National Institute for Health and Care Excellence

Draft for consultation

Cerebral palsy in adults

Methods

NICE guideline TBC

Supplementary material C

July 2018

Draft for Consultation

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

Disclaimer

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

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

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 All NICE guidance is subject to regular review and may be updated or withdrawn.

Copyright

ISBN:

Contents

Development of the guideline **5**
Remit 5
What this guideline covers 5
Groups that are covered 5
Clinical areas that are covered 5
What this guideline does not cover 6
Groups that are not covered 6
Clinical areas that are not covered 6
Methods **7**
Developing the review questions and outcomes 7
Searching for evidence 13
Clinical search literature 13
Health economics search literature 13
Call for evidence 14
Reviewing clinical evidence 14
Systematic review process 14
Type of studies and inclusion/exclusion criteria 14
Methods of combining evidence 15
Appraising the quality of evidence 16
Qualitative reviews 21
Evidence statements 22
Economic evidence 22
Reviewing economic evidence 22
Health economic modelling 23
Cost effectiveness criteria 23
Developing recommendations 23
Guideline recommendations 23
Research recommendations 24
Validation process 24
Updating the guideline 24
Funding 24
References **25**

Development of the guideline

Remit

The National Institute for Health and Care Excellence (NICE) commissioned the National Guideline Alliance (NGA) to develop a new guideline on cerebral palsy in adults.

What this guideline covers

Groups that are covered

- Adults aged 25 and over with cerebral palsy (NICE has published a guideline on cerebral palsy in under 25s).
- Adults aged 19 and over with cerebral palsy, in relation only to the management of spasticity and associated movement disorders such as dystonia (NICE has published a guideline on spasticity in under 19s).

Subgroups

Specific consideration will be given to recognised subgroups within the cerebral palsy population:

- Subgroups with different levels of functional disability (for example, Gross Motor Functional Classification System levels I to V).

Clinical areas that are covered

The guideline covers the following clinical issues:

- Management of abnormal muscle tone in adults aged 19 and over with cerebral palsy, including spasticity and associated movement disorders such as dystonia:
  - pharmacological management
  - neurosurgical management.
- Assessing and monitoring the following complications and comorbidities associated with cerebral palsy in adults aged and 25 over:
  - disorders of bones and joints, including osteoarthritis, osteoporosis and musculoskeletal deformity (especially of the neck, hip and spine)
  - mental health problems
  - feeding and nutritional problems.
- Identifying and managing respiratory disorders associated with cerebral palsy in adults aged 25 and over, including assisted ventilation.
- Interventions that improve function and participation for adults aged 25 and over with cerebral palsy:
  - physical therapy programmes (such as sporting activity, strengthening programmes or training, task-oriented upper limb training)
  - augmentative and alternative communication systems
  - electronic assistive technology
  - equipment to help with mobility (such as orthotics)

    - vocational and independent living skills training.
- Identifying pain, such as musculoskeletal and gastrointestinal pain, in adults aged 25 and over with cerebral palsy.
- Configuration of services for adults aged 25 and over with cerebral palsy:
    - Specialist services.
    - Access to primary and secondary care.

For further details please refer to the scope on the NICE website.

What this guideline does not cover

Groups that are not covered

The guideline does not cover the following groups:

- Children and young people under 25 with cerebral palsy, except for people aged 19 and over in relation to spasticity and associated movement disorders.
- Adults with a progressive movement disorder, spasticity or dystonia that is not associated with cerebral palsy.

Clinical areas that are not covered

This guideline does not cover the following areas:

- Managing pain
- Managing mental health problems.

Methods

This chapter sets out in detail the methods used to review the evidence and to generate recommendations in the guideline. This guideline was developed using the methods described in Developing NICE guidelines: the manual (2014).

Declarations of interest were recorded according to NICE's 2014 conflicts of interest policy from May 2016 until April 2018. From April 2018 onwards they were recorded according to NICE's 2018 conflicts of interest policy.

Developing the review questions and outcomes

The 16 review questions developed for this guideline were based on the key areas identified in the guideline scope. They were drafted by the NGA and refined and validated by the committee. They cover all areas of the scope and were signed-off by NICE (see Table 1).

The review questions were based on the following frameworks:

- intervention reviews: population, intervention, comparator and outcome (PICO)
- diagnostic test accuracy reviews: population, index test, reference standard and outcome (PIRO)
- qualitative reviews: Population or problem, interest (i.e. defined event, activity, experience or process) and context (PICo)

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

Review questions on health monitoring (B1, B3 and C1) were framed as intervention reviews (a comparison of different monitoring protocols or assessments) but in the absence of test and treat studies the diagnostic accuracy of tests used for monitoring was summarised with the assumption that accurate identification of health problems is likely to improve outcome.

