place of death' outcome. The Committee

Outcome	Why is important in EOLC? Why did we choose this particular outcome?	What is the interpretation of the outcome? For example, is an increase in the scale/n of events of the outcome a benefit or harm?	
		From a clinical point of view	From a service delivery point of view (i.e. costs, logistics)
	Where the preference was not reported by the paper, and only the actual place of care was reported, the outcome was extracted but downgraded for indirectness.	change their views about their preferred place of care, depending on the circumstances in which the question is asked (for example: depending on the stage of the disease, the availability of carers).	therefore decided not to generalise on the direction of effect for this outcome and to consider case by case in the context of the single studies.
Longevity and/or physical/psychological health of carer (Continuous)	The Committee deemed it was important to include this outcome, as survival time of carers can be shortened by the death of the patient (or their quality of life in their last year of life). The death of a person can have effects such as for example increase in stress, or increase in suicide rate of carers.	See comments on the 'quality of life of carers' outcome.	The Committee commented that it was difficult to say if the longevity or health of carers might have an impact on service delivery, for example, if carers would access more services if their health deteriorated as a consequence of decreased quality of life of the person in their last year of life. This has not been captured by research so far. The Committee therefore decided not to generalise on the direction of effect for this outcome and to consider case by case in the context of the single studies.

Table 8: Important outcome measures

Outcome	Why it's important in EOLC? Why did we choose this particular outcome?	What is the interpretation of the outcome? Is an increase in the scale/n of events of the outcome a benefit or a harm?	
		From a clinical point of view	From a service delivery point of view (i.e. costs, logistics)
Length of survival (Continuous)	The Committee acknowledged that the length of survival is often not the most important outcome for people in their last year of life, compared to e.g. quality of life. For patients, what matters is the balance between length of survival and quality of survival.	The Committee noted that it was difficult to interpret this outcome, as its importance would depend on where in a disease trajectory it was measured. For example in people with cancer, where there are options to live much longer, this is going to be increasingly important. The Committee also commented that while this is clearly an important outcome on a patient-level, it is difficult to generalise whether prolonging length of survival would be interpreted as a clinically important benefit across a population. See also comments on the 'quality of life' outcome.	See comments on the 'quality of life' outcome.
Length of stay (in hospital) (Continuous) Hospitalisation (Dichotomous)	The Committee agreed that length of stay and hospitalisation were important outcomes. In clinical practice it is important to ensure that the hospital stay is as short as possible for people in their last year of life, provided this is appropriate for the patient.	The Committee agreed that a shorter hospital stay and fewer hospitalisation can be interpreted as a clinically important benefit.	The Committee agreed that a shorter hospital stay and fewer hospitalisation can be interpreted as a service delivery benefit.
Number of hospital visits (outpatients) (Continuous or Dichotomous)	*As above*	The Committee discussed whether attending a higher number of hospital visits would be a clinically important benefit for a person in their last year of life. The Committee commented that providing more access to outpatients might be beneficial in the earlier stages of the last year of life, but in the last months this	The Committee noted that it is difficult to say whether attending a higher number of hospital visits should be interpreted as a service delivery harm or benefit. In fact, sometimes attending more outpatients visits can support the patient to stay at home or in the

Outcome	Why it's important in EOLC? Why did we choose this particular outcome?	What is the interpretation of the outcome? Is an increase in the scale/n of events of the outcome a benefit or a harm?	
		From a clinical point of view	From a service delivery point of view (i.e. costs, logistics)
		could be less and less necessary, as there would be fewer opportunities for prolonging life. Furthermore, the Committee were aware that for some patients, outpatients appointments could be unnecessary and not well coordinated. Overall, in consideration the balance of benefit/harm that changes across the last year of life, the Committee therefore decided not to generalise on the direction of effect for this outcome and to consider case by case in the context of the single studies.	community and reduce hospital admissions. In these cases, increasing the patients' contacts with the hospital could help them stay in their preferred place of residence. However, the Committee agreed that it is difficult to generalise to the whole population at the end of life and decided to assess the direction of effect for this outcome on a case by case basis, in the context of the single studies.
Number of visits to accident and emergency (Dichotomous)	The Committee agreed that the number of visits to A&E was an important outcome, as it is a measure of whether a service works well from a population perspective	The Committee discussed the direction of effect for e.g. an increase in number of visits to A&E and found it difficult to interpret. They acknowledged that whether an increased number of visits to A&E is a benefit or a harm depends on the reason for the visit. For example, a visit may just be just for assessment and not for admission to hospital. Furthermore, the lay members noted that few patients think A&E is a positive experience, although carers might feel different. The Committee therefore decided not to generalise on the direction of effect for this outcome and to consider case by case in the context of the single studies.	The Committee discussed the value of this outcome from a service delivery point of view. In general terms the Committee felt that an increased number of visits to A&E might be interpreted as service delivery harm, as people get to the hospital rather than receiving care in the community. However, the Committee acknowledged that sometimes there might not be any other option then accessing A&E. Therefore, the Committee agreed to assess the direction of effect for this outcome on a case by case basis, in the context of the single studies.

Outcome	Why it's important in EOLC? Why did we choose this particular outcome?	What is the interpretation of the outcome? Is an increase in the scale/n of events of the outcome a benefit or a harm?	
		From a clinical point of view	From a service delivery point of view (i.e. costs, logistics)
Number of unscheduled admissions (Continuous/dichotomous)	The Committee agreed that the number of unscheduled admissions was an important outcome. For the purposes of this guideline, the Committee defined admission as a stay in hospital (i.e. the occupation of an inpatient bed). Where the 'unscheduled' aspect of admission was not reported by the paper, and only the 'number of admissions' was reported, the outcome was extracted but downgraded for indirectness.	The Committee discussed the value of this outcome and found it difficult to interpret. In general terms the Committee felt that an increased number of unscheduled admissions might be interpreted as clinical harm. However, the Committee considered that sometimes there might not be an alternative place for patients to be treated, and an increased number of unscheduled admissions might therefore interpreted as a clinically important benefit. The direction of effect would also depends on the time before death e.g. whether an increased number of unscheduled admission would be a benefit or a harm depends on whether the patient is in their last weeks/days or months before death.	The Committee agreed that an increased number of unscheduled admissions is a service delivery harm. In fact, it indicates how effective the services are to prevent an unscheduled admission (e.g. lack of care planning, alternative community services).
Use of community services (Dichotomous)	The Committee agreed that the use of community services was an important outcome. For the purposes of this guideline, the Committee defined 'community' as any care settings outside hospital, e.g. a patient's usual place of residence, GP surgeries, and hospice-led community care.	The Committee discussed the interpretation of this outcome both from a clinical and service delivery point of view. The Committee were aware of a trend for the NHS to decentralise end of life care from hospitals to the community. It is usually assumed that a decrease in hospital admissions would be mirrored by an increase in community services utilisation. However, the Committee stressed that the use of community services can also be unrelated from hospital admissions. For example, an increase in hospital admission could be related to the availability and quality of community services. Overall, the Committee agreed on the difficult interpretation of this outcome and decided to assess the direction of effect for this outcome on a case by case basis, in the context of the single studies.	

