CHUMS

Children's oUtcome Measurement Study

Systematic Review Protocol

Project title: Informing the NHS outcomes framework: what outcomes of NHS care should be measured for children with neurodisability?

Short Title: 'CHUMS' Study: Children's Outcome Measurement Study

CHUMS Systematic Review Team:

Astrid Janssens
Chris Morris
Jo Thompson-Coon
Morwenna Rogers
Colin Green
Crispin Jenkinson
Alan Tennant

Systematic Review Protocol V7.0 30 July 2012

This project was funded by the National Institute for Health Research Health Services and Delivery Research (NIHR HS&DR) programme (project number 10/2002/16). Visit the HS&DR website for more information.
The views and opinions expressed therein are those of the authors and do not necessarily reflect those of the HS&DR programme, NIHR, NHS or the Department of Health.

Aims & Objectives

The aim is to provide evidence to inform decisions about which PROMs might be used to measure the health of children and young people with neurodisability to assess health outcomes of NHS care.

The objectives of the review are as follows:

1.1 To identify generic (i.e. not condition-specific) patient reported outcome measures (PROMs) used to measure the multidimensional health of children and young people under 18 years old.

1.2 To identify evidence of the psychometric properties and general performance of the generic PROMs when administered to measure the health of children and young people with neurodisability.

1.3 To critically appraise and compare the evidence identified in order to make recommendations about which generic PROMs may provide robust instruments for measuring NHS health outcomes.

Background / context

The UK Government is changing the way health services are commissioned. A new NHS Outcomes Framework will be one way of assessing whether the NHS is working effectively and efficiently for patients. The 'success' of NHS care will be judged partly from the reports of patients using special questionnaires called patient-reported outcome measures, or PROMS. [1]

Around 1 in 20 children in the UK are disabled. Amongst the most common childhood disabilities are cerebral palsy and autism, which are neurological conditions. These conditions are often grouped together under the umbrella of 'neurodisability'. Children affected by neurodisability are amongst the most frequent users of the NHS, and depend on the NHS to improve their health and wellbeing. A recent review of health and social care for children and young people recommended developing a shared vision between families and professionals for what health care is trying to achieve for children. [2] Health is multidimensional, and defined in this study by the World Health Organization's International Classification of Functioning Disability and Health (ICF). [3]

This research will identify which generic PROMs might be the most appropriate to assess the health of children with neurodisability as a potential patient-based indicator of NHS care. The systematic review set out here is one of three streams of related research; with the other two research streams comprising qualitative research with children and parents, and a Delphi survey with clinicians.

Definition of neurodisability
There is no established definition of 'neurodisability'. Therefore for the purposes of this project neurodisability is defined as follows:

- An impairment of functioning relating to any condition that affects the brain and nervous system. This may, for example, result in predominantly physical difficulties (such as cerebral palsy), learning and communication difficulties (such as autism), or other medical conditions (such as the problems associated with epilepsy). Sometimes it is difficult to label a child's condition with a specific diagnosis, however for the purposes of this project we are broadly inclusive. A child with a neurodisability may have greater difficulty, than is expected for their age, in mobility, lifting / carrying, manual dexterity, continence, communication, (speech) hearing, eye sight, memory, understanding, concentration, recognising danger.

This working definition was devised in consultation with paediatricians and parents of disabled children, and with reference to the Disability Discrimination Act.

Methods

The systematic review will follow the general principles published by the NHS Centre for Reviews and Dissemination, [4] and use widely accepted methods for appraising measurement properties, [5] and for assessing the methodological quality of papers that evaluate measurement properties. [6]

The review involves a two stage search; first we will systematically search for candidate instruments, second we will search for and identify those candidate instruments when used with children with neurodisability. Then we will use established standardised frameworks to appraise the evidence on how well each PROM performs when measuring the health of children with neurodisability.

Search Stage 1. Identifying candidate instruments

The purpose of stage 1 is to identify all generic multidimensional patient reported outcome measures (PROMs) used to measure the health of children and young people under 18 years old.