Full literature searches, critical appraisals and evidence reviews were completed for all review questions. Review questions A1, A2 and A3 were searched using a single literature search as were C1, C2 and C3 and D1, D2, D3 and D4.

There are broad topic areas, as indicated by letters, but evidence reviews are presented individually. This was decided because the topics within the sections were sufficiently different to be reviewed and discussed separately and future updates would relate to individual reviews rather than overarching topics.

Table 1: Description of review questions

Chapter or section	Type of review	Review question	Outcomes
A1 Management of abnormal muscle tone – pharmacological treatments for spasticity.	Intervention	A1 Which pharmacological treatments for spasticity (for example, enteral baclofen, tizanidine, diazepam, cannabinoids,	**Critical** • Motor function ○ Swallowing problems ○ Goal Attainment Scale (GAS)

Chapter or section	Type of review	Review question	Outcomes
		and botulinum toxin injections) are most effective for improving motor function, participation and quality of life in adults with cerebral palsy?	○ Functional Independence Measure (FIM) • Muscle tone • Health-related quality of life • Treatment related adverse events ○ Swallowing problems ○ Seizure threshold ○ Undue weakness/loss of function – use of spasticity positively ○ Drowsiness and cognitive change ○ Specific problems in people with low proximal tone and high peripheral tone **Important** • Patient or carer reported satisfaction • Participation
A2 Management of abnormal muscle tone in adults aged 19 and over with cerebral palsy – neurosurgical treatments to reduce spasticity.	Intervention	A2 Are neurosurgical procedures (intrathecal baclofen pump and selective dorsal rhizotomy) effective in adults aged 19 and over with cerebral palsy to reduce spasticity and or dystonia?	**Critical** • Walking (for ambulant people only) • Gross motor function (both upper / lower limb) • Tone (for example Ashworth scale) • Health related quality of life **Important** • Pain • Adverse events (CSF leakage, infection, respiratory depression, baclofen withdrawal and baclofen overdose) • Satisfaction (patient or carer reported) • Use of concurrent medications
A3 Management of abnormal muscle tone in adults aged	Intervention	A3 Which treatments (pharmacological treatment (levodopa,	**Critical outcomes** • Health related quality of life

Chapter or section	Type of review	Review question	Outcomes

19 and over with cerebral palsy – treatments to reduce dystonia.		anticholinergic drugs, and botulinum toxin injections), neurosurgical procedure (deep brain stimulation, ITB) are most effective for managing dystonia in adults with cerebral palsy where dystonia is the predominant abnormality of tone?	• Dystonia rating scales
			  ○ DMFRS
			  ○ Fahn-Marsden Rating Scale
			• Patient or carer reported satisfaction
			Important outcomes
			• Motor function using functional measures
			• Goal attainment scores
			• Adverse events
			• Pain

B1. Assessing and monitoring complications and comorbidities - disorders of bones and joints.	Intervention and diagnostic test accuracy	B1 What is the most effective protocol for monitoring the following disorders of bones and joints in adults with cerebral palsy:
		• osteoarthritis
		• osteoporosis (including osteopenia and osteomalacia)
		• hip displacement
		• spinal deformity, including scoliosis, kyphosis and lordosis
		• cervical instability leading to cervical myelopathy
			Critical
			• Incidence of bone or joint disorders
			• Severity of bone or joint disorders
			• Diagnostic accuracy:
			  ○ Sensitivity
			  ○ Specificity
			  ○ Negative /positive likelihood ratios
			• Validity reliability
			Important
			• Patient satisfaction

B2. Assessing and monitoring complications and comorbidities - mental health problems.	Diagnostic test accuracy	B2 Which mental health assessment tools are clinically useful for adults with cerebral palsy?	**Critical**
			• Diagnostic accuracy:
			  ○ Sensitivity
			  ○ Specificity
			  ○ Positive/Negative likelihood ratio
			• Validity and reliability
			Important
			• Patient satisfaction

B3. Assessing and monitoring complications and comorbidities - feeding and nutrition.	Intervention and diagnostic test accuracy	B3 What is the best way to assess and monitor the safety (of swallowing and risk of aspiration) and effectiveness of feeding and maintaining nutrition in adults with cerebral palsy?	**Critical** • Function • HR-QoL • Chest infection **Important** • Patient satisfaction • Mortality • Weight

Chapter or section	Type of review	Review question	Outcomes

			• Skin integrity
			• Feeding time
			• TOMS
			• Diagnostic accuracy:
			  ○ Sensitivity
			  ○ Specificity
			  ○ Positive/Negative likelihood ratio