Outcome	Why it's important in EOLC? Why did we choose this particular outcome?	What is the interpretation of the outcome? Is an increase in the scale/n of events of the outcome a benefit or a harm?	
		From a clinical point of view	From a service delivery point of view (i.e. costs, logistics)
Avoidable/inappropriate admissions to ICU (Dichotomous)	The Committee agreed that avoidable/inappropriate admissions to ICU would be an important outcome. The Committee acknowledged that defining inappropriate or avoidable is difficult, therefore it is unlikely to be reported by studies. Where the 'inappropriate/avoidable' aspect was not reported by individual papers, the team extracted the outcome 'admissions to ICU', downgrading for indirectness.	The Committee agreed that an increase in avoidable/inappropriate admissions to ICU is a clinical harm. However, when papers do not report whether the admission was 'avoidable/inappropriate', the interpretation of the outcome becomes more difficult. The Committee also acknowledged that the case of sudden death is different, as people would be more likely to be admitted to ICU. The Committee decided to assess the direction of effect for this outcome on a case by case basis, in the context of the single studies.	The Committee agreed that an increase in avoidable/inappropriate admissions to ICU is a service delivery harm. However, when the 'avoidable/inappropriate' aspect is not reported, it's difficult to say. The Committee decided to assess the direction of effect for this outcome on a case by case basis, in the context of the single studies.
Inappropriate attempts at cardiopulmonary resuscitation (Dichotomous)	The Committee agreed that inappropriate attempts at cardiopulmonary resuscitation would be an important outcome.	The Committee agreed that an increase in inappropriate attempts at cardiopulmonary resuscitation would be a clinically important harm.	The Committee agreed that an increase in inappropriate attempts at cardiopulmonary resuscitation would be a service delivery harm.
Staff satisfaction (Continuous)	The Committee agreed staff satisfaction would be an important outcome.	The Committee agreed that an increase in staff satisfaction is a clinical benefit.	The Committee agreed that an increase in staff satisfaction is a service delivery benefit.
Patient/carer reported outcomes (satisfaction) (Continuous)	The Committee agreed patient or carer reported outcomes such as satisfaction would be an important outcome.	The Committee agreed that an increase in patient satisfaction is a clinical benefit. However, they noted that an increase in carer satisfaction might not be a clinical benefit for the patient. For this reason, the Committee decided to assess the direction of effect for this outcome on a case by case basis, in the context of the single studies.	The Committee agreed that an increase in patient satisfaction is a service delivery benefit. However, an increase in carer satisfaction might not be a service delivery benefit. For this reason, the Committee decided to assess the direction of effect for this outcome on a case by case basis, in the context of the single studies.

Outcome	Why it's important in EOLC? Why did we choose this particular outcome?	What is the interpretation of the outcome? Is an increase in the scale/n of events of the outcome a benefit or a harm?	
		From a clinical point of view	From a service delivery point of view (i.e. costs, logistics)
Carer health (for example:GP visits, mental health, school/work attendance) (Continuous/Dichotomous)	The Committee agreed that carer health would be an important outcome.	An improvement in carer health is a clinical benefit. However, sometimes the worsening of carer health is due to caring for the patient, resulting in better health for the patient. For this reason, the Committee decided to assess the direction of effect for this outcome on a case by case basis, in the context of the single studies.	The Committee acknowledged that the interpretation of the carer health outcome from a service delivery perspective was difficult to measure. The Committee decided to assess the direction of effect for this outcome on a case by case basis, in the context of the single studies.

3.3.6 Measuring outcomes

A number of validated scores/scales were used in the evidence base to report the identified outcomes. Table 9 outlines and describes the commonly used scales/scores used throughout the research.

Table 9: Validated scores/scales

Score/scale	Description
BDI-II	Beck Depression Inventory-II. A 21-item measure of depressive symptoms. A threshold of 20 or greater corresponds to moderate and severe depression. Scores range from 0-63; high score reflects a poor outcome.
BHS	Beck Hopelessness Scale. A 20-item true/false scale developed to quantify hopelessness and negative expectancies. Scores range from 0-20; high scores reflecting increased hopelessness.
CBI	Caregiver Burden Inventory (physical burden sub-scale), a 4-item scale used to assess the impact of caregiving on caregivers' health. Scores range from 0-4; high score reflects a poor outcome.
CBS	Caregiver Burden Scale. Demand subscale measures how much time a caregiver devoted to 14 caregiving tasks. Difficulty subscale assesses subjective caregiving burden and measures how difficult caregivers appraised the undertaking of the 14 different tasks. Scores on each subscale range from 0 to 70 with higher scores reflective of greater objective and/or subjective caregiver burden.
CES-D	Centre for Epidemiologic Studies Depression Scale. Asks the frequency of depressive symptoms over the past week, with a score of ≥16 indicating clinically significant levels of depression. Scores range from 0-60; high score reflects a poor outcome.
CGI	Complicated Grief Inventory. Identifies the grief symptoms proposed as criteria for prolonged grief disorder (PGD). Scores range from 0-52; high score reflects a poor outcome.
CQOL-C	Caregiver Quality of Life Scale – Cancer. A 35-item self-report measure; items measure impact of caregiving on a person's physical, emotional, and spiritual well-being and on his or her relationship with the care recipient and family. Scores range from 0 to 140; higher scores indicate worse QOL.
DAQ	Death anxiety questionnaire. Scores range from 15 to 75; higher scores indicate worse anxiety. Igbo version of tool.
DASS	Depression Anxiety Stress Scales. Questionnaire consisting of 14 questions, rated on a 4-point Likert scale. Scores range from 0-42; high score reflects a poor outcome.
GHQ	General Health Questionnaire. A 12-item measure of psychological distress. High score reflects a poor outcome.
HADS	Hospital Anxiety and Depression Scale provides self-rated levels of anxiety and depression. Score of 8-10 indicates possible anxiety or depression, score of ≥11 indicates definite anxiety or depression. Scores range from 0-21; high score reflects a poor outcome.
K10	Kessler Psychological Distress Scale. A 10-item questionnaire intended to yield a global measure of distress based on questions about anxiety and depressive symptoms that a person has experienced in the most recent 4 week period. Scores range from 10-50; high score reflects a poor outcome. Igbo version of tool.
LASA	Linear Analogue Self-Assessment measures overall quality of life, physical wellbeing, social activity, spiritual wellbeing, pain, fatigue, support, financial concerns, and legal concerns. Scores range from 0-100; high score reflects a positive outcome.

