

## Neonatal parenteral nutrition

### Supplementary material C - Methods

*NICE guideline tbc*

*Supplementary material C*

*September 2019*

*Draft for consultation*

*These evidence reviews were developed by  
the National Guideline Alliance which is part of  
the Royal College of Obstetricians and  
Gynaecologists*



## **Disclaimer**

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

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

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 All NICE guidance is subject to regular review and may be updated or withdrawn.

## **Copyright**

© NICE 2019. All rights reserved. Subject to [Notice of Rights](#).

ISBN:

## Contents

<b>Development of the guideline.....</b>	<b>5</b>
Remit.....	5
What this guideline covers.....	5
Groups that are covered.....	5
Clinical areas that are covered .....	5
<b>Methods .....</b>	<b>6</b>
Introduction .....	6
Developing the review questions and outcomes .....	6
Searching for evidence .....	9
Clinical literature search .....	9
Health economic literature search .....	9
Reviewing evidence.....	10
Systematic review process .....	10
Type of studies and inclusion/exclusion criteria .....	10
Methods of combining evidence .....	11
Appraising the quality of evidence .....	13
Formal consensus reviews .....	21
Collaboration with Cochrane (rapid update of a lipid formulation review – D4) ....	23
Evidence statements .....	24
Economic evidence .....	24
Reviewing economic evidence .....	24
Health economic modelling.....	25
Cost effectiveness criteria .....	26
Developing recommendations .....	26
Guideline recommendations.....	26
Research recommendations.....	26
Validation process .....	26
Updating the guideline.....	27
Funding .....	27
References.....	27

# 1 Development of the guideline

## 2 Remit

- 3 The National Institute for Health and Care Excellence (NICE) commissioned the National  
4 Guideline Alliance (NGA) to develop a new guideline for the management of parenteral  
5 nutrition in neonates.

## 6 What this guideline covers

### 7 Groups that are covered

- 8 • Babies born preterm, up to 28 days after their due birth date (preterm babies)  
9 • Babies born at term, up to 28 days after their birth (term babies)
- 10 Specific consideration will be given to those babies who:
- 11 • Are critically ill **or**  
12 • Need surgery

### 13 Clinical areas that are covered

- 14 1. Indications for, and approaches to, starting parenteral nutrition in preterm and term babies  
15 2. Energy needs of preterm and term babies receiving parenteral nutrition  
16 3. Individual constituents in parenteral nutrition for preterm and term babies:  
17     o macronutrients (amino acids, carbohydrates and lipids)  
18     o minerals and iron  
19     o chloride and acetate balance.  
20 4. Venous access for parenteral nutrition in preterm and term babies  
21 5. Monitoring parenteral nutrition in preterm and term babies  
22 6. Stopping parenteral nutrition in preterm and term babies  
23 7. Service design  
24 8. Information and support for parents and carers

# 1 Methods

## 2 Introduction

3 This section summarises methods used to identify and review the evidence, to consider cost  
4 effectiveness, and to develop guideline recommendations. This guideline was developed in  
5 accordance with methods described in [Developing NICE guidelines: the manual](#) (NICE  
6 2014). A more up to date version is now available of the NICE manual; however  
7 development of this guideline was initiated before publication of the new manual and to  
8 ensure consistent methods the 2014 version was used throughout.

9 Until March 2018, declarations of interest were recorded and managed in accordance with  
10 NICE's 2014 conflicts of interest policy. From April 2018, declarations were recorded and  
11 managed in accordance with NICE's 2018 [Policy on declaring and managing interests for](#)  
12 [NICE advisory committees](#).

## 13 Developing the review questions and outcomes

14 The 23 review questions considered in this guideline were based on the key areas identified  
15 in the guideline [scope](#). They were drafted by the NGA technical team, and refined and  
16 validated by the guideline committee (see Table 1). The review questions were based on the  
17 following frameworks:

- 18 • intervention reviews – using population, intervention, comparison and outcome (PICO)
- 19 • prognostic reviews – population, prognostic factor, outcome (PPO)
- 20 • qualitative reviews – using population, phenomenon of interest and context (PICo).
- 21 • formal consensus – using the nominal group consensus technique.

22 These frameworks guided the development of review protocols, the literature searching  
23 process, and critical appraisal and synthesis of evidence. They also facilitated development  
24 of recommendations by the committee.

25 Full literature searches, critical appraisal and evidence reviews were completed for all review  
26 questions.

27 The review questions and corresponding review types for each evidence report are provided  
28 in Table 1.

29 **Table 1: Summary of review questions and index to evidence reports**

Chapter/Evidence report	Subtopic	Review question	Type of review
[A1] Indications for, and approaches to, starting parenteral nutrition in preterm and term babies: predictors of feeding success	Predictors of feeding success	• What are the predictors for enteral feeding success?	Prognostic
[A2] Indications for, and approaches to, starting parenteral nutrition in preterm and term babies:	Optimal timeframe	• For those neonates where parenteral nutrition is required, what is the optimal timeframe for doing this?	Intervention

Chapter/Evidence report	Subtopic	Review question	Type of review
optimal timeframe to starting PN			
[B] Venous access	Venous access	<ul style="list-style-type: none"> <li>• What overall osmolality, concentration of calcium, and glucose/dextrose in parenteral nutrition can determine whether to administer centrally or peripherally?</li> </ul>	Intervention
[C] Energy needs of preterm and term babies receiving parenteral nutrition	Energy needs	<ul style="list-style-type: none"> <li>• How many kcal/kg/day should be given to preterm and term babies receiving parenteral nutrition?</li> </ul>	Intervention
[D1] Individual constituents in parenteral nutrition for preterm and term babies	Constituents of neonatal parenteral nutrition - glucose	<ul style="list-style-type: none"> <li>• What is the optimal target dose for carbohydrates in preterm and term babies who are receiving parenteral nutrition?</li> <li>• What is the optimal way (starting dose and approach to increment, if employed) to achieve this target dose of carbohydrates?</li> </ul>	Intervention
[D2] Individual constituents in parenteral nutrition for preterm and term babies	Constituents of neonatal parenteral nutrition – amino acids	<ul style="list-style-type: none"> <li>• What is the optimal target dose for amino acids in preterm and term babies who are receiving parenteral nutrition?</li> <li>• What is the optimal way (starting dose and approach to increment, if employed) to achieve this target dose of amino acids?</li> </ul>	Intervention
[D3] Individual constituents in parenteral nutrition for preterm and term babies	Constituents of neonatal parenteral nutrition - lipids	<ul style="list-style-type: none"> <li>• What is the optimal target dose for lipids in preterm and term babies who are receiving parenteral nutrition?</li> <li>• What is the optimal way (starting dose and approach to increment, if employed) to achieve this target dose of lipid?</li> </ul>	Intervention
[D4] Individual constituents in parenteral nutrition for preterm and term babies	Constituents of neonatal parenteral nutrition -lipid emulsions	<ul style="list-style-type: none"> <li>• What is the clinical effectiveness, efficacy and safety of lipid formulations from different sources (for example soya, fish oil, or mixed sources)?</li> </ul>	Intervention
[D5] Individual constituents in parenteral nutrition for preterm and term babies	Constituents of neonatal parenteral nutrition - Iron	<ul style="list-style-type: none"> <li>• What is the effective and safe iron supplementation compared to no iron or different dosage/formulations in preterm and term babies who are receiving parenteral nutrition?</li> </ul>	Intervention