C1. Identifying and managing respiratory disorders associated with cerebral palsy – protocols to monitor respiratory disorders.	Intervention and diagnostic test accuracy	C1 What is the most effective protocol for monitoring respiratory health in adults with cerebral palsy?	**Critical**
			• Respiratory health
			• Overall survival
			• Hospital admission
			Important
			• Secondary conditions (e.g. colds, asthma, sleep apnoea, daytime sleepiness)
			• Respiratory function
			• Health related quality of life
			• Satisfaction
			• Diagnostic accuracy:
			  ○ Sensitivity
			  ○ Specificity
			  ○ Positive and negative likelihood ratios

C2. Identifying and managing respiratory disorders associated with cerebral palsy – assisted ventilation.	Intervention	C2 Does assisted ventilation improve quality of life for adults with cerebral palsy who have a chronic respiratory disorder (including respiratory failure)?	**Critical** • Hospital admissions • Overall survival • Quality of life (carer or self-reported) **Important** • Treatment complications • Daytime sleepiness and fatigue
C3 Identifying and managing respiratory disorders associated with cerebral palsy – prophylactic treatments.	Intervention	C3 Are prophylactic treatments (for example, antibiotics, chest physiotherapy, cough assistance) effective in preventing respiratory infections in adults with cerebral palsy?	**Critical** • Respiratory infections • Hospital admission • Overall survival **Important** • Health related quality of life • Satisfaction
D1. Interventions that improve function and	Intervention	D1 Which interventions (for example, vocational and independent living	**Critical** • Participation

Chapter or section	Type of review	Review question	Outcomes
participation – vocational and independent living skills.		skills training) promote participation in adults with cerebral palsy?	○ occupation ○ employment ○ vocational activity ○ leisure ○ (AUS)TOMS ○ GAS • Independence • Health related quality of life **Important** • Function ○ COPM ○ FIM/FAM • Self-efficacy / self-determination
D2. Interventions that improve function and participation – physical function	Intervention	D2 Which interventions are effective for maintaining physical function and mobility in adults with cerebral palsy?	**Critical** • Participation (incorporating mobility) • Physical function • Health related quality of life & psychological wellbeing **Important** • Independence • Fatigue • Frequency of falls [in a subset] • Complications of treatment • Adherence
D3. Interventions that improve function and participation – vocational and independent living skills	Intervention	D3 What is the effectiveness of electronic assistive technology in promoting independence in adults with cerebral palsy?	**Critical** • Participation • Function • Independence • Health related quality of life **Important** • Frequency and duration of healthcare worker / carer contact • Person & carer satisfaction • Admission to long term residential care
D4. Interventions that improve function and	Intervention	D4 Which interventions (for example augmentative and	**Critical** • Participation

Chapter or section	Type of review	Review question	Outcomes
participation – communication		alternative communication systems) are effective in promoting communication for adults with cerebral palsy who have communication difficulties?	• Function (expressive and receptive communication) • Independence (communication in different situations) **Important** • Health related quality of life • Patient satisfaction
E. Identifying pain, such as musculoskeletal and gastrointestinal pain.	Diagnostic test accuracy	E1 What is the value of self-report and observational techniques for providing a standardised way of identifying and localising pain in adults with cerebral palsy?	**Critical** • Psychometric properties ○ Concurrent validity ○ Internal consistency ○ Inter- or intra-rater reliability • Test accuracy: ○ Sensitivity ○ Specificity
F1. Configuration of services– service design.	Intervention	F1 What is the most clinical and cost-effective configuration of services (setting and staffing) for adult with cerebral palsy?	**Critical** • Health-related quality of life • Time to treatment • Hospital admissions (unplanned) **Important** • Satisfaction (patient or carer reported) • Adverse effects (from delayed identification or management) • Residential care admissions (unplanned) • Length of hospital stay • Mortality
F2. Configuration of services– access to primary and secondary care.	Intervention	F2 What service configuration and what interventions can facilitate access to health care in adults with cerebral palsy, and what are the perceived barriers and facilitators for access to care in adults with cerebral palsy?	**Critical** • Qualitative outcomes: ○ Perceived barriers to health care - Personal - Organisational - Financial **Important** • Quantitative outcomes: ○ Service availability

Chapter or section	Type of review	Review question	Outcomes
			○ Utilisation of services ○ Secondary care services ○ Social care ○ Primary care surveillance ○ Dental

(AUS)TOMS: (Australian) Therapy Outcome Measures for Occupational Therapy; COPM: Canadian Occupational Performance Measure; CSF: cerebrospinal fluid; FAM: functional ability measure; FIM: functional independence measure; GAS: goal attainment scale; HR-QoL: Health-Related Quality of Life; TOMS: Therapy Outcome Measures-Swallowing.