Score/scale	Description
MBCB	Montgomery–Borgatta CG Burden Scale, which includes objective, demand, and stress burden subscales. High objective burden score (range, 6 to 30; > 23 indicates clinical significance) suggests interference with the CG's private, social, and recreational time and normal daily routine; high demand burden score (range, 4 to 20; > 15 indicates clinical significance) indicates that the CG feels overstrained by his or her caregiving demands; high stress burden score (range, 4 to 20; > 13.5 indicates clinical significance) signals strained emotional demands related to caregiving.
MSAS	Memorial Symptom Assessment Scale. Assesses burden; includes 24 patient symptoms such as pain, lack of energy, diarrhoea, and shortness of breath. Adapted for carers. High score reflects a poor outcome.
MQOL	McGill Quality of Life Questionnaire. A 17-item patient rated measure of quality of life for patients receiving palliative care. Scores range from 0-10; high score reflects a positive outcome.
PG-13	Prigerson Inventory of Complicated Grief-Short Form. A 13-item short form measure including questions to assess pathological grief over the past month. Scores range from 0-52; high score reflects a poor outcome.
PHQ-9	Patient Health Questionnaire. A tool for screening, diagnosing, monitoring, and measuring the severity of depression. Scores range from 0-27; high score reflects a poor outcome.
PROMIS short form anxiety scale	6-item Patient Reported Outcomes Measurement Information System – short form anxiety measure assesses fear, anxious misery, and hyper-arousal over the past week. Scores range from 6-30; high score reflects a poor outcome.
PROMIS short form depression scale	6-item Patient Reported Outcomes Measurement Information System short-form depression measure assesses negative mood and views of the self over the past week. Scores range from 6-30; high score reflects a poor outcome.
SV-POMS	A subset of negative mood items from the Shortened Version Profile of Mood States used to assess caregiver negative mood. Scores range from 0-4; high score reflects a poor outcome.
SWC-EOLD	Satisfaction with care at the end of life in dementia. Scores range from 0-42; high score reflects a positive outcome.
SWLS	Satisfaction With Life Scale. Scores range from 5-25; high score reflects a positive outcome.
ZBI	The 12-item short form of the Zarit Burden Interview taps the constructs of personal and role strain. Score ranges vary; high score reflects a poor outcome.

3.3.7 Clinical evidence statements

Clinical evidence statements are summary statements that are included in each evidence report, and which summarise the key features of the clinical effectiveness evidence presented. The wording of the evidence statements reflects the certainty or uncertainty in the estimate of effect. The evidence statements are presented by outcome and encompass the following key features of the evidence:

- The number of studies and the number of participants for a particular outcome.
- An indication of the direction of clinical importance (if one treatment is beneficial or harmful compared to the other, or whether there is no difference between the 2 tested treatments).
- A description of the overall quality of the evidence (GRADE overall quality).

3.4 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.[15]

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 literature.
- Undertook new cost-effectiveness analysis in priority areas.

3.4.1 Literature review

The health economists:

- Identified potentially relevant studies for each review question 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.[15]
- Extracted key information about the studies' methods and results into health economic evidence tables (which can be found in appendices to the relevant evidence reports).
- Generated summaries of the evidence in NICE health economic evidence profile tables (included in the relevant evidence report for each review question) – see below for details.

3.4.1.1 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.

Remaining health economic studies were prioritised for inclusion based on their relative applicability to the development of this guideline and the study limitations. However, in this guideline, no economic studies were excluded on the basis that more applicable evidence was available.

For more details about the assessment of applicability and methodological quality see Table 10 below and the economic evaluation checklist (appendix H of the NICE guidelines manual[15]) and the health economics review protocol, which can be found in each of the evidence reports.

When no relevant health economic studies were found from the economic literature review, relevant UK NHS unit costs related to the compared interventions were presented to the committee to inform the possible economic implications of the recommendations.

3.4.1.2 NICE health economic evidence profiles

NICE health economic evidence profile tables were used to summarise cost and cost-effectiveness estimates for the included health economic studies in each evidence review report. 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.[15] 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 in the study, as well as information about the assessment of uncertainty in the analysis. See Table 10 for more details.

When a non-UK study was included in the profile, the results were converted into pounds sterling using the appropriate purchasing power parity.[18]

Table 10: 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[15]

3.4.2 Undertaking new health economic analysis

As well as reviewing the published health economic literature for each review question, as described above, new health economic costing analysis was undertaken by the health economist in selected areas. Priority areas for new analysis were agreed by the committee after formation of the review questions and consideration of the existing health economic evidence.

The committee identified out-of-hours services and community based end-of-life services as the highest priority areas for original health economic modelling. Their rationale for prioritising these areas was largely due to the high likeliness that potential recommendations would result in significant resource impact, and the committee's view that service delivery improvements in these areas could benefit people in the last year of life and/or people caring for those in the last year of life. The following general principles were adhered to in developing the cost analysis:

- Methods were consistent with the NICE reference case for interventions with health outcomes in NHS settings.[15, 17]
- The committee was involved in the design of the model, selection of inputs and interpretation of the results.
- Inputs were based on data from grey literature reports identified through the call for evidence conducted for the guideline supplemented with other published data sources where possible.
- When data was not available committee expert opinion was used to populate resource use estimates in the analysis.
- Inputs and assumptions were reported fully and transparently.
- Limitations of the analysis were discussed.
- The analysis was peer-reviewed by another health economist at the NGC.

Full methods and results of the costing analysis for out-of-hours, community based interventions are described in a separate economic analysis report.

3.4.3 Cost-effectiveness criteria

NICE's report 'Social value judgements: principles for the development of NICE guidance' sets out the principles that committees should consider when judging whether an intervention offers good value for money.[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 the factors set out in 'Social value judgements: principles for the development of NICE guidance'.[16]

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.

As QALYs were not commonly reported in the evidence identified throughout the guideline, a cost per QALY approach was not appropriate to determine cost effectiveness. A different costing approach was therefore taken for the economic analysis. The committee considered recommending interventions where it was deemed possible that the upfront costs of the intervention or service changes could be offset by the long term cost savings they might produce.

As well as their being little evidence identified that reported QALYs, a QALY maximising approach was not considered appropriate for end of life services. This is because the majority of end of life services are not intended to increase survival or improve quality of life (the two elements that make up a QALY), they are intended to maintain quality of life (if possible) but most often they are aimed at improving the quality of care, for both the person being cared for, those important to them, and their carers if appropriate.

3.4.4 In the absence of health economic evidence

When no relevant published health economic studies were found, and a new analysis was not prioritised, the committee made a qualitative judgement about cost effectiveness by considering expected differences in resource use between options and relevant UK NHS unit costs, alongside the results of the review of clinical effectiveness evidence.

The UK NHS costs reported in the guideline are those that were presented to the committee and were correct at the time recommendations were drafted. They may have changed subsequently before the time of publication. However, we have no reason to believe they have changed substantially.