Chapter/Evidence report	Subtopic	Review question	Type of review
[D6] Individual constituents in parenteral nutrition for preterm and term babies	Constituents of neonatal parenteral nutrition - Acetate	<ul style="list-style-type: none"> <li>How much (if any) intravenous acetate should be provided to preterm and term babies who are receiving parenteral nutrition?</li> </ul>	Intervention
[D7] Individual constituents in parenteral nutrition for preterm and term babies	Constituents of neonatal parenteral nutrition – relative amounts of non-nitrogen energy to nitrogen	<ul style="list-style-type: none"> <li>What are the most effective relative amounts of nitrogen and non-nitrogen energy?</li> </ul>	Intervention
[D8] Individual constituents in parenteral nutrition for preterm and term babies	Constituents of neonatal parenteral nutrition – relative amounts of carbohydrates and lipids	<ul style="list-style-type: none"> <li>What are the most effective relative amounts of carbohydrates and lipids?</li> </ul>	Intervention
[D9] Individual constituents in parenteral nutrition for preterm and term babies	Constituents of neonatal parenteral nutrition - Calcium and phosphate	<ul style="list-style-type: none"> <li>What are the optimal target dosages for calcium and phosphate in preterm and term babies receiving parenteral nutrition?</li> </ul>	Intervention
[D10] Individual constituents in parenteral nutrition for preterm and term babies	Relative amounts amino acids to phosphate	<ul style="list-style-type: none"> <li>What is the optimal ratio of phosphate to amino acid in preterm and term babies receiving parenteral nutrition?</li> </ul>	Intervention
[E] Standardised neonatal parenteral nutrition	Standardised bags	<ul style="list-style-type: none"> <li>What is the effectiveness, efficacy and safety of standardised parenteral nutrition bags as compared to individualised bags?</li> </ul>	Intervention
[F] Monitoring neonatal parenteral nutrition	Monitoring	<ul style="list-style-type: none"> <li>In babies on parenteral nutrition, what is the optimal frequency of blood sampling for monitoring glucose, calcium, phosphate, potassium and serum triglycerides?</li> </ul>	Intervention
[G] Stopping neonatal parenteral nutrition	Stopping PN	<ul style="list-style-type: none"> <li>What amount of enteral nutrition (measured in terms of ml/kg/day or kcal/kg/day) should be given when starting enteral feeds?</li> </ul>	Intervention
[H] Service design	Service design	<ul style="list-style-type: none"> <li>Are nutrition care teams (pharmacists and dietician) effective and safe in providing parenteral nutrition in preterm and term babies?</li> </ul>	Intervention
[I] Information and support	Information and support	<ul style="list-style-type: none"> <li>What are the most effective methods of information provision about parenteral nutrition, and what information and support to parents/carers perceive as useful?</li> </ul>	Mixed methods-intervention (quantitative) and qualitative

Chapter/Evidence report	Subtopic	Review question	Type of review
[J] General principles of neonatal parenteral nutrition	General principles	<ul style="list-style-type: none"><li>• What are the general principles of neonatal parenteral nutrition?</li></ul>	Formal nominal group consensus

- 1 Additional information related to development of the guideline is contained in:
- 2
  - Supplementary material A (NGA staff list)
- 3
  - Supplementary material B (Glossary and abbreviations)
- 4
  - Supplementary material D (Economic study selection).

## 5 Searching for evidence

### 6 Clinical literature search

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

9 Databases were searched using medical subject headings, free-text terms and study type  
10 filters where appropriate. Where possible, searches were restricted to retrieve articles  
11 published in English. All searches were conducted in the following databases: Medline,  
12 Medline-in-Process, Cochrane Central Register of Controlled Trials (CCTR), Cochrane  
13 Database of Systematic Reviews (CDSR), Database of Abstracts of Reviews of Effects  
14 (DARE), Health Technology Assessments (HTA) and Embase. All searches, except the  
15 searches for 9.1 (the General principles chapter), were updated at least 6–8 weeks in  
16 advance of the final guideline committee meetings before consultation on the draft guideline;  
17 these updates were completed during April 2019. Any studies added to the databases after  
18 April 2019 (including those published before April 2019 but not yet indexed) were not  
19 considered for inclusion. The searches for 9.1 (the General Principles chapter), were not  
20 rerun as the evidence included for the formal consensus exercise were published guidelines,  
21 and the committee were not aware of any recently published guidelines which had been  
22 released following the undertaking of this review. .

23 Search strategies were quality assured by cross-checking reference lists of relevant articles,  
24 analysing search strategies from other systematic reviews and asking members of the  
25 committee to highlight key studies. All search strategies were also quality assured by an  
26 information scientist who was not involved in developing the primary search strategy. Details  
27 of the search strategies, including study-design filters applied and databases searched, are  
28 presented in Appendix B of each evidence report.

29 All publications highlighted by stakeholders at the time of the consultation on the draft scope  
30 were considered for inclusion. During the scoping phase, searches were conducted for  
31 guidelines, health technology assessments, systematic reviews, economic evaluations and  
32 reports on biomedical databases and websites of organisations relevant to the topic. Formal  
33 searching for grey literature and unpublished literature was not undertaken.

### 34 Health economic literature search

35 Systematic literature searches were also undertaken to identify published health economic  
36 evidence. A broad search was conducted to identify health economic evidence related to  
37 parenteral nutrition in neonates in the following databases: NHS Economic Evaluation  
38 Database (NHS EED) and HTA. A broad search was also conducted to identify health  
39 economic evidence related to parenteral nutrition in neonates in the following databases with  
40 an economic search filter applied: Medline, CCTR and Embase. Where possible, the

1 searches were restricted to retrieve articles published in English; studies published in  
2 languages other than English were not eligible for inclusion.

3 The search strategies for the health economic literature search are included in Supplement D  
4 (Health economics). As for the clinical literature searches, economic literature searches were  
5 updated at least 6–8 weeks in advance of the final committee meeting before consultation on  
6 the draft guideline; these updates were completed during April 2019.

## 7 **Reviewing evidence**

### 8 **Systematic review process**

9 The evidence was reviewed in accordance with the following approach.

- 10 • Potentially relevant articles were identified from the search results for each review  
11 question by screening titles and abstracts. Full-text copies of the articles were then  
12 obtained.
- 13 • Full-text articles were reviewed against pre-specified inclusion and exclusion criteria in the  
14 review protocol (see appendix A of each evidence report).
- 15 • Key information was extracted from each article on study methods and results, in  
16 accordance with factors specified in the review protocol. The information was presented in  
17 a summary table in the corresponding evidence report and in a more detailed evidence  
18 table (see appendix D of each evidence report).
- 19 • Included studies were critically appraised using an appropriate checklist as specified in  
20 [Developing NICE guidelines: the manual](#) (NICE 2014). Further detail on appraisal of the  
21 evidence is provided below.
- 22 • Summaries of evidence by outcome were presented in the corresponding evidence report  
23 and discussed by the committee.
- 24 • Results were summarised and reported in GRADE profiles (for intervention reviews), or  
25 reported in evidence tables including the risk of bias assessment for prognostic reviews.

26 Review questions selected as high priorities for health economic analysis (and those  
27 selected as medium priorities and where health economic analysis could influence  
28 recommendations) were subject to dual screening and study selection through a 10%  
29 random sample of articles. Any discrepancies were resolved by discussion between the first  
30 and second reviewers or by reference to a third (senior) reviewer. For the remaining review  
31 questions, internal (NGA) quality assurance processes included consideration of the  
32 outcomes of screening, study selection and data extraction and the committee reviewed the  
33 results of study selection and data extraction. The review protocol for each question specifies  
34 whether dual screening and study selection was undertaken for that particular question.