Searching for evidence

Clinical search literature

Systematic literature searches were undertaken to identify all published clinical evidence relevant to the review questions.

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. All searches were conducted in MEDLINE, Embase and The Cochrane Library, with some additional database searching in AMED, PsycINFO and CINAHL for certain topic areas (for example PsycINFO for topic B2).

Re-run searches were carried out on 22nd March 2018. Any studies added to the databases after the date of the last search (even those published prior to this date) were not included unless specifically stated in the text.

Search strategies were quality assured by cross-checking reference lists of relevant papers, analysing search strategies in other systematic reviews and asking committee members to highlight any additional studies. The questions, the study types applied, the databases searched and the years covered can be found in appendix B in each evidence review chapter.

Searching for grey literature or unpublished literature was not undertaken. During the scoping stage, a search was conducted for guidelines and reports on websites of organisations relevant to the topic. Any references suggested by stakeholders at the scoping consultation were considered. Clinical search strategies can be found in appendix B of each evidence review.

Health economics search literature

A global search of economic evidence was undertaken in December 2016 and re-run in March 2018 The following databases were searched:

- MEDLINE (Ovid)
- EMBASE (Ovid)
- Health Technology Assessment database (HTA)
- NHS Economic Evaluations Database (NHS EED).

Further to the database searches, the committee was contacted with a request for details of relevant published and unpublished studies of which they may have knowledge; reference lists of key identified studies were also reviewed for any potentially relevant studies. Finally, the NICE website was searched for any recently published guidance relating to cerebral palsy that had not been already identified via the database searches.

The search strategy for existing economic evaluations combined terms capturing the target condition (cerebral palsy) and, for searches undertaken in MEDLINE and EMBASE, terms to capture economic evaluations. No restrictions on language or setting were applied to any of the searches, but a standard exclusions filter was applied (letters, animals, etc.). Full details of the search strategy are presented in Supplementary material D: Health economic literature review.

Call for evidence

No call for evidence was made.

Reviewing clinical evidence

Systematic review process

The evidence was reviewed following these steps.

- Potentially relevant studies were identified for each review question from the relevant search results by reviewing titles and abstracts. Full papers were then obtained.
- Full papers were reviewed against pre-specified inclusion and exclusion criteria in the review protocols (in appendix A of each evidence review chapter).
- Key information was extracted on the study's methods, according to the factors specified in the protocols and results. These were presented in summary tables (in each review chapter) and evidence tables (in appendix D of each evidence review chapter).
- Relevant studies were critically appraised using the appropriate checklist as specified in Developing NICE guidelines: the manual 2014.
- Summaries of evidence were generated by outcome (included in the relevant review chapters) and were presented in committee meetings.
- Results were summarised and reported in GRADE profiles (for intervention reviews) or their equivalent (for diagnostic test accuracy and qualitative reviews)
- Model performance studies: data were presented individually by study.

All drafts of reviews were checked by a senior reviewer.

Type of studies and inclusion/exclusion criteria

Systematic reviews (SRs) with meta-analyses (for diagnostic or intervention reviews) or SRs of qualitative studies were considered the highest quality evidence to be selected for inclusion.

For intervention reviews, randomised controlled trials (RCTs) were included because they are considered the most robust study design for unbiased estimation of intervention effects. Based on their judgement, if the committee believed RCT data

were not appropriate or there was limited evidence from RCTs, they agreed to include cohort studies with a comparative group.

Posters, letters, editorials, comment articles, unpublished studies and studies not in the English language were excluded. Narrative reviews were also excluded, but individual references were checked for inclusion. Conference abstracts were not included due to insufficient information to assess their quality.

For quality assurance of study identification, a 10% random sample of the literature search results for every review was sifted by a second reviewer.

The inclusion and exclusion of studies was based on the review protocols, which can be found in appendix A of each evidence review chapter. Excluded studies and the reasons for their exclusion are listed in appendix K of each evidence review. In addition, the committee was consulted to resolve any uncertainty about inclusion or exclusion.

Methods of combining evidence

Data synthesis for intervention reviews

Pairwise meta-analysis of homogenous randomised trials was done using Review Manager 5 (RevMan 5) software. For binary outcomes, such as occurrence of adverse events, the Mantel-Haenszel method of statistical analysis was used to calculate risk ratios (relative risks, RRs) with 95% confidence intervals (CIs).