3.5 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 reports [A–M]).
- Evidence tables of the clinical and health economic evidence reviewed from the literature. All evidence tables can be found in appendices to the relevant evidence reports.
- Forest plots (in appendices to the relevant evidence reports).
- A description of the methods and results of the cost-effectiveness analysis undertaken for the guideline (in a separate economic analysis report).

Recommendations were drafted 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. Firstly, the net clinical benefit over harm (clinical effectiveness) was considered, focusing on the critical outcomes. When this was done informally, the committee took into account the clinical benefits and harms when one intervention 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 interventions.

When clinical and health economic evidence was of poor quality, conflicting or absent, the committee drafted recommendations 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 and to make no recommendation, taking into account the potential harm of failing to make a clear recommendation (see section 3.5.1 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[15]).

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

3.5.1 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.

3.5.2 Validation process

This guidance is subject to a 6-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.

3.5.3 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.

3.5.4 Disclaimer

Healthcare providers need to use clinical judgement, knowledge and expertise when deciding whether it is appropriate to apply guidelines. The recommendations cited here are a guide and may not be appropriate for use in all situations. The decision to adopt any of the recommendations cited here must be made by practitioners in light of individual patient circumstances, the wishes of the patient, clinical expertise and resources.

The National Guideline Centre disclaims any responsibility for damages arising out of the use or non-use of this guideline and the literature used in support of this guideline.

3.5.5 Funding

The National Guideline Centre was commissioned by the National Institute for Health and Care Excellence to undertake the work on this guideline.

4 Acronyms and abbreviations

Acronym or abbreviation	Description
ACP	Advanced care planning
ADRT	Advance decision to refuse treatment
A&E	Accident and emergency
CHF/HF	Chronic heart failure
COPD	Chronic obstructive pulmonary disease
DNACPR	Do not attempt cardiopulmonary resuscitation
ED	Emergency Department
EOLC	End of life care
Committee	Guideline committee
GP	General Practitioner
GRADE	Grading of Recommendations Assessment, Development and Evaluation
HCP	Health care professional
HHU	Home hospice unit
ITU/ICU	Intensive care units
LYOL	Last year of life
MDT	Multidisciplinary team
MID	Minimal important difference
NGC	National Guideline Centre
NHS	National Health Service
NICE	National Institute for Health and Care Excellence
NRS	Non-randomised study
PC	Palliative care
PCU	Palliative care unit
QALY	Quality adjusted life years
RCT	Randomised control trial
RRT/RRS	Rapid Response Team
SC	Standard care
SR	Systematic review
TEP	Treatment escalation plan

5 Glossary

The NICE Glossary can be found at www.nice.org.uk/glossary.

5.1 Clinical terms

Term	Definition
Advance care planning	Advance Care Planning is the process of discussing your preferences and wishes about future treatment and care with those close to you and your healthcare team. This may include: • talking about where you want to be cared for • identifying the people you'd like to be consulted about your care • making treatment decisions in advance Advance Care Planning helps health and care professionals, and those close to you, to understand how you want to be cared for if you become too ill to make decisions or speak for yourself. You can also formally document your wishes as part of this process in an Advance Care Plan. (https://compassionindying.org.uk/making-decisions-and-planning-your-care/planning-ahead/advance-care-planning/)
Community support	Community support refers to any support outside of hospital that is provided by NHS services.
Dementia	An acquired organic mental disorder with loss of intellectual abilities of sufficient severity to interfere with social or occupational functioning. The dysfunction is multifaceted and involves memory, behaviour, personality, judgment, attention, spatial relations, language, abstract thought, and other executive functions. The intellectual decline is usually progressive, and initially spares the level of consciousness.
Emergency Department	A medical facility specialising in emergency medicine and the acute care of patients who present without a prior appointment. Also known as an accident and emergency department (A&E)
General Practitioner	A doctor based in the community who treats patients with minor or chronic illnesses.
Hospice	Facilities or services which are especially devoted to providing palliative and supportive care to the patient with a terminal illness and to the patient's family.
Hospital	An institution providing medical and surgical treatment and nursing care for sick or injured people.
Multidisciplinary team/ multiprofessional team	All members of the healthcare and social care team that provide care, including clinical staff and social care staff in hospital, community and nursing home or residential settings.
Palliative care	Care alleviating symptoms without curing the underlying disease.
Pharmacist	A person who is professionally qualified to prepare and dispense medicinal drugs.

5.2 Methodological terms

Term	Definition
Abstract	Summary of a study, which may be published alone or as an introduction to a full scientific paper.

Term	Definition
Algorithm (in guidelines)	A flow chart of the clinical decision pathway described in the guideline, where decision points are represented with boxes, linked with arrows.
Allocation concealment	The process used to prevent advance knowledge of group assignment in an RCT. The allocation process should be impervious to any influence by the individual making the allocation, by being administered by someone who is not responsible for recruiting participants.
Applicability	How well the results of a study or NICE evidence review can answer a clinical question or be applied to the population being considered.
Arm (of a clinical study)	Subsection of individuals within a study who receive one particular intervention, for example placebo arm.
Association	Statistical relationship between 2 or more events, characteristics or other variables. The relationship may or may not be causal.
Base case analysis	In an economic evaluation, this is the main analysis based on the most plausible estimate of each input. In contrast, see Sensitivity analysis.
Baseline	The initial set of measurements at the beginning of a study (after run-in period where applicable), with which subsequent results are compared.
Bayesian analysis	A method of statistics, where a statistic is estimated by combining established information or belief (the 'prior') with new evidence (the 'likelihood') to give a revised estimate (the 'posterior').
Before-and-after study	A study that investigates the effects of an intervention by measuring particular characteristics of a population both before and after taking the intervention, and assessing any change that occurs.
Bias	Influences on a study that can make the results look better or worse than they really are. (Bias can even make it look as if a treatment works when it does not.) Bias can occur by chance, deliberately or as a result of systematic errors in the design and execution of a study. It can also occur at different stages in the research process, for example, during the collection, analysis, interpretation, publication or review of research data. For examples see selection bias, performance bias, information bias, confounding factor, and publication bias.
Blinding	A way to prevent researchers, doctors and patients in a clinical trial from knowing which study group each patient is in so they cannot influence the results. The best way to do this is by sorting patients into study groups randomly. The purpose of 'blinding' or 'masking' is to protect against bias. A single-blinded study is one in which patients do not know which study group they are in (for example whether they are taking the experimental drug or a placebo). A double-blinded study is one in which neither patients nor the researchers and doctors know which study group the patients are in. A triple blind study is one in which neither the patients, clinicians or the people carrying out the statistical analysis know which treatment patients received.
Carer	A carer is someone who helps another person, usually a relative, partner or friend, in their day to day life. This term does not refer to someone who provides care professionally or through a voluntary organisation
Case–control study	A study to find out the cause(s) of a disease or condition. This is done by comparing a group of patients who have the disease or condition (cases) with a group of people who do not have it (controls) but who are otherwise as similar as possible (in characteristics thought to be unrelated to the causes of the disease or condition).