35 Drafts of all evidence reviews were checked by a senior reviewer.

### 36 **Type of studies and inclusion/exclusion criteria**

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

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

41 For intervention reviews, randomised controlled trials (RCTs) were prioritised for inclusion  
42 because they are considered to be the most robust type of study design that could produce  
43 an unbiased estimate of intervention effects. Where there was limited evidence from RCT,  
44 non-randomised controlled trials and/or observational studies were considered for inclusion,

- 1 including cohort studies, case–control studies, cross-sectional studies and case series, and  
2 this was specified in each evidence review.
- 3 For prognostic reviews, prospective and retrospective cohort and case–control studies and  
4 case series were considered for inclusion.
- 5 For qualitative reviews, studies using focus groups, structured interviews or semi-structured  
6 interviews were considered for inclusion. Where qualitative evidence was sought, data from  
7 surveys or other types of questionnaire were considered for inclusion only if they provided  
8 data from open-ended questions, but not if they reported only quantitative data.
- 9 For the formal consensus review, published guidelines were included (see formal consensus  
10 review section).
- 11 The committee was consulted about any uncertainty regarding inclusion or exclusion of  
12 studies. A list of excluded studies for each review question, including reasons for exclusion,  
13 is presented in appendix K of the corresponding evidence report.
- 14 Narrative reviews, posters, letters, editorials, comment articles, unpublished studies and  
15 studies published in languages other than English were excluded. Conference abstracts  
16 were not considered for inclusion (see the review protocols for details).

## 17 **Methods of combining evidence**

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

## 20 **Data synthesis for intervention reviews**

- 21 Meta-analysis to pool results was conducted where possible using Cochrane Review  
22 Manager (RevMan5) software. The potential for bias and inconsistency across studies is  
23 higher in observational studies compared with RCTs. Therefore, evidence from RCTs and  
24 observational studies was analysed and presented to the committee separately. Differences  
25 and similarities between RCT and observational evidence were considered and discussed by  
26 the committee when making recommendations.
- 27 For dichotomous outcomes, such as mortality, the Mantel–Haenszel method with a fixed  
28 effect model was used to calculate risk ratios (RRs) with 95% confidence intervals (CIs). For  
29 all outcomes with zero events in both arms the risk difference (RD) was calculated and  
30 presented. For outcomes with zero events in only one arm, Peto odds ratios (PORs) was  
31 calculated as this method performs well when events are rare (Bradburn 2007). Both POR  
32 and RD were presented.
- 33 For continuous outcomes, measures of central tendency (mean) and variation (standard  
34 deviation; SD) are required for meta-analysis. Data for continuous outcomes, such as  
35 duration of hospital stay, were meta-analysed using an inverse-variance method for pooling  
36 weighted mean differences (WMDs). If a study reported only a summary statistic for a  
37 continuous outcome and a measure of variability (mean difference and 95% confidence  
38 interval (CI) or the mean difference and its SD or standard error) the generic-inverse  
39 variance method was used to enter data and conduct the meta-analysis.
- 40 Subgroups for stratified analyses were agreed a priori for some review questions as part of  
41 protocol development.
- 42 When meta-analysis was undertaken, the results were presented visually using forest plots  
43 generated using RevMan5 (see appendix E of relevant evidence reports). Forest plots were  
44 not presented for outcomes reported by single studies.

1 Statistical heterogeneity was assessed by visually examining forest plots and calculating the  
2 I-squared inconsistency statistic (with an I-squared value of more than 50% indicating  
3 considerable heterogeneity, and more than 80% indicating very serious heterogeneity).  
4 When considerable heterogeneity was present, the meta-analysis was re-run using the Der-  
5 Simonian and Laird method with a random effects model and the results from the two models  
6 were compared. In addition, predefined subgroup analyses were performed where possible.  
7 In the case of unexplained heterogeneity, sensitivity analyses were planned based on the  
8 quality of studies, eliminating studies at high risk of bias (in relation to randomisation,  
9 allocation concealment and blinding, and/or missing outcome data). In cases where there  
10 was no plausible explanation for the heterogeneity a random effects meta-analysis model  
11 was used.

## 12 **Data synthesis for prognostic reviews**

13 Odds ratios or RRs with 95% CIs reported in published studies would be extracted or  
14 calculated by the NGA technical team to examine relationships between a given factor and  
15 the outcome of interest. Analyses adjusting for key confounders (such as gestational age)  
16 would have been included in preference to unadjusted analyses. Recognising variation  
17 across studies in terms of populations, risk factors, outcomes and statistical analysis  
18 methods (including adjustments for confounding factors), prognostic data would not usually  
19 be pooled, but results from individual studies would be presented in the evidence reports.

20 None of these methods were used because no evidence was identified for this topic.

## 21 **Data synthesis for qualitative reviews**

22 Where possible, a meta-synthesis was going to be conducted to combine evidence from  
23 qualitative studies. The main aim of qualitative data synthesis in this guideline was to  
24 determine what information parent or carers of preterm or term babies may receive, and how  
25 this may influence satisfaction with care, uptake of services and anxiety levels. The aim was  
26 also to determine the views of healthcare professionals regarding the provision of information  
27 to parents and carers. Whenever studies identified a qualitative theme, this would be  
28 extracted and the main characteristics would be summarised. When all themes would be  
29 extracted from studies, common concepts would be categorised and tabulated. This would  
30 include information on how many studies had contributed to each theme identified by the  
31 NGA technical team.

32 In qualitative synthesis, a theme being reported more than other themes across included  
33 studies does not necessarily mean that the theme is more important than other themes. The  
34 aim of qualitative research is to identify new perspectives on a particular topic. Study types  
35 and populations in qualitative research can differ widely, meaning that themes identified by  
36 just one or a few studies can provide important new information on a given topic. Therefore,  
37 for the purpose of the qualitative reviews in this guideline, it was planned that further studies  
38 would not be added when they reported only the same themes as had already been  
39 identified from other studies because the emphasis was to be on conceptual robustness  
40 rather than quantitative completeness of the evidence. This would have implications for the  
41 types and numbers of studies included in the qualitative reviews, with study inclusion  
42 continuing until no new relevant data could be found regarding a topic that would add to or  
43 refute it. This concept is referred to in the literature as 'theoretical saturation' (Dixon-Woods  
44 2005). However, there was no evidence available for the qualitative review considered in this  
45 guideline, and so the methods for managing data saturation were not applied.

46 Themes from individual studies would be integrated into a wider context and overarching  
47 categories of themes with sub-themes may be identified, for example if there was a theme of  
48 'logistical support' this may link into a wider theme of 'satisfaction with overall care'. Themes  
49 would be derived from data presented in individual studies. When themes are extracted from  
50 one primary study only, theme names used in the guideline mirrored those in the source

- 1 study. However, when themes would be based on evidence from multiple studies, the theme  
2 names would be assigned by the NGA technical team. The names of overarching categories  
3 of themes would also be assigned by the NGA technical team.
- 4 Emerging themes would be placed into a thematic map. The purpose of such a map is to  
5 show relationships between overarching categories and associated themes.
- 6 None of these methods were used because no evidence was identified for this topic.