For continuous outcomes, measures of central tendency (mean) and variation (standard deviation (SD)) are required for meta-analysis. Data for continuous outcomes (such as health-related quality of life score or length of hospital stay) were analysed using an inverse-variance method for pooling weighted mean differences.

Statistical heterogeneity was assessed by visually examining the forest plots, and by considering the chi-squared test for significance with heterogeneity defined as a $p<0.1$ or an I-squared inconsistency statistic value of 50% or more. Where heterogeneity was present, predefined subgroup analyses were performed. If the heterogeneity still remained, a random effects (DerSimonian 2015) model was employed to provide a more conservative estimate of the effect.

Results from multiple observational studies of the same comparison were not pooled but presented as a range of effects. This was due the high risk of selection bias in observational studies whereby differences in participant characteristics between treatment arms leads to a biased estimate of treatment effect.

Forest plots were generated to present the results (please see appendix E of each intervention evidence review).

Data synthesis for diagnostic test accuracy reviews

Meta-analysis of diagnostic test accuracy was not done because there were no reviews with multiple studies reporting the same test. Results were presented individually for each study.

Sensitivity and specificity plots were generated to present the results (please see appendix E of each diagnostic test accuracy evidence review chapter).

Data synthesis for qualitative reviews

Each qualitative study was summarised by theme and meta-synthesis was carried out where appropriate to identify an overarching framework of themes and their subthemes. This framework was illustrated graphically using a theme-map showing how the themes and sub-themes were connected.

Appraising the quality of evidence

Intervention reviews

GRADE methodology (the Grading of Recommendations Assessment, Development and Evaluation)

For intervention reviews, the evidence for outcomes from the included studies was evaluated and presented using GRADE, which was developed by the international GRADE working group.

The software developed by the GRADE working group (GRADEpro) was used to assess the quality of each outcome, taking into account individual study quality factors and the meta-analysis results. The clinical evidence profile tables include details of the quality assessment and pooled outcome data, where appropriate, an absolute measure of intervention effect and the summary of quality of evidence for that outcome. In this table, the columns for intervention and control indicate summary measures of effect and measures of dispersion (such as mean and SD or median and range) for continuous outcomes and frequency of events (n/N; the sum across studies of the number of participants with events divided by sum of the number of completers) for binary outcomes. Reporting or publication bias was taken into consideration in the quality assessment and reported in the clinical evidence profile tables if it was apparent.

The selection of outcomes for each review question was decided when each review protocol was discussed with the committee, and was informed by committee discussion and by key papers.

The evidence for each outcome in the intervention reviews was examined separately for the quality elements listed and defined in Table 2. Each element was graded using the quality levels listed in Table 3.

The main criteria considered in the rating of these elements are discussed below. Footnotes were used in the GRADE profiles to describe reasons for grading a quality element as having serious or very serious limitations. The ratings for each component were combined to obtain an overall assessment for each outcome (Table 4).

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

Quality element	Description
Risk of bias	Limitations in the study design and implementation may bias the estimates of the treatment effect. High risk of bias for the majority of the evidence decreases confidence in the estimate of the effect.
Inconsistency	Inconsistency refers to an unexplained heterogeneity of results or findings.

Quality element	Description
Indirectness	Indirectness refers to differences in study population, intervention, comparator and outcomes between the available evidence and the review question, such that the effect estimate is changed. This is also related to applicability or generalisability of findings.
Imprecision	Results are imprecise when studies include relatively few patients and / or few events and thus have wide confidence intervals around the estimate of the effect. Imprecision results if the confidence interval includes the clinically important threshold (minimally important difference – see below).
Publication bias	Publication bias is a systematic underestimate or an overestimate of the underlying beneficial or harmful effect due to selective publication of studies.

Table 3: Levels of quality elements in GRADE

Levels of quality elements in GRADE	Description
None/ no serious	There are no serious issues with the evidence.
Serious	The issues are serious enough to downgrade the outcome evidence by 1 level.
Very serious	The issues are serious enough to downgrade the outcome evidence by 2 levels.

Table 4: Levels of overall quality of outcome evidence in GRADE

Overall quality of outcome evidence in GRADE	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.

Assessing risk of bias in intervention reviews

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

Risk of bias in RCTs was assessed using the Cochrane Risk of Bias Tool (see appendix H in Developing NICE guidelines: the manual 2014).

It should be noted that a study with a poor methodological design does not automatically imply high risk of bias; the bias is considered individually for each outcome and it is assessed whether this poor design will impact on the estimation of the intervention effect.

For observational studies methodological quality was assessed using the Newcastle-Ottawa Scale (Wells 2008) for cohort and cross-sectional studies or the Effective

Practice and Organisation of Care (EPOC) risk of bias tool for before-and-after studies (see appendix H in Developing NICE guidelines: the manual 2014).