Term	Definition
	This means the researcher can look for aspects of their lives that differ to see if they may cause the condition. For example, a group of people with lung cancer might be compared with a group of people the same age that do not have lung cancer. The researcher could compare how long both groups had been exposed to tobacco smoke. Such studies are retrospective because they look back in time from the outcome to the possible causes of a disease or condition.
Case series	Report of a number of cases of a given disease, usually covering the course of the disease and the response to treatment. There is no comparison (control) group of patients.
Clinical efficacy	The extent to which an intervention is active when studied under controlled research conditions.
Clinical effectiveness	How well a specific test or treatment works when used in the 'real world' (for example, when used by a doctor with a patient at home), rather than in a carefully controlled clinical trial. Trials that assess clinical effectiveness are sometimes called management trials. Clinical effectiveness is not the same as efficacy.
Clinician	A healthcare professional who provides patient care. For example, a doctor, nurse or physiotherapist.
Cochrane Review	The Cochrane Library consists of a regularly updated collection of evidence-based medicine databases including the Cochrane Database of Systematic Reviews (reviews of randomised controlled trials prepared by the Cochrane Collaboration).
Cohort study	A study with 2 or more groups of people – cohorts – with similar characteristics. One group receives a treatment, is exposed to a risk factor or has a particular symptom and the other group does not. The study follows their progress over time and records what happens. See also observational study.
Comorbidity	A disease or condition that someone has in addition to the health problem being studied or treated.
Comparability	Similarity of the groups in characteristics likely to affect the study results (such as health status or age).
Concordance	This is a recent term whose meaning has changed. It was initially applied to the consultation process in which doctor and patient agree therapeutic decisions that incorporate their respective views, but now includes patient support in medicine taking as well as prescribing communication. Concordance reflects social values but does not address medicine-taking and may not lead to improved adherence.
Confidence interval (CI)	There is always some uncertainty in research. This is because a small group of patients is studied to predict the effects of a treatment on the wider population. The confidence interval is a way of expressing how certain we are about the findings from a study, using statistics. It gives a range of results that is likely to include the 'true' value for the population. The CI is usually stated as '95% CI', which means that the range of values has a 95 in a 100 chance of including the 'true' value. For example, a study may state that "based on our sample findings, we are 95% certain that the 'true' population blood pressure is not higher than 150 and not lower than 110". In such a case the 95% CI would be 110 to 150. A wide confidence interval indicates a lack of certainty about the true effect of the test or treatment – often because a small group of patients has been studied. A narrow confidence interval indicates a more precise estimate (for example, if a large number of patients have been studied).

Term	Definition
Confounding factor	Something that influences a study and can result in misleading findings if it is not understood or appropriately dealt with. For example, a study of heart disease may look at a group of people that exercises regularly and a group that does not exercise. If the ages of the people in the 2 groups are different, then any difference in heart disease rates between the 2 groups could be because of age rather than exercise. Therefore age is a confounding factor.
Consensus methods	Techniques used to reach agreement on a particular issue. Consensus methods may be used to develop NICE guidance if there is not enough good quality research evidence to give a clear answer to a question. Formal consensus methods include Delphi and nominal group techniques.
Control group	A group of people in a study who do not receive the treatment or test being studied. Instead, they may receive the standard treatment (sometimes called 'usual care') or a dummy treatment (placebo). The results for the control group are compared with those for a group receiving the treatment being tested. The aim is to check for any differences. Ideally, the people in the control group should be as similar as possible to those in the treatment group, to make it as easy as possible to detect any effects due to the treatment.
Cost–benefit analysis (CBA)	Cost–benefit analysis is one of the tools used to carry out an economic evaluation. The costs and benefits are measured using the same monetary units (for example, pounds sterling) to see whether the benefits exceed the costs.
Cost–consequences analysis (CCA)	Cost–consequences analysis is one of the tools used to carry out an economic evaluation. This compares the costs (such as treatment and hospital care) and the consequences (such as health outcomes) of a test or treatment with a suitable alternative. Unlike cost–benefit analysis or cost-effectiveness analysis, it does not attempt to summarise outcomes in a single measure (like the quality-adjusted life year) or in financial terms. Instead, outcomes are shown in their natural units (some of which may be monetary) and it is left to decision-makers to determine whether, overall, the treatment is worth carrying out.
Cost-effectiveness analysis (CEA)	Cost-effectiveness analysis is one of the tools used to carry out an economic evaluation. The benefits are expressed in non-monetary terms related to health, such as symptom-free days, heart attacks avoided, deaths avoided or life years gained (that is, the number of years by which life is extended as a result of the intervention).
Cost-effectiveness model	An explicit mathematical framework, which is used to represent clinical decision problems and incorporate evidence from a variety of sources in order to estimate the costs and health outcomes.
Cost–utility analysis (CUA)	Cost–utility analysis is one of the tools used to carry out an economic evaluation. The benefits are assessed in terms of both quality and duration of life, and expressed as quality-adjusted life years (QALYs). See also utility.
Credible interval (CrI)	The Bayesian equivalent of a confidence interval.
Deterministic analysis	In economic evaluation, this is an analysis that uses a point estimate for each input. In contrast, see Probabilistic analysis
Discounting	Costs and perhaps benefits incurred today have a higher value than costs and benefits occurring in the future. Discounting health benefits reflects individual preference for benefits to be experienced in the present rather than the future. Discounting costs reflects individual preference for costs to be experienced in the future rather than the present.

Term	Definition
Disutility	The loss of quality of life associated with having a disease or condition. See Utility
Dominance	A health economics term. When comparing tests or treatments, an option that is both less effective and costs more is said to be 'dominated' by the alternative.
Drop-out	A participant who withdraws from a trial before the end.
Economic evaluation	An economic evaluation is used to assess the cost effectiveness of healthcare interventions (that is, to compare the costs and benefits of a healthcare intervention to assess whether it is worth doing). The aim of an economic evaluation is to maximise the level of benefits – health effects – relative to the resources available. It should be used to inform and support the decision-making process; it is not supposed to replace the judgement of healthcare professionals. There are several types of economic evaluation: cost–benefit analysis, cost–consequences analysis, cost-effectiveness analysis, cost-minimisation analysis and cost–utility analysis. They use similar methods to define and evaluate costs, but differ in the way they estimate the benefits of a particular drug, programme or intervention.
Effect (as in effect measure, treatment effect, estimate of effect, effect size)	A measure that shows the magnitude of the outcome in one group compared with that in a control group. For example, if the absolute risk reduction is shown to be 5% and it is the outcome of interest, the effect size is 5%. The effect size is usually tested, using statistics, to find out how likely it is that the effect is a result of the treatment and has not just happened by chance (that is, to see if it is statistically significant).
Effectiveness	How beneficial a test or treatment is under usual or everyday conditions, compared with doing nothing or opting for another type of care.
Efficacy	How beneficial a test, treatment or public health intervention is under ideal conditions (for example, in a laboratory), compared with doing nothing or opting for another type of care.
Epidemiological study	The study of a disease within a population, defining its incidence and prevalence and examining the roles of external influences (for example, infection, diet) and interventions.
EQ-5D (EuroQol 5 dimensions)	A standardised instrument used to measure health-related quality of life. It provides a single index value for health status.
Evidence	Information on which a decision or guidance is based. Evidence is obtained from a range of sources including randomised controlled trials, observational studies, expert opinion (of clinical professionals or patients).
Exclusion criteria (literature review)	Explicit standards used to decide which studies should be excluded from consideration as potential sources of evidence.
Exclusion criteria (clinical study)	Criteria that define who is not eligible to participate in a clinical study.
Extended dominance	If Option A is both more clinically effective than Option B and has a lower cost per unit of effect, when both are compared with a do-nothing alternative then Option A is said to have extended dominance over Option B. Option A is therefore cost effective and should be preferred, other things remaining equal.
Extrapolation	An assumption that the results of studies of a specific population will also hold true for another population with similar characteristics.
Follow-up	Observation over a period of time of an individual, group or initially defined population whose appropriate characteristics have been assessed in order to observe changes in health status or health-related variables.