## 7 Appraising the quality of evidence

### 8 Intervention studies

#### 9 *GRADE methodology for intervention reviews*

10 For intervention reviews, the evidence for outcomes from included RCTs and comparative  
11 observational studies was evaluated and presented using the Grading of Recommendations  
12 Assessment, Development and Evaluation (GRADE) methodology developed by the  
13 international [GRADE working group](#). More information about this tool can be found on the  
14 [developer's website](#).

15 When GRADE was applied, software developed by the GRADE working group (GRADEpro)  
16 was used to assess the quality of each outcome, taking account of individual study quality  
17 factors and any meta-analysis results. Results were presented in GRADE profiles (GRADE  
18 tables). The clinical evidence profile tables include details of the quality assessment and  
19 pooled outcome data, where appropriate, a relative and an absolute measure of intervention  
20 effect and the summary of quality of evidence for that outcome. In this table, the columns for  
21 intervention and control indicate the sum across studies of the number of participants in each  
22 arm for continuous outcomes and frequency of events (n/N; the sum across studies of the  
23 number of participants with events divided by sum of the number of participants) for  
24 dichotomous outcomes.

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

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

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

Quality element	Description
Risk of bias ('Study limitations')	Limitations in study design and implementation may bias estimates of treatment effect. High risk of bias for the majority of the evidence reduces confidence in the estimated effect
Inconsistency	This refers to unexplained heterogeneity in the results

Quality element	Description
Indirectness	This refers to differences in study populations, interventions, comparators or outcomes between the available evidence and inclusion criteria specified in the review protocol
Imprecision	This occurs when a study has relatively few participants or few events of interest, resulting in wide confidence intervals around estimates of effect that include clinically important thresholds
Publication bias	This refers to systematic under- or over-estimation of the underlying benefit or harm resulting from selective publication of study results

1 **Table 3: GRADE quality ratings (by quality element)**

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

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

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

3 *Assessing risk of bias in intervention reviews*

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

6 Risk of bias in RCTs was assessed using the Cochrane risk of bias tool (see appendix H in  
 7 [Developing NICE guidelines: the manual](#); NICE 2014).

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

- 9 • selection bias
- 10 • performance bias
- 11 • attrition bias
- 12 • detection bias
- 13 • reporting bias.

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

16 For systematic reviews of RCT the ROBIS checklist was used (see appendix H in [Developing  
 17 NICE guidelines: the manual](#); NICE 2014).

1 For observational studies the Cochrane risk of bias for non-randomised studies (ROBINS-I)  
2 was used (see appendix H in [Developing NICE guidelines: the manual](#); NICE 2014).

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

#### 6 *Assessing inconsistency in intervention reviews*

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

14 Inconsistency was assessed visually by inspecting forest plots and observing whether there  
15 was considerable heterogeneity in the results of the meta-analysis. Statistical heterogeneity  
16 was assessed by calculating the I-squared statistic for the meta-analysis with an I-squared  
17 value of more than 50% indicating considerable heterogeneity, and more than 80% indicating  
18 very serious heterogeneity. When considerable or very serious heterogeneity was observed,  
19 possible reasons were explored using sensitivity and subgroup analyses as pre-specified in  
20 the review protocol where possible. Sensitivity analyses were planned based on the quality  
21 of studies, eliminating studies at high risk of bias (in relation to randomisation, allocation  
22 concealment and blinding, and/or missing outcome data).

23 When no plausible explanation for the heterogeneity could be found, the quality of the  
24 evidence was downgraded in GRADE for inconsistency and a random effects meta-analysis  
25 was presented.

#### 26 *Assessing indirectness in intervention reviews*

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

#### 33 *Assessing imprecision and clinical importance in intervention reviews*

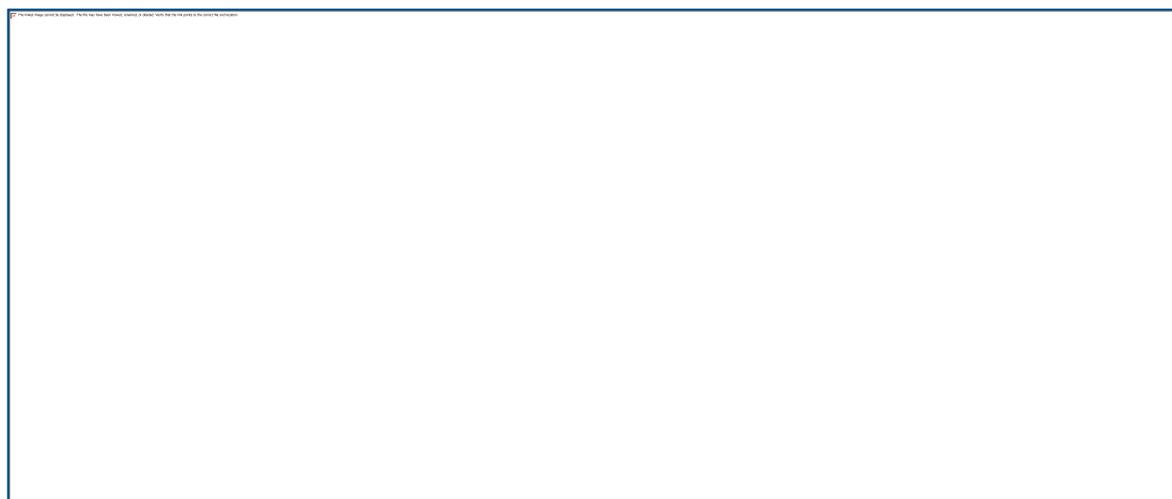
34 Imprecision in GRADE methodology refers to uncertainty around the effect estimate and  
35 whether or not there is a clinically important difference between interventions (that is,  
36 whether the evidence clearly supports a particular recommendation or appears to be  
37 consistent with several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

42 The 95% CI is defined as the range of values within which the population value will fall on  
43 95% of repeated samples, were the procedure to be repeated. The larger the study, the  
44 smaller the 95% CI will be and the more certain the effect estimate.

45 Imprecision was assessed in the guideline evidence reviews by considering whether the  
46 width of the 95% CI of the effect estimate was relevant to decision making, considering each  
47 outcome independently. This is illustrated in Figure 1, which considers a positive outcome for  
48 the comparison of treatment 'A' versus treatment 'B'. Three decision-making zones can be

- 1 differentiated, bounded by the thresholds for clinical importance (minimally important  
2 differences; MIDs) for benefit and harm. The MID for harm for a positive outcome means the  
3 threshold at which treatment A is less effective than treatment B by an amount that is  
4 clinically important (favours B).
- 5 When the CI of the effect estimate is wholly contained in 1 of the 3 zones there is no  
6 uncertainty about the size and direction of effect, therefore, the effect estimate is considered  
7 precise; that is, there is no imprecision.
- 8 When the CI crosses 2 zones, it is uncertain in which zone the true value of the effect  
9 estimate lies and therefore there is uncertainty over which decision to make. The CI is  
10 consistent with 2 possible decisions, therefore, the effect estimate is considered to be  
11 imprecise in the GRADE analysis and the evidence is downgraded by 1 level ('serious  
12 imprecision').
- 13 When the CI crosses all 3 zones, the effect estimate is considered to be very imprecise  
14 because the CI is consistent with 3 possible clinical decisions and there is therefore a  
15 considerable lack of confidence in the results. The evidence is therefore downgraded by 2  
16 levels in the GRADE analysis ('very serious imprecision').
- 17 Implicitly, assessing whether a CI is in, or partially in, a clinically important zone, requires the  
18 guideline committee to estimate an MID or to say whether they would make different  
19 decisions for the 2 confidence limits.
- 20 For outcomes with zero events in both arms, and where there were zero events in only one  
21 arm, the risk difference was calculated. However, there was no agreement on an equivalent  
22 to an MID for these cases. Due to the low event rate and usually associated wide CIs it was  
23 decided to downgrade these outcomes to 'serious' imprecision to prevent quality inflation.