Assessing inconsistency in intervention reviews

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

Statistical heterogeneity was assessed by visually examining the forest plots, and by considering the chi-squared test for significance at $p<0.1$ and the I-squared inconsistency statistic (with an I-squared value of 50 to 80% indicating potentially serious inconsistency and I-squared value of over 80% indicating very serious inconsistency).. When no plausible explanation for the heterogeneity could be found, the quality of the evidence was downgraded in GRADE by 1 or 2 levels for the domain of inconsistency, depending on the extent of heterogeneity in the results.

Assessing indirectness in intervention reviews

Directness refers to the extent to which the populations, intervention, comparisons and outcome measures are similar 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.

Assessing imprecision and clinical significance in intervention reviews

Imprecision in guidelines concerns whether the uncertainty (CI) around the effect estimate means that it is not clear whether there is a clinically important difference between interventions or not (that is, whether the evidence would clearly support one recommendation or appear to be consistent with several different types of recommendations). Therefore, imprecision differs from the other aspects of evidence quality because it is not really concerned with whether the point estimate is accurate or correct (has internal or external validity). Instead, it is concerned with the uncertainty around the point estimate actually is. This uncertainty is reflected in the width of the CI.

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

Imprecision in the evidence reviews is assessed by considering whether the width of the 95% CI of the effect estimate is relevant to decision-making, taking each outcome in isolation. This assessment also involves effect size thresholds for clinical importance (the minimally important difference, MID) for benefit and for harm.

If the effect estimate CI includes clinically important benefit (or harm) there is uncertainty over which decision to make (based on this outcome alone). The CI is consistent with 2 possible decisions and so this is considered to be imprecise in the GRADE analysis and the evidence is downgraded by 1 level ('serious imprecision').

An effect CI including clinically important benefit, clinically important harm and no effect is consistent with 3 possible decisions. This is considered to be very imprecise in the GRADE analysis and the evidence is downgraded by 2 levels ('very serious imprecision').

If the effect estimate did not include clinically important benefit (or harm), it was considered whether the criterion for Optimal Information Size (OIS) was met (see below), if not, the outcome was downgraded one level.

Minimally important differences

The literature was searched for established MIDs for the selected outcomes in the evidence reviews. In addition, the committee was asked whether they were aware of any acceptable MIDs in the clinical community.

If no published or acceptable MIDs were identified, the committee considered whether it was clinically acceptable to use the GRADE default MIDs to assess imprecision. For binary outcomes, GRADE default MIDs are RRs of 0.8 and 1.25 (due to the statistical distribution of this measure this means that this is t symmetrical on a log [RR] scale). For continuous outcomes, GRADE default MIDs are half of the SD of the control group.

- There were published MID values (compiled in the Rehabilitation Measures Database: RMD 2018) available for the following measurement scales for level of functional ability or disability, pain and independence:
  - Goal Attainment Scale: 7 units
  - Modified Ashworth Scale: 1 unit
  - Quality of Upper Extremities Test: 5 units
  - International Classification of Functioning (ICF) – Measure of Participation and Activities Screener: 2 units
  - Community Balance and Mobility Scale: 10 units
  - Canadian Occupational Performance Measure: 2 units
  - Five Times Sit to Stand Test: 2.5 seconds
  - Seated Shot-Put: 40 cm
  - Timed Up and Go: 5 seconds
  - Australian Therapy Outcome Measures for Occupational Therapy: 0.5 units
  - Assessment of Life Habits: use minimal detectable change for each subdomain reported on rehabmeasures.org
  - Pain: 30% reduction – corresponding to 'much improved' or 'very much improved' on a global impression of change, or 2 points on a 0 to 11 pain intensity numerical rating scale
  - Assessment of Life Habits: use minimal detectable change for each subdomain reported on rehabmeasures.org
  - Functional Independence measure (FIM) total score 20 points
  - Functional Assessment measure (FAM) total score 20 points
- For all other outcomes, GRADE default MID values were used as a starting point and decisions on clinical importance were then considered based on the absolute risk difference.