Term	Definition
Generalisability	The extent to which the results of a study hold true for groups that did not participate in the research. See also external validity.
Gold standard	A method, procedure or measurement that is widely accepted as being the best available to test for or treat a disease.
GRADE, GRADE profile	A system developed by the GRADE Working Group to address the shortcomings of present grading systems in healthcare. The GRADE system uses a common, sensible and transparent approach to grading the quality of evidence. The results of applying the GRADE system to clinical trial data are displayed in a table known as a GRADE profile.
Harms	Adverse effects of an intervention.
Health economics	Study or analysis of the cost of using and distributing healthcare resources.
Health-related quality of life (HRQoL)	A measure of the effects of an illness to see how it affects someone's day-to-day life.
Heterogeneity or Lack of homogeneity	The term is used in meta-analyses and systematic reviews to describe when the results of a test or treatment (or estimates of its effect) differ significantly in different studies. Such differences may occur as a result of differences in the populations studied, the outcome measures used or because of different definitions of the variables involved. It is the opposite of homogeneity.
Imprecision	Results are imprecise when studies include relatively few patients and few events and thus have wide confidence intervals around the estimate of effect.
Inclusion criteria (literature review)	Explicit criteria used to decide which studies should be considered as potential sources of evidence.
Incremental analysis	The analysis of additional costs and additional clinical outcomes with different interventions.
Incremental cost	The extra cost linked to using one test or treatment rather than another. Or the additional cost of doing a test or providing a treatment more frequently.
Incremental cost-effectiveness ratio (ICER)	The difference in the mean costs in the population of interest divided by the differences in the mean outcomes in the population of interest for one treatment compared with another.
Incremental net benefit (INB)	The value (usually in monetary terms) of an intervention net of its cost compared with a comparator intervention. The INB can be calculated for a given cost-effectiveness (willingness to pay) threshold. If the threshold is £20,000 per QALY gained then the INB is calculated as: (£20,000 × QALYs gained) − Incremental cost.
Indirectness	The available evidence is different to the review question being addressed, in terms of PICO (population, intervention, comparison and outcome).
Intention-to-treat analysis (ITT)	An assessment of the people taking part in a clinical trial, based on the group they were initially (and randomly) allocated to. This is regardless of whether or not they dropped out, fully complied with the treatment or switched to an alternative treatment. Intention-to-treat analyses are often used to assess clinical effectiveness because they mirror actual practice: that is, not everyone complies with treatment and the treatment people receive may be changed according to how they respond to it.
Intervention	In medical terms this could be a drug treatment, surgical procedure, diagnostic or psychological therapy. Examples of public health interventions could include action to help someone to be physically active or to eat a more healthy diet.

Term	Definition
Kappa statistic	A statistical measure of inter-rater agreement that takes into account the agreement occurring by chance.
Length of stay	The total number of days a participant stays in hospital.
Licence	See 'Product licence'.
Life years gained	Mean average years of life gained per person as a result of the intervention compared with an alternative intervention.
Long-term care	Residential care in a home that may include skilled nursing care and help with everyday activities. This includes nursing homes and residential homes.
Logistic regression or Logit model	In statistics, logistic regression is a type of analysis used for predicting the outcome of a binary dependent variable based on one or more predictor variables. It can be used to estimate the log of the odds (known as the 'logit').
Loss to follow-up	A patient, or the proportion of patients, actively participating in a clinical trial at the beginning, but whom the researchers were unable to trace or contact by the point of follow-up in the trial
Markov model	A method for estimating long-term costs and effects for recurrent or chronic conditions, based on health states and the probability of transition between them within a given time period (cycle).
Meta-analysis	A method often used in systematic reviews. Results from several studies of the same test or treatment are combined to estimate the overall effect of the treatment.
Multivariate model	A statistical model for analysis of the relationship between 2 or more predictor (independent) variables and the outcome (dependent) variable.
Net monetary benefit (NMB)	The value in monetary terms of an intervention net of its cost. The NMB can be calculated for a given cost-effectiveness threshold. If the threshold is £20,000 per QALY gained then the NMB for an intervention is calculated as: (£20,000 × mean QALYs) − mean cost. The most preferable option (that is, the most clinically effective option to have an ICER below the threshold selected) will be the treatment with the highest NMB.
Non-randomised intervention study	A quantitative study investigating the effectiveness of an intervention that does not use randomisation to allocate patients (or units) to treatment groups. Non-randomised studies include observational studies, where allocation to groups occurs through usual treatment decisions or people's preferences. Non-randomised studies can also be experimental, where the investigator has some degree of control over the allocation of treatments. Non-randomised intervention studies can use a number of different study designs, and include cohort studies, case–control studies, controlled before-and-after studies, interrupted-time-series studies and quasi-randomised controlled trials.
Number needed to treat (NNT)	The average number of patients who need to be treated to get a positive outcome. For example, if the NNT is 4, then 4 patients would have to be treated to ensure 1 of them gets better. The closer the NNT is to 1, the better the treatment. For example, if you give a stroke prevention drug to 20 people before 1 stroke is prevented, the number needed to treat is 20. See also number needed to harm, absolute risk reduction.
Observational study	Individuals or groups are observed or certain factors are measured. No attempt is made to affect the outcome. For example, an observational study of a disease or treatment would allow 'nature' or usual medical care to take its course. Changes or differences in one