24 **Figure 1: Assessment of imprecision and clinical importance in intervention reviews**  
25 **using GRADE**



26  
27 *MID, minimally important difference*

28 *Defining minimally important differences for intervention reviews*

29 The committee was asked whether there were any recognised or acceptable MIDs in the  
30 clinical literature and community relevant to the review questions under consideration. The  
31 committee was not aware of any MIDs that could be used for the guideline.

32 In the absence of published or accepted MIDs, the committee agreed to use the GRADE  
33 default MIDs to assess imprecision. For dichotomous outcomes, RRs of 0.80 and 1.25 were  
34 considered the clinically important thresholds and were used as default MIDs in the

1 guideline. The same thresholds were used as default MIDs in the guideline for all  
 2 dichotomous outcomes considered in intervention evidence reviews. For all outcomes with 0  
 3 events in both arms, or where there were 0 events in only one arm, the Peto odds ratio and  
 4 risk difference was calculated. However, there was no agreement on the equivalent to an  
 5 MID for these cases. Due to the low event rate which is usually associated with wide  
 6 confidence intervals, it was decided to give those cases a 'serious' imprecision rating to  
 7 prevent quality inflation for these outcomes.

8 For continuous outcomes, default MIDs are equal to half the mean SD of the control groups  
 9 at baseline, or at follow-up if the SD is not available a baseline. If within one review question  
 10 some outcomes have baseline SDs, yet other outcomes do not, then all MIDs will be  
 11 calculated using follow-up SDs to ensure consistency.

## 12 Prognostic studies

### 13 *Adapted GRADE methodology for prognostic reviews*

14 For prognostic reviews, with evidence from comparative observational studies, an adapted  
 15 GRADE approach was used. As noted above, GRADE methodology is designed for  
 16 intervention reviews but the quality assessment elements of GRADE were adapted by the  
 17 technical team for prognostic reviews. Adapted GRADE was not used for evidence from case  
 18 series; instead quality of case series evidence was assessed using the Checklist for Case  
 19 Series developed by the Joanna Briggs Institute. More information about this tool can be  
 20 found on the [developer's website](#).

21 The evidence for each outcome in the prognostic reviews was examined separately for the  
 22 quality elements listed and defined in Table 5. The criteria considered in the rating of these  
 23 elements are discussed below. Each element was graded using the quality levels  
 24 summarised in Table 3. Footnotes to GRADE tables were used to record reasons for grading  
 25 a particular quality element as having 'serious' or 'very serious' quality issues. The ratings for  
 26 each component were combined to obtain an overall assessment of quality for each outcome  
 27 as described in Table 4.

28 **Table 5: Adaptation of GRADE quality elements for prognostic reviews**

Quality element	Description
Risk of bias ('Study limitations')	Limitations in study design and implementation may bias estimates and interpretation of the effect of the prognostic/risk factor. High risk of bias for the majority of the evidence reduces confidence in the estimated effect. Prognostic studies are not usually randomised and therefore would not be downgraded for study design from the outset (they start as high quality)
Inconsistency	This refers to unexplained heterogeneity between studies looking at the same prognostic/risk factor, resulting in wide variability in estimates of association (such as RRs or ORs), with little or no overlap in confidence intervals
Indirectness	This refers to any departure from inclusion criteria listed in the review protocol (such as differences in study populations or prognostic/risk factors), that may affect the generalisability of results
Imprecision	This occurs when a study has relatively few participants and also when the number of participants is too small for a multivariable analysis (as a rule of thumb, 10 participants are needed per variable). This was assessed by considering the confidence interval in relation to the point estimate for each outcome reported in the included studies

29 *RR, relative risk; OR, odds ratio*

1 *Assessing risk of bias in prognostic reviews*

2 The Quality in Prognosis Studies (QUIPS) tool developed by Hayden 2013 was used to  
3 assess risk of bias in studies included in prognostic reviews (see appendix H in [Developing](#)  
4 [NICE guidelines: the manual](#); NICE 2014). The risk of bias in each study was determined by  
5 assessing the following domains:

- 6 • selection bias
- 7 • attrition bias
- 8 • prognostic factor bias
- 9 • outcome measurement bias
- 10 • control for confounders
- 11 • appropriate statistical analysis.

12 *Assessing inconsistency in prognostic reviews*

13 No meta-analysis was performed for prognostic reviews in this guideline. 'No serious  
14 inconsistency' was nevertheless used to describe this quality assessment in the GRADE  
15 tables for outcomes from single studies.

16 *Assessing indirectness in prognostic reviews*

17 Indirectness in prognostic reviews was assessed by comparing the populations, prognostic  
18 factors and outcomes in the evidence to those defined in the review protocol.

19 *Assessing imprecision and clinical importance in prognostic reviews*

20 Prognostic studies may have a variety of purposes, for example, establishing typical  
21 prognosis in a broad population, establishing the effect of patient characteristics on  
22 prognosis, and developing a prognostic model. While by definition MIDs relate to treatment  
23 effects, the committee agreed to use GRADE default MIDs for intervention studies as a  
24 starting point from which to assess whether the size of an outcome effect in a prognostic  
25 study would be large enough to be meaningful in clinical practice.

## 26 Qualitative reviews

### 27 **GRADE-CERQual methodology for qualitative reviews**

28 For qualitative reviews the GRADE Confidence in the Evidence from Reviews of Qualitative  
29 research (GRADE-CERQual) approach (Lewin 2015) was used. In this approach the quality  
30 of evidence is considered according to themes in the evidence. The themes may have been  
31 identified in the primary studies or they may have been identified by considering the reports  
32 of a number of studies. Quality elements assessed using GRADE-CERQual are listed and  
33 defined in Table 6. Each element was graded using the levels of concern summarised in  
34 Table 7. The ratings for each component were combined (as with other types of evidence) to  
35 obtain an overall assessment of quality for each theme as described in Table 8.

### 36 **Table 6: GRADE-CERQual elements for qualitative reviews**

Quality element	Description
Risk of bias ('Methodological limitations')	Limitations in study design and implementation may bias interpretation of qualitative themes identified. High risk of bias for the majority of the evidence reduces confidence in review findings. Qualitative studies are not usually randomised and therefore would not be downgraded for study design from the outset (they start as high quality)
Relevance (or applicability) of evidence	This refers to the extent to which the evidence supporting the review findings is applicable to the context specified in the review question

Quality element	Description
Coherence of findings	This refers to the extent to which review findings are well grounded in data from the contributing primary studies and provide a credible explanation for patterns identified in the evidence
Adequacy of data (theme saturation or sufficiency)	This corresponds to a similar concept in primary qualitative research, that is, whether a theoretical point of theme saturation was achieved, at which point no further citations or observations would provide more insight or suggest a different interpretation of the particular theme. Individual studies that may have contributed to a theme or sub-theme may have been conducted in a manner that by design would have not reached theoretical saturation at an individual study level

1 **Table 7: CERQual levels of concern (by quality element)**

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

2 **Table 8: Overall confidence in the evidence in CERQual (by review finding)**

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

3 *Assessing risk of bias in qualitative reviews*

4 The risk of bias in qualitative studies was assessed using the Critical Appraisal Skills  
5 Programme (CASP) checklist for qualitative studies (see appendix H in [Developing NICE  
6 guidelines: the manual](#); NICE 2014). The overall risk of bias was derived by assessing the  
7 risk of bias across the 6 domains summarised in Table 9.