Search strategy
A search strategy has been designed to identify generic PROMs used to measure the health of children and young people under 18 years old. The search strategy for stage 1 has been developed with reference to the methodological filters published by the COSMIN group, [7] and the construct filters developed by the Oxford PROMs group. The strategy was further refined with input from all members of the systematic review team. Our strategy has been used to tailor the search across multiple bibliographic databases, with a focus on children and young people under 18 years.

An example of the search strategy (for MEDLINE) is shown in Appendix 1. The terms within each group will be combined with a Boolean OR command and are searched in combination using a Boolean AND command. Search terms are grouped as follows:

Group 1: generic names for measures (e.g., questionnaires, instruments, tools);
Group 2: health construct terms that are multidimensional (e.g., quality of life, health status);
Group 3: terms to describe children and young people (e.g., children, teenagers, adolescents);

In the piloting of these search strategies combining these three groups of terms produced 38,893 references. This output is judged unmanageable for the scope of this review, i.e. systematic screening would be too burdensome. Therefore we created a fourth set of terms (*Group 4*). This fourth grouping of terms is used to narrow the search to PROMS references that mention and refer to at least one form of evidence in terms of psychometric performance (i.e. aspects of validity or reliability) to narrow the search to identify papers only describing instruments that would be eligible (Appendix 1). This reduces the pilot search output to 6,519 records which is feasible to screen systematically. This narrower search, whilst less sensitive, is expected to be more specific and is judged to be both practical and feasible, identifies references that are more likely to be useful in the second stage of the search and subsequent review.

The search strategy will be run in the following databases; the terms will be modified as appropriate for each database:

- MEDLINE
- EMBASE
- CINAHL
- PsycINFO (via OvidSP)
- Oxford PROM bibliographic database
- ProQolid

A combination of controlled syntax (MeSH) and free-text terms will be used to search MEDLINE, EMBASE, CINAHL and PsycINFO. The search will also be limited to studies published from 1992 to the present and to English language publications only. Search results will be managed in reference management software and the dates of searches recorded so they may be updated as required. Duplicate references will be removed.

Inclusion/exclusion criteria

The inclusion criteria for stage 1 are patient reported outcome measures that are generic (i.e. not specific to any condition), questionnaires completed by a child and/or parent (or primary carer, in English language, used to measure the multidimensional health of children aged under 18 years. Measures used in economic evaluations are included. Sub-groups of children within this age-group are eligible. Interviewer-administered instruments are excluded; also excluded are instruments where the proxy respondent is not a parent or primary carer (e.g. clinicians, teachers).

Criteria	Specification
Population	Children and young people < 18 years old
Instruments	Generic patient reported outcome measures (PROMS) used in the English language; child self-report and parent (primary carer) reported measures are eligible.
Evidence	Indication of some testing/reporting of psychometric performance, such as aspects of validity or reliability.
Study design	Any type of study design
Date	1992 onwards
Language	English language

See table below for more detailed exclusion criteria.

Criteria	Specification
Exclusion criteria	▪ Any instrument that has not been used in a population of children (<18 years). ▪ Instruments that are administered by an interviewer. ▪ Any instrument for which an English language version has not been developed and used.

Screening for PROMs

Titles and abstracts of the articles resulting from this search will be screened independently by two reviewers to identify the names of instruments that meet the inclusion criteria. It is not envisaged that full texts will be obtained at this preliminary stage as only the names and acronyms used to describe instruments are required.

Data extraction and synthesis

The end of stage 1 will be a list of eligible candidate instruments. Experts in the field will be consulted to identify any additional PROMs meeting the inclusion criteria that were not identified by the search. Authors will be contacted to see if there are manuals for their instruments.

The constructs assessed by each candidate PROMs will be transcribed into lay terms to inform the qualitative work with families and professionals in the qualitative and Delphi survey work. The interpretability of definitions will be checked with members of the PenCRU Family Faculty.

The items and domains within each instrument will be mapped to codes within the World Health Organization's International Classification of Functioning Disability and Health (ICF), [3] broadly following the linking rules developed for this purpose. [9] The number and density of ICF domains represented within each instrument will be reported. Mapping will be performed by one reviewer and checked by a second, with disagreements being resolved by discussion.