Optimal information size (OIS)

Evaluating the CI is not sufficient to assess imprecision. When there is a small number of events the CI can be narrow but the results may be fragile. Therefore, it is suggested that in addition to considering whether the CI crosses thresholds for MIDs, the OIS, representing the number of patients generated by a conventional single-trial sample size calculation, should be considered (Schünemann 2013). In statistical hypothesis testing alpha is probability of rejecting the null hypothesis given that it is true and beta is the probability of failing to reject the null hypothesis given that it is false. For continuous outcomes, using the standard alpha and beta values of 0.05 and 0.20 respectively, a total sample size (across both arms) of approximately 400 would be required to detect an effect size of 0.2; therefore if N < 400 for an outcome, the evidence would be considered imprecise and downgraded by 1 level ('serious imprecision'). For binary outcomes, evidence should be considered imprecise and downgraded by 1 level ('serious imprecision') if the total number of events (across both arms) is less than 300. For outcomes where any statistically significant change was considered by the committee to be clinically important, imprecision was rated based on OIS alone; for all other outcomes, imprecision was determined based on the width of the CI and the OIS.

Diagnostic test accuracy reviews

Modified GRADE methodology for diagnostic test accuracy reviews

The GRADE approach was modified to assess the quality of evidence about diagnostic test accuracy by adapting the principles of GRADE for intervention reviews as described below. Four domains were considered: risk of bias, indirectness, inconsistency and imprecision. Each domain was rated as 'no serious...', 'serious...' or 'very serious...'. These domains were then combined to give the overall certainty in the body of evidence, rated as 'very low', 'low', 'moderate' or 'high'.

Assessing risk of bias in diagnostic test accuracy reviews

Risk of bias in diagnostic test accuracy studies was assessed using the risk of bias items from the QUADAS-2 checklist (see appendix H in Developing NICE guidelines: the manual 2014). An overall risk of bias judgement was for each study was reached by considering the QUADAS-2 bias domains together. The risk of bias for the body of diagnostic test accuracy evidence was based on the risk of bias from the individual studies but with consideration of how much each study contributed to the overall evidence base.

Assessing indirectness in diagnostic test accuracy reviews

Indirectness was assessed using the applicability items from the QUADAS-2 checklist. An overall indirectness judgement was for each study was reached by considering the QUADAS-2 applicability domains together. The indirectness for the body of diagnostic test accuracy evidence was based on the indirectness of the individual studies but with consideration of how much each study contributed to the overall evidence base.

Assessing inconsistency in diagnostic test accuracy reviews

Where there were multiple studies the body of evidence was downgraded for serious inconsistency if there was unexplained variability between studies, when viewed on a

forest plot or Receiver Operating Characteristics (ROC) curve. If there was only one study then inconsistency was rated as 'not applicable'.

Assessing imprecision in diagnostic test accuracy reviews

Imprecision was judged by comparing the CI of the estimate of sensitivity or specificity to clinical decision thresholds agreed beforehand by the committee. The committee decided whether sensitivity or specificity was the most important for decision making and agreed two threshold values. First a threshold for high sensitivity/specificity (above which the test would be definitely recommended) and second a threshold for low sensitivity/specificity (below which the test would not be recommended). If the CI of the estimate of sensitivity or specificity included one of these thresholds then the evidence was downgraded for serious imprecision, because it was consistent with two possible decisions. If the CI included both these thresholds then the evidence was downgraded for very serious imprecision because it was consistent with three possible decisions. In this guideline sensitivity was prioritised for decision making about diagnostic tests and thresholds of 0.75 and 0.90 were chosen for low and high sensitivity respectively.

Qualitative reviews

GRADE CERQual methodology for qualitative reviews

The GRADE-CERQual (Confidence in the Evidence from Reviews of Qualitative research; Lewin 2015) approach was used to summarise the confidence in qualitative evidence. Each qualitative study was summarised by theme and meta-synthesis was carried out where appropriate to identify an overarching framework of themes and subthemes.

The overall confidence in evidence about each theme or sub-theme was rated as high, moderate, low or very low based on four dimensions: methodological limitations, applicability, coherence and adequacy of data.

Methodological limitations refer to the extent to which there were problems in the design or conduct of the studies that contributed evidence to the findings of the review.

Applicability of evidence was assessed by looking at the extent to which the body of evidence from the primary studies supporting the review findings is applicable to the review protocol

Coherence of findings was assessed by looking at the extent to which the review findings were well grounded in data from the contributing primary studies

Adequacy of data was assessed by looking at the degree of richness and quantity of data supporting the findings of the review

Assessing risk of bias in qualitative reviews

For qualitative studies, quality was assessed using a checklist for qualitative studies (as suggested in appendix H in Developing NICE guidelines: the manual 2014). This was based on the Critical Appraisal Skills Programme (CASP) checklist for qualitative studies.

Evidence statements

Evidence statements are summary statements presented after the GRADE profiles, highlighting the key features of the clinical 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 or theme and encompass the following key features of the evidence:

- the quality of the evidence
- the number of studies and the number of participants for a particular outcome
- a brief description of the participants
- the clinical significance of the effect and an indication of its direction (for example, if a treatment is clinically significant (beneficial or harmful) compared with another, or whether there is no clinically significant difference between the tested treatments).