8 **Table 9: Risk of bias in qualitative studies**

Aim and appropriateness of qualitative evidence	This domain assesses whether the aims and relevance of the study were described clearly and whether qualitative research methods were appropriate for investigating the research question
Rigour in study design or validity of theoretical approach	This domain assesses whether the study approach was documented clearly and whether it was based on a theoretical framework (such as ethnography or grounded theory). This does not necessarily mean that the framework has to be stated explicitly, but a detailed description ensuring transparency and reproducibility should be provided

Sample selection	This domain assesses the background, the procedure and reasons for the method of selecting participants. The assessment should include consideration of any relationship between the researcher and the participants, and how this might have influenced the findings
Data collection	This domain assesses the documentation of the method of data collection (in-depth interviews, semi-structured interviews, focus groups or observations). It also assesses who conducted any interviews, how long they lasted and where they took place
Data analysis	This domain assesses whether sufficient detail was documented for the analytical process and whether it was in accordance with the theoretical approach. For example, if a thematic analysis was used, the assessment would focus on the description of the approach used to generate themes. Consideration of data saturation would also form part of this assessment (it could be reported directly or it might be inferred from the citations documented that more themes could be found)
Results	This domain assesses any reasoning accompanying reporting of results (for example, whether a theoretical proposal or framework is provided)

1 *Assessing relevance of evidence in qualitative reviews*

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

5 *Assessing coherence of findings in qualitative reviews*

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

16 *Assessing adequacy of data in qualitative reviews*

17 Adequacy of data (theme saturation or sufficiency) corresponds to a similar concept in  
 18 primary qualitative research in which consideration is made of whether a theoretical point of  
 19 theme saturation was achieved, meaning that no further citations or observations would  
 20 provide more insight or suggest a different interpretation of the theme concerned. As noted  
 21 above, it is not equivalent to the number of studies contributing to a theme, but rather to the  
 22 depth of evidence and whether sufficient quotations or observations were provided to  
 23 underpin the findings.

1 *Assessing clinical importance in qualitative reviews*

2 For themes stemming from qualitative findings, clinical importance was agreed by the  
3 committee taking account of the generalisability of the context from which the theme was  
4 derived and whether it was sufficiently convincing to support or warrant a change in current  
5 practice, as well as the quality of the evidence.

**6 Formal consensus reviews**

7 Formal consensus was carried out using the nominal group technique (Murphy 1998) for  
8 evidence review J (General principles). This is a structured method focusing on the opinions  
9 of individuals within a group. Due to this focus on individuals it is referred to as a 'nominal  
10 group' technique. It usually involves anonymous voting with an opportunity to provide  
11 comments. It is usually conducted by an iterative process in which options with low  
12 agreement are eliminated and options with high agreement are retained. Using the  
13 comments that individuals provided, options with medium agreement are revised and then  
14 considered in a second round.

**15 Topics addressed with this method (evidence report J - general principles)**

16 It was agreed that clinical evidence reviews would not be the most appropriate for topics  
17 where administration of care would be guided by physiological, pathophysiological and  
18 clinical principles; for example, vitamins, minerals, fluid volume and electrolytes. However,  
19 guidance on parenteral nutrition would not be complete without including general  
20 recommendations about these constituents. Therefore formal consensus was agreed as the  
21 most appropriate method to review this evidence. Agreed topics to be covered by formal  
22 consensus were:

- 23 • overall level of included vitamins
- 24 • general practice for fluid volume
- 25 • overall levels of blood and urinary electrolytes
- 26 • overall level of included minerals
- 27 • overall level of included trace elements
- 28 • delivery of lipids via syringe or bags
- 29 • filtration and protection from light.

**30 Details of the nominal group technique as used in this guideline**

31 A search was conducted for published guidelines on neonatal parenteral nutrition. Only  
32 international, national or regional guidelines/standard protocols were included, it was agreed  
33 that local guidelines would likely be based on these, lack wider applicability, and may not be  
34 generalisable; therefore, local guidelines were excluded. In order to identify most relevant  
35 literature, only those published in the last 10 years (January 2008 onwards) were included.  
36 All potentially relevant guidelines were identified and were assessed for quality using the  
37 Appraisal of Guidelines for Research and Evaluation (AGREE II) instrument (see assessing  
38 quality below).

39 Once the guidelines had been assessed and rated as high quality the NGA technical team  
40 extracted relevant recommendations from these guidelines and derived a set of statements  
41 for all included topics; guidelines rated as low quality were excluded (scores of below 70% in  
42 any two domains of the AGREE II instrument - see assessing quality of guidelines below).

43 All statements were checked for clinical content by the NGA clinical advisor and the  
44 committee chair. If no recommendations existed within the included guidelines for a particular  
45 review area then no statement was produced. The derived statements can be seen in the  
46 relevant evidence review (evidence review J – general principles).

1 The formal consensus exercise was conducted over two committee meetings. At the initial  
2 meeting the statements were presented to the committee in a questionnaire format (this  
3 questionnaire can be seen in the relevant evidence review (evidence review J – general  
4 principles, appendix M). All committee members were invited to take part in the formal  
5 consensus exercise (this did not include the chair as he had been involved in deriving the  
6 statements, nor co-opted members). Committee members were asked to rate each  
7 statement based on their personal opinion of what they believed 'best clinical practice' would  
8 be. The statements were rated using a 9-point Likert scale, where 1 represents 'strongly  
9 disagree', 5 represents 'neither agree nor disagree', and 9 represents 'strongly agree'. The  
10 participants were also able to state that they believed they had insufficient knowledge to  
11 provide a rating. There was also space for written comments about each statement. The  
12 questionnaire was completed anonymously with no prior discussion on the topic. Once this  
13 first round of consensus had been conducted, the NGA technical team calculated overall  
14 percentage agreement for each individual statement. The ratings were grouped into three  
15 categories: 1 to 3 (disagree), 4 to 6 (neither agree nor disagree), or 7 to 9 (agree). If a  
16 committee member indicated they had insufficient knowledge to provide a rating for a  
17 particular statement this was excluded from the calculation of agreement. Statements with  
18 80% or greater agreement were kept, and were to be used to inform recommendations.  
19 Statements with less than 60% agreement were discarded, unless there were obvious and  
20 addressable issues identified from any comments. Those statements with between 60-80%  
21 agreement were re-drafted by the NGA technical team (using the written comments if  
22 provided).

23 The redrafted statements were placed into the same questionnaire format as round 1 of the  
24 formal consensus process. Committee members were sent these revised statements  
25 electronically, and asked to rate them in the same way as in the first round. Responses were  
26 emailed back to the NGA technical team, who calculated agreement as above.

27 At the following committee meeting, statements with 80% or greater agreement (from rounds  
28 1 and 2) were presented as the evidence to inform the development of recommendations.  
29 For each topic (vitamins, fluid volume, electrolytes, minerals, trace elements, delivery of  
30 lipids, filtration and protection from light), the statements with >80% agreement were  
31 presented, the committee then discussed how these statements reflect practice, and how  
32 they could inform or drive best practice. The committee used these statements, along with  
33 their knowledge and experience to develop the recommendations for each topic. Statements  
34 below 80% were discarded in round 2 because they had already been redrafted after  
35 insufficient agreement in round 1 and further redrafting was not expected to lead to a higher  
36 level of agreement in a subsequent round of voting.