Search Stage 2. Identifying evidence of PROMs performance in neurodisability

The purpose of stage 2 is to appraise evidence of the psychometric properties of the PROMS identified in stage 1, preferably but not exclusively when used with children and young people with neurodisability.

Search strategy
The search strategy will comprise blocks of terms to identify papers describing use of candidate PROMs, with children. Lines of searching will then (i) identify studies that are specifically set out to evaluate the psychometric properties in general populations, and (ii) identify use of PROMs with neurodisability. Both searches use the following groups of search terms:

Group 1: names and acronyms of the candidate instruments identified in stage 1;
Group 2: terms to describe children and young people (e.g., children, teenagers, adolescents);

2.1 To identify studies that evaluate psychometric properties of PROMs in general populations

Search term groups 1 and 2 plus a group of psychometric terms, see Appendix 2.

This search strategy will be deployed across the following bibliographic databases:

- MEDLINE (including Pre-Medline)
- EMBASE
- PsycINFO

Inclusion/exclusion criteria

Criteria	Specification
Population	General populations and neurodisability populations of children and young people < 18 years old.
Instruments	Generic patient reported outcome measures (PROMS) as listed as a result of Stage 1; child self-report and parent (primary carer) reported measures are eligible.
Evidence	Reporting of the psychometric performance of candidate PROMS: reliability, validity, responsiveness, precision, interpretability, acceptability & feasibility.
Study design	Studies specifically set up to evaluate psychometric properties with, primarily, participants in the English language. Cross-cultural studies are included if referencing an English language version of the instrument.
Date	1992 onwards
Language	English language

2.2 To identify use of PROMs with neurodisability

Search term groups 1 and 2 plus a group of neurodisability terms, see Appendix 3. Note that there is no comprehensive list of conditions that represent neurodisability. We propose using the exemplar conditions from the full protocol plus relevant MeSH and general terms.

This search strategy will be deployed across the following bibliographic databases:

- MEDLINE (including Pre-Medline)
- EMBASE
- CINAHL
- PsycINFO
- AMED
- NHS Database of Economic Evaluations

Inclusion/exclusion criteria
Studies will only be included if they report a generic PROM that is available in the English language and has been used in children (under the age of 18 years) with neurodisability. The table below provides a summary of the inclusion criteria.

Criteria	Specification
Population	Children and young people < 18 years old
Neurodisability	Any condition that affects the brain and nervous system including but not exclusively motor (e.g. cerebral palsy), neurological (e.g. epilepsy), or neuropsychiatric (e.g. autism) impairments.
Instruments	All instruments resulting from stage 1. Generic patient reported outcome measures (PROMS) used in the English language with children with neurodisability; child self-report and parent (primary carer) reported measures are eligible.
Evidence	Reporting of the psychometric performance of candidate PROMS, in which at least one of the following constructs has been evaluated: reliability, validity, responsiveness, precision, interpretability, acceptability & feasibility.
Study design	Any type of study design
Date	1992 onwards
Language	English language

We will carry out citation searches (to help to confirm saturation of our initial searches)

- Backwards citation chasing (1 generation) from included references
- 'Forwards' citation chasing on included references using citation databases (Science Citation Index/Social Science Citation Index)

The detailed strategies will be retained and recorded. The search results will be interrogated to ensure that key known 'marker' papers for known instruments are returned. Search results will be managed in reference management software and the dates of searches recorded so they may be updated as required.

Study selection

The titles and abstracts of articles resulting from the search will be screened independently by two reviewers, using the inclusion criteria described above. Discrepancies will be resolved by discussion. If either reviewer believes a paper is likely to yield evidence of the psychometric performance of candidate generic PROMs, the paper will be retrieved as full text.
Using the same methods, the retrieved papers will be assessed for inclusion in the review. Discrepancies will be resolved by discussion with a third reviewer, if necessary. Any duplicate papers will be recorded, double checked and excluded.

A PRISMA-style flow chart will be prepared to record each stage of the study selection process.