Term	Definition
	characteristic (for example, whether or not people received a specific treatment or intervention) are studied without intervening. There is a greater risk of selection bias than in experimental studies.
Odds ratio	Odds are a way to represent how likely it is that something will happen (the probability). An odds ratio compares the probability of something in one group with the probability of the same thing in another. An odds ratio of 1 between 2 groups would show that the probability of the event (for example a person developing a disease, or a treatment working) is the same for both. An odds ratio greater than 1 means the event is more likely in the first group. An odds ratio less than 1 means that the event is less likely in the first group. Sometimes probability can be compared across more than 2 groups – in this case, one of the groups is chosen as the 'reference category', and the odds ratio is calculated for each group compared with the reference category. For example, to compare the risk of dying from lung cancer for non-smokers, occasional smokers and regular smokers, non-smokers could be used as the reference category. Odds ratios would be worked out for occasional smokers compared with non-smokers and for regular smokers compared with non-smokers. See also confidence interval, risk ratio.
Opportunity cost	The loss of other healthcare programmes displaced by investment in or introduction of another intervention. This may be best measured by the health benefits that could have been achieved had the money been spent on the next best alternative healthcare intervention.
Outcome	The impact that a test, treatment, policy, programme or other intervention has on a person, group or population. Outcomes from interventions to improve the public's health could include changes in knowledge and behaviour related to health, societal changes (for example, a reduction in crime rates) and a change in people's health and wellbeing or health status. In clinical terms, outcomes could include the number of patients who fully recover from an illness or the number of hospital admissions, and an improvement or deterioration in someone's health, functional ability, symptoms or situation. Researchers should decide what outcomes to measure before a study begins.
P value	The p value is a statistical measure that indicates whether or not an effect is statistically significant. For example, if a study comparing 2 treatments found that one seems more effective than the other, the p value is the probability of obtaining these results by chance. By convention, if the p value is below 0.05 (that is, there is less than a 5% probability that the results occurred by chance) it is considered that there probably is a real difference between treatments. If the p value is 0.001 or less (less than a 1% probability that the results occurred by chance), the result is seen as highly significant. If the p value shows that there is likely to be a difference between treatments, the confidence interval describes how big the difference in effect might be.
Placebo	A fake (or dummy) treatment given to participants in the control group of a clinical trial. It is indistinguishable from the actual treatment (which is given to participants in the experimental group). The aim is to determine what effect the experimental treatment has had – over and above any placebo effect caused because someone has received (or thinks they have received) care or attention.
Polypharmacy	The use or prescription of multiple medications.

Term	Definition
Posterior distribution	In Bayesian statistics this is the probability distribution for a statistic based after combining established information or belief (the prior) with new evidence (the likelihood).
Power (statistical)	The ability to demonstrate an association when one exists. Power is related to sample size; the larger the sample size, the greater the power and the lower the risk that a possible association could be missed.
Prior distribution	In Bayesian statistics this is the probability distribution for a statistic based on previous evidence or belief.
Primary care	Healthcare delivered outside hospitals. It includes a range of services provided by GPs, nurses, health visitors, midwives and other healthcare professionals and allied health professionals such as dentists, pharmacists and opticians.
Primary outcome	The outcome of greatest importance, usually the one in a study that the power calculation is based on.
Probabilistic analysis	In economic evaluation, this is an analysis that uses a probability distribution for each input. In contrast, see Deterministic analysis.
Product licence	An authorisation from the MHRA to market a medicinal product.
Prognosis	A probable course or outcome of a disease. Prognostic factors are patient or disease characteristics that influence the course. Good prognosis is associated with low rate of undesirable outcomes; poor prognosis is associated with a high rate of undesirable outcomes.
Prospective study	A research study in which the health or other characteristic of participants is monitored (or 'followed up') for a period of time, with events recorded as they happen. This contrasts with retrospective studies.
Publication bias	Publication bias occurs when researchers publish the results of studies showing that a treatment works well and don't publish those showing it did not have any effect. If this happens, analysis of the published results will not give an accurate idea of how well the treatment works. This type of bias can be assessed by a funnel plot.
Quality of life	See 'Health-related quality of life'.
Quality-adjusted life year (QALY)	A measure of the state of health of a person or group in which the benefits, in terms of length of life, are adjusted to reflect the quality of life. One QALY is equal to 1 year of life in perfect health. QALYS are calculated by estimating the years of life remaining for a patient following a particular treatment or intervention and weighting each year with a quality of life score (on a scale of 0 to 1). It is often measured in terms of the person's ability to perform the activities of daily life, freedom from pain and mental disturbance.
Randomisation	Assigning participants in a research study to different groups without taking any similarities or differences between them into account. For example, it could involve using a random numbers table or a computer-generated random sequence. It means that each individual (or each group in the case of cluster randomisation) has the same chance of receiving each intervention.
Randomised controlled trial (RCT)	A study in which a number of similar people are randomly assigned to 2 (or more) groups to test a specific drug or treatment. One group (the experimental group) receives the treatment being tested, the other (the comparison or control group) receives an alternative treatment, a dummy treatment (placebo) or no treatment at all. The groups are followed up to see how effective the experimental treatment was. Outcomes are measured at specific times and any difference in response between the groups is assessed statistically. This method is also used to reduce bias.

Term	Definition
Reference standard	The test that is considered to be the best available method to establish the presence or absence of the outcome – this may not be the one that is routinely used in practice.
Reporting bias	See 'Publication bias'.
Resource implication	The likely impact in terms of finance, workforce or other NHS resources.
Retrospective study	A research study that focuses on the past and present. The study examines past exposure to suspected risk factors for the disease or condition. Unlike prospective studies, it does not cover events that occur after the study group is selected.
Review question	In guideline development, this term refers to the questions about treatment and care that are formulated to guide the development of evidence-based recommendations.
Risk ratio (RR)	The ratio of the risk of disease or death among those exposed to certain conditions compared with the risk for those who are not exposed to the same conditions (for example, the risk of people who smoke getting lung cancer compared with the risk for people who do not smoke). If both groups face the same level of risk, the risk ratio is 1. If the first group had a risk ratio of 2, subjects in that group would be twice as likely to have the event happen. A risk ratio of less than 1 means the outcome is less likely in the first group. The risk ratio is sometimes referred to as relative risk.
Secondary outcome	An outcome used to evaluate additional effects of the intervention deemed a priori as being less important than the primary outcomes.
Selection bias	Selection bias occurs if: a) The characteristics of the people selected for a study differ from the wider population from which they have been drawn, or b) There are differences between groups of participants in a study in terms of how likely they are to get better.
Sensitivity analysis	A means of representing uncertainty in the results of economic evaluations. Uncertainty may arise from missing data, imprecise estimates or methodological controversy. Sensitivity analysis also allows for exploring the generalisability of results to other settings. The analysis is repeated using different assumptions to examine the effect on the results. One-way simple sensitivity analysis (univariate analysis): each parameter is varied individually in order to isolate the consequences of each parameter on the results of the study. Multi-way simple sensitivity analysis (scenario analysis): 2 or more parameters are varied at the same time and the overall effect on the results is evaluated. Threshold sensitivity analysis: the critical value of parameters above or below which the conclusions of the study will change are identified. Probabilistic sensitivity analysis: probability distributions are assigned to the uncertain parameters and are incorporated into evaluation models based on decision analytical techniques (for example, Monte Carlo simulation).
Significance (statistical)	A result is deemed statistically significant if the probability of the result occurring by chance is less than 1 in 20 ($p<0.05$).
Stakeholder	An organisation with an interest in a topic that NICE is developing a guideline or piece of public health guidance on. Organisations that register as stakeholders can comment on the draft scope and the draft guidance. Stakeholders may be: • manufacturers of drugs or equipment