### 37 **Assessing quality of guidelines**

38 The identified potentially relevant guidelines were assessed for quality using the Appraisal of  
39 Guidelines for Research and Evaluation ([AGREE II](#)) instrument (Table 10). The tool assesses  
40 6 domains: scope and purpose, stakeholder involvement, rigour of development, clarity of  
41 presentation, applicability and editorial independence.

42 Within each domain there is a set of questions, each of which is scored using a 7-point scale  
43 (1 – 'strongly disagree' to 7 – 'strongly agree'). Each section is rated and then an overall  
44 score for that domain is calculated. Two reviewers independently rated all identified  
45 guidelines using this method (see the [AGREE II](#) for detailed instructions). Only those  
46 guidelines which scored 70% or above in any two domains (and therefore rated as high  
47 quality) were included. Guidelines below this level were excluded.

1 **Table 10: Assessing quality of guidelines**

Domain	Aim of assessment
Scope and purpose	Assesses the aim of the guideline, the relevant health questions, and the population
Stakeholder involvement	Assesses the extent to which the guideline was developed by stakeholders, and whether it represents the views the guidelines intended users
Rigour of development	Assesses the processes used to collect, analyse and synthesise the evidence. It also assesses the methods used to generate and update the recommendations
Clarity of presentation	Assesses the language, and layout of the guideline
Applicability	Assesses potential barriers and facilitators of implementation, uptake of the guideline. It also assesses the potential resource implications.
Editorial independence	Assesses the possibility of the recommendations being biased and assesses potential conflict of interests

2 **Collaboration with Cochrane (rapid update of a lipid formulation review – D4)**

3 During the guideline scoping phase it was identified that Cochrane Neonatal were intending  
 4 to conduct a rapid update of a review on different lipid formulations for parenteral nutrition.  
 5 Cochrane Neonatal received funding from the National Institute of Health Research (NIHR)  
 6 to update their published review to support development of the guideline. It was therefore  
 7 agreed that the NGA technical team as well as the committee would work with Cochrane  
 8 Neonatal to develop the review protocol and would have input into the structure of the  
 9 evidence review to ensure that it would fit the purpose of the guideline. The Cochrane  
 10 Neonatal team also agreed to conduct a second (new rather than updated) review to cover  
 11 term babies. The NGA developed a standard protocol in accordance with NICE processes  
 12 and then the Cochrane team used this to inform their full protocol, ensuring all critical and  
 13 important outcomes as agreed by the committee were included. The draft Cochrane  
 14 Neonatal protocols (developed using Cochrane processes) were then presented at a  
 15 committee meeting to get committee input and sign off. The final review protocols were then  
 16 peer reviewed in accordance with Cochrane processes.

17 Cochrane Neonatal agreed to conduct the review in accordance with the methods described  
 18 above, including GRADE assessment of clinically relevant outcomes. Additionally, the NGA  
 19 technical team reviewed the Cochrane reviews, including assessment of quality using the  
 20 ROBIS tool, (ROBIS: Tool to assess risk of bias in systematic reviews), in line with those  
 21 methods outlined in other sections of this document. The GRADE assessments completed  
 22 by Cochrane Neonatal differed in several ways to those outlined in other sections of this  
 23 document:

- 24 • Optimal information size was considered when judging imprecision:
  - 25 ○ Dichotomous outcomes were downgraded if the total number of events was less than
  - 26 300.
  - 27 ○ Continuous outcomes were downgraded if the total number of participants was less
  - 28 than 400.
- 29 • GRADE default MIDs were not used when judging imprecision:
  - 30 ○ For dichotomous outcomes, RRs of 0.75 and 1.25 were considered the thresholds for
  - 31 appreciable benefit or harm, which is similar to the concept of MIDs, when judging
  - 32 imprecision
  - 33 ○ Continuous outcomes were downgraded for imprecision if the 95% confidence interval
  - 34 cross the null effect and were wide. If the confidence interval was wider than half a
  - 35 standardised mean difference, this was often considered wide. However, this was not

1 an absolute rule and the final decision was made based on consensus among the  
2 Cochrane authors, taking into account the clinical context.

3 • Some outcomes were further downgraded when only a single study contributed to the  
4 effect. This was dependent on sample size and consensus among the Cochrane authors.

5 Despite these minor differences in methods the NGA technical team was confident that it is  
6 unlikely any of the overall decisions would be significantly different. The draft Cochrane  
7 reviews were presented to the committee by Cochrane Neonatal and the committee and  
8 NGA technical team had an opportunity to review the draft and provide feedback. Additionally  
9 the NGA technical team summarised the Cochrane Neonatal reviews to develop a chapter  
10 for this guideline (Evidence Review - D4).

## 11 Evidence statements

12 Evidence statements are presented after the GRADE tables in each evidence report. They  
13 summarise key features in the available clinical evidence. The wording reflects the certainty  
14 or uncertainty in the estimate of effect (quantitative evidence) or review finding (qualitative  
15 evidence). Evidence statements are presented by outcome or theme, and encompass the  
16 following features in the evidence:

17 • the quality of the evidence

18 • the numbers of studies and participants for the outcome concerned or prognostic/risk  
19 factor or prediction model (quantitative evidence) or that contributed to themes (qualitative  
20 evidence)

21 • where relevant, whether or not the estimate of effect is clinically important.

22 • where relevant, an indication of the direction of effect (for example, if a treatment is  
23 beneficial or harmful compared with another,

24 • a brief description of the participants

25 • the imprecision of the effect: for serious imprecision, the statement ‘there was uncertainty  
26 around the effect’ was included and for very serious imprecision, the statement ‘there was  
27 high uncertainty around the effect was used.

## 28 Economic evidence

29 The aim of the health economic input to the guideline was to inform the committee of  
30 potential economic issues related to preterm babies and term babies who require parenteral  
31 nutrition and to ensure that recommendations represented a cost effective use of healthcare  
32 resources. Health economic evaluations aim to integrate data on healthcare benefits (ideally  
33 in terms of quality-adjusted life-years [QALYs]) with the costs of different care options. In  
34 addition, the health economic input aimed to identify areas of high resource impact. These  
35 are recommendations which might have a large impact on Clinical Commissioning Groups’ or  
36 Trusts’ finances and so need special attention.

## 37 Reviewing economic evidence

38 Titles and abstracts of articles identified through the economic literature searches were  
39 assessed for inclusion using the predefined eligibility criteria listed in Table 11.

### 40 Table 11: Inclusion and exclusion criteria for systematic reviews of economic 41 evaluations

Inclusion criteria
Intervention or comparators in accordance with the guideline scope
Study population in accordance with the guideline scope

#### Inclusion criteria

Full economic evaluations (cost-utility, cost-effectiveness, cost-benefit or cost-consequence analyses) assessing both costs and outcomes associated with interventions of interest

Only studies from Organisation for Economic Co-operation and Development countries were included, because the aim of the review was to identify economic information transferable to the UK context

#### Exclusion criteria

Abstracts containing insufficient methodological details

Cost-of-illness type studies

Studies which adopted a very narrow perspective and included only intervention costs or reported only unit cost data

1 Once the screening of titles and abstracts was completed, full-text copies of potentially  
2 relevant articles were requested for detailed assessment. The economic evidence study  
3 selection is presented in supplementary material D. Existing economic evidence considered  
4 in the guideline is provided in the respective evidence reports, following presentation of the  
5 relevant clinical evidence. The references to included studies and the respective evidence  
6 tables with the study characteristics and results are provided in appendix H of the relevant  
7 evidence report, the economic evidence profiles are provided in the corresponding appendix  
8 I.