Data extraction and synthesis

Summary data for each included paper: author and year, name of candidate instruments, participants' characteristics (i.e., age etc.), study characteristics (i.e., study design, child or parent proxy respondent etc) will be extracted by one reviewer into a standardised, piloted data extraction form. For each instrument we will catalogue number of domains and number of items in instrument and domains. Data extraction will be checked by a second reviewer, with disagreements resolved by discussion with a third, if required.

For each instrument, studies will be categorised as either (i) those which specifically set out to evaluate psychometric properties of one or more PROMs, (ii) those in which a candidate PROM has been used in a trial or observational study, and performance indicators are reported incidentally.

Appraisal of evidence for each PROM

We will utilise the COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN) checklist to appraise the methodological quality of studies that were specifically designed to evaluate the psychometric performance of candidate instruments. [9] For each included paper, in which the psychometric properties of an instrument have been evaluated, the checklist will be administered by one reviewer and checked by a second. Any discrepancies will be resolved by discussion with the involvement of a third reviewer, if necessary.

We will use the appraisal framework developed by the Oxford PROMs group to appraise the psychometric performance and operational characteristics of each identified instrument in terms of reliability, validity, responsiveness, precision, interpretability, acceptability & feasibility (Appendix 2). [10] These performance indices will be appraised using a checklist by one reviewer be checked by a second reviewer.

Given the purpose of the review is to identify and recommend a generic PROM for children under 18 years, and with different neurodisability diagnoses, we will look particularly for evidence of group invariance across age groups and different conditions. This would indicate that valid comparisons can be made across age and diagnostic groups.

Once the evidence has been amassed for each instrument the properties will be scored according to the following scheme; 0 – not reported; - no evidence in favour; + some evidence in favour; ++ some good evidence in favour; +++ good evidence in favour.

References

1. Department of Health. (2010) The NHS Outcomes Framework 2011/12.
2. Kennedy I. (2010) Getting it right for children and young people: overcoming cultural barriers in the NHS so as to meet their needs. Department of Health.
3. World Health Organisation (2006) International Classification of Functioning, Disability and Health, Version for Children. Geneva: WHO.
4. NHS Centre for Reviews and Dissemination (2008) Systematic Reviews CRD's guidance for undertaking reviews in health care. Centre for Reviews and Dissemination, University of York www.york.ac.uk/inst/crd/index_guidance.htm (Accessed 20 Jan 2012)
5. Fitzpatrick R, Davey C, Buxton MJ, Jones DR. (1998) Evaluating patient-based outcome measures for use in clinical trials. Health Technol Assess. 2(14):1-74.
6. Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, Bouter LM, de Vet HC. (2010) The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. Qual Life Res. 19(4):539-49.
7. Terwee CB, Jansma EP, Riphagen II, de Vet HCW. (2009) Development of a methodological PubMed search filter for finding studies on measurement properties of measurement instruments. Qual Life Res 18:115-1123.
8. Cieza A, Geyh S, Chatterji S, Kostanjsek N, Ustun B, Stucki G. ICF Linking Rules: An Update Based on Lessons Learned. J. Rehabil. Med. 2005; 37: 212-8.
9. Terwee CB, Bot SD, de Boer MR, van der Windt DA, Knol DL, Dekker J, Bouter LM, de Vet HC. (2007) Quality criteria were proposed for measurement properties of health status questionnaires. J Clin Epidemiol. 60(1):34-42.
10. Jenkinson C, Gibbons E, Fitzpatrick R (2009) A structured review of patient-reported outcome measures in relation to stroke. Report for the UK Department of Health.

APPENDICES

Appendix 1: Search Strategy Phase 1 (used in MEDLINE and adjusted for other databases)

Database: Ovid MEDLINE(R) In-Process & Other Non-Indexed Citations and Ovid MEDLINE(R) <1946 to Present>

Search Strategy:

---

1 "outcome assessment (Health Care)"/ (40965)
2 tool*.ti,ab. (308785)
3 instrument*.ti,ab. (157950)
4 questionnaire*.ti,ab. (243679)
5 index.ti,ab. (373387)
6 indices.ti,ab. (90301)
7 scale*.ti,ab. (357043)
8 survey*.ti,ab. (328860)
9 feedback.ti,ab. (68589)
10 interview*.ti,ab. (183532)
11 (outcome* adj2 measure*).ti,ab. (125415)
12 (outcome* adj2 assessment*).ti,ab. (4843)
13 PROMS.ti,ab. (73)
14 (measur* adj2 (quality or health or outcomes)).ti,ab. (28899)
15 (assess* adj2 (quality or health or outcomes)).ti,ab. (42576)
16 (patient report* adj2 outcome* adj2 (measure* or assessment*)).ti,ab. (647)
17 (self report* adj2 outcome* adj2 (measure* or assessment*)).ti,ab. (649)
18 (parent report* adj2 outcome* adj2 (measure* or assessment*)).ti,ab. (19)
19 (child report* adj2 outcome* adj2 (measure* or assessment*)).ti,ab. (1)
20 (patient assess* adj2 outcome* adj2 (measure* or assessment*)).ti,ab. (34)
21 (self assess* adj2 outcome* adj2 (measure* or assessment*)).ti,ab. (30)
22 (parent assess* adj2 outcome* adj2 (measure* or assessment*)).ti,ab. (0)
23 (child assess* adj2 outcome* adj2 (measure* or assessment*)).ti,ab. (0)
24 (carer assess* adj2 outcome* adj2 (measure* or assessment*)).ti,ab. (0)
25 (caregiver assess* adj2 outcome* adj2 (measure* or assessment*)).ti,ab. (0)
26 or/1-25 (1873850)
27 "quality of life"/ (96741)
28 quality of life.ti,ab. (115639)
29 QOL.ti,ab. (14551)
30 HRQOL.ti,ab. (5283)
31 QL.ti,ab. (964)

32 HRQL.ti,ab. (1898)
33 health utilit*.ti,ab. (840)
34 health outcomes.ti,ab. (13863)
35 patient outcome*.ti,ab. (18435)
36 (patient reported adj2 outcome*).ti,ab. (2050)
37 (self reported adj2 outcome*).ti,ab. (985)
38 (parent reported adj2 outcome*).ti,ab. (39)
39 (proxy reported adj2 outcome*).ti,ab. (2)
40 (child* adj3 outcome*).ti,ab. (9788)
41 (patient assessed adj2 outcome*).ti,ab. (39)
42 (self assessed adj2 outcome*).ti,ab. (42)
43 (parent assessed adj2 outcome*).ti,ab. (0)
44 ((health or functional) adj status).ti,ab. (44722)
45 (well being or wellbeing).ti,ab. (35861)
46 functioning.ti,ab. (95250)
47 activit*.ti,ab. (1958455)
48 participation.ti,ab. (78838)
49 or/27-48 (2307513)
50 child*.ti,ab. (868046)
51 infant*.ti,ab. (272309)
52 (young adj people).ti,ab. (13148)
53 (pediatric or paediatric).ti,ab. (158091)
54 adolescen*.ti,ab. (143244)
55 teenager*.ti,ab. (9186)
56 or/50-55 (1215533)
57 reliab*.ti,ab. (257762)
58 valid*.ti,ab. (311006)
59 evaluation.ti,ab. (685686)
60 repeatability.ti,ab. (11261)
61 acceptability.ti,ab. (14430)
62 responsiveness.ti,ab. (72761)
63 feasibility.ti,ab. (78302)
64 psychometric.ti,ab. (19424)
65 57 or 58 or 59 or 60 or 61 or 62 or 63 or 64 (1303060)
66 26 and 49 and 56 and 65 (8557)
67 limit 66 to (english language and yr="1992 -Current") **(7253)**

****

Appendix 2: Search Strategy Phase 2.1, to identify evidence of psychometric performance of candidate PROMs (used in MEDLINE and adjusted for other databases).