Term	Definition
	• national patient and carer organisations • NHS organisations • organisations representing healthcare professionals.
State transition model	See Markov model
Systematic review	A review in which evidence from scientific studies has been identified, appraised and synthesised in a methodical way according to predetermined criteria. It may include a meta-analysis.
Time horizon	The time span over which costs and health outcomes are considered in a decision analysis or economic evaluation.
Transition probability	In a state transition model (Markov model), this is the probability of moving from one health state to another over a specific period of time.
Treatment allocation	Assigning a participant to a particular arm of a trial.
Univariate	Analysis which separately explores each variable in a data set.
Utility	In health economics, a 'utility' is the measure of the preference or value that an individual or society places upon a particular health state. It is generally a number between 0 (representing death) and 1 (perfect health). The most widely used measure of benefit in cost–utility analysis is the quality-adjusted life year, but other measures include disability-adjusted life years (DALYs) and healthy year equivalents (HYEs).

References

1. Aoun S, Deas K, Toye C, Ewing G, Grande G, Stajduhar K. Supporting family caregivers to identify their own needs in end-of-life care: Qualitative findings from a stepped wedge cluster trial. Palliative Medicine. 2015; 29(6):508-517

2. Aoun SM, Grande G, Howting D, Deas K, Toye C, Troeung L et al. The impact of the carer support needs assessment tool (CSNAT) in community palliative care using a stepped wedge cluster trial. PloS One. 2015; 10(4):e0123012

3. Bajwah S, Ross JR, Wells AU, Mohammed K, Oyebode C, Birring SS et al. Palliative care for patients with advanced fibrotic lung disease: A randomised controlled phase II and feasibility trial of a community case conference intervention. Thorax. 2015; 70(9):830-9

4. Candy B, Holman A, Leurent B, Davis S, Jones L. Hospice care delivered at home, in nursing homes and in dedicated hospice facilities: A systematic review of quantitative and qualitative evidence. International Journal of Nursing Studies. 2011; 48(1):121-33

5. Chapman M, Johnston N, Lovell C, Forbat L, Liu WM. Avoiding costly hospitalisation at end of life: findings from a specialist palliative care pilot in residential care for older adults. BMJ Support Palliat Care. 2016;

6. Gomes B, Calanzani N, Higginson IJ. Benefits and costs of home palliative care compared with usual care for patients with advanced illness and their family caregivers. JAMA. 2014; 311(10):1060-1061

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

8. Grande GE, Austin L, Ewing G, O'Leary N, Roberts C. Assessing the impact of a Carer Support Needs Assessment Tool (CSNAT) intervention in palliative home care: a stepped wedge cluster trial. BMJ Support Palliat Care. 2017; 7(3):326-334

9. Harding R, Epiphaniou E, Chidgey-Clark J. Needs, experiences, and preferences of sexual minorities for end-of-life care and palliative care: a systematic review. Journal of Palliative Medicine. 2012; 15(5):602-11

10. Johnston N, Lovell C, Liu WM, Chapman M, Forbat L. Normalising and planning for death in residential care: findings from a qualitative focus group study of a specialist palliative care intervention. BMJ Supportive & Palliative Care. 2016; Epublication

11. Lamont T, Richardson A. Better endings: A review of end-of-life services. Nursing Times. 2016; 112(12):16-18

12. Lucas B, Small N, Greasley P, Daley A. A reassuring presence: An evaluation of Bradford District Hospice at Home service. BMC Palliative Care. 2008; 7:9

13. Mason B, Epiphaniou E, Nanton V, Donaldson A, Shipman C, Daveson BA et al. Coordination of care for individuals with advanced progressive conditions: A multi-site ethnographic and serial interview study. British Journal of General Practice. 2013; 63(613):e580-8

14. Nakajima M, Benson D, Mucci E. 29 DOES THE PEACE TOOL ENABLE DEATH AT HOME FOR FRAIL ELDERLY PATIENTS AT THE END OF THEIR LIVES? Age and Ageing. 2015; 44(suppl_1):i7-i7

15. 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

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. 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

18. Organisation for Economic Co-operation and Development (OECD). Purchasing power parities (PPP). 2017. Available from: https://data.oecd.org/conversion/purchasing-power-parities-ppp.htm Last accessed: 25/08/2017

19. Perkins GD, Griffiths F, Slowther AM, George R, Fritz Z, Satherley P et al. Do-not-attempt-cardiopulmonary-resuscitation decisions: an evidence synthesis. Health Services and Delivery Research. 2016;

20. Petrova M, Riley J, Abel J, Barclay S. Crash course in EPaCCS (Electronic Palliative Care Coordination Systems): 8 years of successes and failures in patient data sharing to learn from. BMJ Supportive & Palliative Care. 2016; Epublication

21. Review Manager (RevMan) [Computer program]. Version 5. Copenhagen. The Nordic Cochrane Centre, The Cochrane Collaboration, 2015. Available from: http://tech.cochrane.org/Revman

22. Sandsdalen T, Grondahl VA, Hov R, Hoye S, Rystedt I, Wilde-Larsson B. Patients' perceptions of palliative care quality in hospice inpatient care, hospice day care, palliative units in nursing homes, and home care: a cross-sectional study. BMC Palliative Care. 2016; 15(1):79

23. Smith C, Whitmore D, Hough L, Gough N, Cheung C-C, Thick M et al. Implementation of a locality end of life care register, now called Electronic Palliative Care Coordination System (EPCCS) – has it improved the information shared with out of hours services? . Journal of Health Service Research. 2012;

24. Wilkinson E, Randhawa G, Brown E, Da Silva Gane M, Stoves J, Warwick G et al. Exploring access to end of life care for ethnic minorities with end stage kidney disease through recruitment in action research. BMC Palliative Care. 2016; 15:57

25. Wilkinson E, Randhawa G, Brown EA, Da Silva Gane M, Stoves J, Warwick G et al. Communication as care at end of life: an emerging issue from an exploratory action research study of renal end-of-life care for ethnic minorities in the UK. J Ren Care. 2014; 40 Suppl 1:23-9

26. Wye L, Lasseter G, Percival J, Duncan L, Simmonds B, Purdy S. What works in 'real life' to facilitate home deaths and fewer hospital admissions for those at end of life?:

Results from a realist evaluation of new palliative care services in two English counties. BMC Palliative Care. 2014; 13:37