## 9 Health economic modelling

10 As well as reviewing the published economic literature, as described above, new economic  
11 analysis was undertaken in selected areas prioritised by the committee in conjunction with  
12 the health economist. Topics were prioritised on the basis of the following criteria, in  
13 accordance with [Developing NICE guidelines: the manual 2014](#):

- 14 • the overall importance of the recommendation, which may be a function of the number of  
15 people affected and the potential impact on costs and health outcomes per patient
- 16 • the current extent of uncertainty over cost effectiveness, and the likelihood that economic  
17 analysis will reduce this uncertainty
- 18 • the feasibility of building an economic model.

19 The guideline committee prioritised the following review questions where it was thought that  
20 economic considerations would be particularly important in formulating recommendations.

- 21 • What are the predictors for enteral feeding success?
- 22 • For those neonates where parenteral nutrition is required, what is the optimal timeframe  
23 for doing this?
- 24 • What is the clinical effectiveness, efficacy and safety of lipid formulations from different  
25 sources (for example soya, fish oil, or mixed sources)?
- 26 • What is the effectiveness, efficacy and safety of standardised parenteral nutrition bags as  
27 compared to individualised bags?
- 28 • What overall osmolality, concentration of calcium, and glucose/dextrose in parenteral  
29 nutrition can determine whether to administer centrally or peripherally?
- 30 • In babies on parenteral nutrition, what is the optimal frequency of blood sampling for  
31 monitoring glucose, calcium, phosphate, potassium and serum triglycerides?
- 32 • What amount of enteral nutrition (measured in terms of ml/kg/day or kcal/kg/day) should  
33 be given when starting enteral feeds?
- 34 • Are nutrition care teams (pharmacists and dietician) effective and safe in providing  
35 parenteral nutrition in preterm and term babies?

36 Ultimately, original health economic modelling was undertaken only for the review question  
37 looking at the effectiveness, efficacy and safety of standardised parenteral nutrition bags as

1 compared to individualised bags. The clinical evidence was insufficient to allow de novo  
2 economic modelling in any of the other prioritised areas. Detail of the cost effectiveness  
3 analysis undertaken for the guideline is presented in the evidence chapter G with summary  
4 provided following the presentation of the relevant clinical evidence, with full methods and  
5 results provided in appendix J.

## 6 Cost effectiveness criteria

7 NICE's report [Social value judgements: principles for the development of NICE guidance](#)  
8 sets out the principles that committees should consider when judging whether an intervention  
9 offers good value for money. In general, an intervention was considered to be cost effective if  
10 any of the following criteria applied (provided that the estimate was considered plausible):

- 11 • the intervention dominated other relevant strategies (that is, it was both less costly in  
12 terms of resource use and more clinically effective compared with all the other relevant  
13 alternative strategies)
- 14 • the intervention cost less than £20,000 per QALY gained compared with the next best  
15 strategy
- 16 • the intervention provided clinically important benefits at an acceptable additional cost  
17 when compared with the next best strategy.

18 The committee's considerations of cost effectiveness are discussed explicitly under the  
19 heading 'Consideration of economic benefits and harms' in the relevant evidence reports.

## 20 Developing recommendations

### 21 Guideline recommendations

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

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

32 For further details refer to [Developing NICE guidelines: the manual](#) (NICE 2014).

### 33 Research recommendations

34 When areas were identified for which evidence was lacking, the committee considered  
35 making recommendations for future research. For further details refer to [Developing NICE](#)  
36 [guidelines: the manual](#) (NICE 2014).

## 37 Validation process

38 This guideline was subject to a 6-week public consultation and feedback process. All  
39 comments received from registered stakeholders were responded to in writing and posted on  
40 the NICE website at publication. For further details refer to [Developing NICE guidelines: the](#)  
41 [manual](#) (NICE 2014).

## 1 Updating the guideline

2 Following publication, NICE will undertake a surveillance review to determine whether the  
3 evidence base has progressed sufficiently to consider altering the guideline  
4 recommendations and warrant an update. For further details refer to [Developing NICE](#)  
5 [guidelines: the manual](#) (NICE 2014).

## 6 Funding

7 The NGA was commissioned by NICE to develop this guideline.

## 8 References

- 9  
10 **Bradburn 2007**  
11  
12 Bradburn, M. J., Deeks, J. J., Berlin, J. A., & Localio, A. R. Much ado about nothing:  
13 A comparison of the performance of meta-analytical methods with rare events.  
14 *Statistics in Medicine*, 26, 53–77, 2007.
- 15 **Dixon-Woods 2005**  
16 Dixon-Woods M, Agarwal S, Jones D et al. (2005) Synthesising qualitative and quantitative  
17 evidence: a review of possible methods. *Journal of Health Services Research & Policy* 10(1),  
18 45–53
- 19 **Hayden 2013**  
20 Jill A. Hayden, Danielle A. van der Windt, Jennifer L. Cartwright, Pierre Côté, Claire  
21 Bombardier. Assessing Bias in Studies of Prognostic Factors. *Ann Intern Med*.  
22 2013;158:280–286. doi: 10.7326/0003-4819-158-4-201302190-00009
- 23 **Higgins 2011**  
24 Higgins JPT, Green S (editors) (2011) *Cochrane Handbook for Systematic Reviews of*  
25 *Interventions* Version 5.1.0 [updated March 2011] The Cochrane Collaboration. Available  
26 from [www.handbook.cochrane.org](http://www.handbook.cochrane.org) (accessed 15 June 2018)
- 27 **Lewin 2015**  
28 Lewin S, Glenton C, Munthe-Kaas H et al. (2015) Using qualitative evidence in decision  
29 making for health and social interventions: an approach to assess confidence in findings from  
30 qualitative evidence syntheses (GRADE-CERQual). *PLoS Med* 12(10), e1001895
- 31
- 32 **Murphy 1998**  
33 Murphy, M,K,. Black, N,A,. Lamping, D,L,. McKee, C,M,. Standerson, C,F,B,. Askam, J,.  
34 Consensus development methods, and their use in clinical guideline development. *Heath*  
35 *Technology Assessment*, 2, 1998.
- 36
- 37 **NICE 2014**  
38 National Institute for Health and Care Excellence (NICE) (2014) *Developing NICE guidelines:*  
39 *the manual* (updated 2017). Available from

- 1 <https://www.nice.org.uk/process/pmg20/chapter/introduction-and-overview> (accessed 15  
2 June 2018)
- 3 **NICE 2018**
- 4 National Institute for Health and Care Excellence (NICE) (2014) NICE Policy on conflict of  
5 interest (updated 2017). Available from [https://www.nice.org.uk/Media/Default/About/Who-  
6 we-are/Policies-and-procedures/declaration-of-interests-policy.pdf](https://www.nice.org.uk/Media/Default/About/Who-we-are/Policies-and-procedures/declaration-of-interests-policy.pdf) (accessed 15 June 2018)
- 7 **Santesso 2016**
- 8 Santesso N, Carrasco-Labra A, Langendam M et al. (2016) Improving GRADE evidence  
9 tables part 3: detailed guidance for explanatory footnotes supports creating and  
10 understanding GRADE certainty in the evidence judgments. *Journal of clinical epidemiology*  
11 74, 28-39