Economic evidence

The aim of the health economic input to the guideline was to inform the committee of potential economic issues related to management of adults with cerebral palsy and to ensure that recommendations represented a cost effective use of healthcare resources. Health economic evaluations aim to integrate data on healthcare benefits (ideally in terms of quality-adjusted life-years (QALYs)) with the costs of different care options. In addition, the health economic input aimed to identify areas of high resource impact. These are recommendations which might have a large impact on Clinical Commissioning Groups' or Trusts' finances and so need special attention.

Reviewing economic evidence

The titles and abstracts of papers identified through the searches were independently assessed for inclusion using predefined eligibility criteria summarised in Table 5.

Table 5: Inclusion and exclusion criteria for the systematic reviews of economic evaluations

Inclusion criteria
Intervention or comparators according to the scope
Study population according to the scope
Full economic evaluations (cost utility, cost effectiveness, cost benefit or cost consequence analyses) that assess both the costs and outcomes associated with the interventions of interest
Exclusion criteria
Abstracts with insufficient methodological details
Cost of illness type studies

Once the screening of titles and abstracts was complete, full versions of the selected papers were acquired for assessment. The quality of evidence was assessed using the economic evaluations checklist as specified in Developing NICE guidelines: the manual 2014.

Health economic modelling

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

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

The full methods and results of de novo economic analyses are reported in appendix J of each evidence review that was modelled (topics A3 and F1). When new economic analysis was not prioritised, the committee made a qualitative judgement regarding cost effectiveness by considering expected differences in resource and cost use between options, alongside clinical effectiveness evidence identified from the clinical evidence review.

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

- the intervention dominated other relevant strategies (that is, it was both less costly in terms of resource use and more 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, or
- the intervention provided clinically significant benefits at an acceptable additional cost when compared with the next best strategy.

The committee's considerations of cost effectiveness are discussed explicitly under the 'Cost effectiveness and resource use' headings of the relevant sections.

Developing recommendations

Guideline recommendations

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. When clinical and economic evidence was of poor quality, conflicting or absent, the committee drafted recommendations based on the members' expert opinion. The considerations for making consensus-based recommendations include the balance between potential harms and benefits, the economic costs or implications compared with the economic benefits, current practices, recommendations made in other relevant guidelines, patient preferences and equality issues.

The main considerations specific to each recommendation are outlined under the 'The committee's discussion of the evidence' headings within each chapter as well as the 'rationale and impact' section in the short guideline.

For further details please refer to Developing NICE guidelines: the manual 2014.

Research recommendations

When areas were identified for which good evidence was lacking, the committee considered making recommendations for future research. For further details please refer to Developing NICE guidelines: the manual 2014.

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 at publication. For further details please refer to Developing NICE guidelines: the manual 2014.

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. For further details please refer to Developing NICE guidelines: the manual 2014.

Funding

The NGA was commissioned by NICE to develop this guideline.

References

DerSimonian2015

DerSimonian, R., Laird, N. Meta-analysis in clinical trials revisited. Contemporary clinical trials. 2015 Nov 30; 45:139-45.

NICE 2014

National Institute for Health and Care Excellence (NICE), Developing NICE guidelines: the manual (accessed 8 June 2017)

RMD 2018

Rehabilitation Measures Database, Rehabilitation Measures Database – Cerebral Palsy (accessed 8 June 2017)

Santesso 2016

Santesso, N., Carrasco-Labra, A., Langendam, M., Brignardello-Petersen, R., Mustafa, R.A., Heus, P., Lasserson, T., Opiyo, N., Kunnamo, I., Sinclair, D. and Garner, P., Improving GRADE evidence tables part 3: detailed guidance for explanatory footnotes supports creating and understanding GRADE certainty in the evidence judgments, Journal of clinical epidemiology, 74, 28-39, 2016.

Schünemann 2013

Schünemann, H., Brożek, J, Guyatt, G., Oxman, A., (eds.). Handbook for grading the quality of evidence and the strength of recommendations using the GRADE approach (updated October 2013). GRADE Working Group, 2013. Available from gdt.guidelinedevelopment.org/app/handbook/handbook.html (accessed 26 November 2017)

Wells 2008

Wells, G.A., Shea, B., O'Connel,l D., Peterson, J., Welch, V., Losos, M., Tugwel,l P. The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses, 2008. Available from http://www.ohri.ca/programs/clinical_epidemiology/oxford.asp (accessed 26 November 2017)