Database: Ovid MEDLINE(R) In-Process & Other Non-Indexed Citations and Ovid MEDLINE(R) <1946 to Present>
Search Strategy:
---

1 child/ (1273968)
2 child*.ti,ab. (891994)
3 adolescent/ (1484571)
4 adolescent*.ti,ab. (133832)
5 infant/ (596913)
6 infant*.ti,ab. (279971)
7 1 or 2 or 3 or 4 or 5 or 6 (2675732)
8 reliab*.ti,ab. (266400)
9 valid*.ti,ab. (326487)
10 responsive*.ti,ab. (151564)
11 evaluation.ti,ab. (706297)
12 repeatab*.ti,ab. (16575)
13 feasib*.ti,ab. (147155)
14 acceptab*.ti,ab. (93849)
15 psychometric.ti,ab. (20216)
16 8 or 9 or 10 or 11 or 12 or 13 or 14 or 15 (1533356)

For each candidate PROM: 16 AND [Name of measure, including variants & acronyms]

****

Appendix 2: Search Strategy Phase 2.2, to identify use of candidate PROMs with neurodisability (used in MEDLINE and adjusted for other databases).

Database: Ovid MEDLINE(R) In-Process & Other Non-Indexed Citations and Ovid MEDLINE(R) <1946 to Present>
Search Strategy:

---

1 child/ (1273968)
2 child*.ti,ab. (891994)
3 adolescent/ (1484571)
4 adolescent*.ti,ab. (133832)
5 infant/ (596913)
6 infant*.ti,ab. (279971)
7 1 or 2 or 3 or 4 or 5 or 6 (2675732)
8 exp Nervous System Diseases/ (1896492)
9 exp Autistic Disorder/ (14556)
10 exp Neurologic Manifestations/ (724201)
11 exp cerebral palsy/ (14344)
12 (cerebral adj palsy).ti,ab. (13034)
13 epilep*.ti,ab. (86738)
14 exp autism/ (14556)
15 autis*.ti,ab. (18478)
16 (neuro-motor adj disease*).ti,ab. (2)
17 (neuromotor adj disease*).ti,ab. (18)
18 (neuromotor adj disorder*).ti,ab. (50)
19 (neuromotor adj dysfunction*).ti,ab. (59)
20 neurodisabilit*.ti,ab. (82)
21 (neuropsychiatric adj disease*).ti,ab. (987)
22 (neuropsychiatric adj dysfunction*).ti,ab. (63)
23 ((Child* or infant* or adolescen*) adj4 disab*).ti,ab. (8404)
24 or/8-23 (1940912)
25 7 and 24 (475944)

For each candidate PROM: 25 AND [Name of measure, including variants & acronyms]

****

Appendix 4: Appraisal Framework - taken from Jenkinson et al (2009) [8]

Appraisal Component	Definition/test	Criteria for acceptability
Reliability		
Reproducibility/Test retest reliability	The stability of a measuring instrument over time; assessed by administering the instrument to respondents on two different occasions and examining the correlation between test and re-test scores	Test re-test reliability correlations for summary scores 0.70 for group comparisons
Internal Consistency	The extent to which items comprising a scale measure the same construct (e.g. homogeneity of items in a scale); assessed by Cronbach's alpha's and item-total correlations	Cronbach's alphas for summary scores ≥0.70 for group comparisons Item-total correlations ≥ 0.20
Validity		
Content Validity	The extent to which the content of a scale is representative of the conceptual domain it is intended to cover; assessed qualitatively during the questionnaire development phase through pre-testing with patients. Expert opinion and literature review	Qualitative evidence from pre-testing with patients, expert opinion and literature review that items in the scale represent the construct being measured Patients involved in the development stage and item generation
Construct Validity	Evidence that the scale is correlated with other measures of the same or similar constructs in the hypothesised direction; The ability of the scale to differentiate known-groups; assessed by comparing scores for sub-groups who are expected to differ on the construct being measured (e.g a clinical group and control group) assessed on the basis of correlations between the measure and other similar measures	High correlations between the scale and relevant constructs preferably based on a priori hypothesis with predicted strength of correlation. Statistically significant differences between known groups and/or a difference of expected magnitude

Responsiveness	The ability of a scale to detect significant change over time; assessed by comparing scores before and after an intervention of known efficacy (on the basis of various methods including t-tests, effect sizes (ES), standardised response means (SRM) or responsiveness statistics	Statistically significant changes on scores from pre to post treatment and/or difference of expected magnitude. The recommended index of responsiveness is the effect size, calculated by subtracting the baseline score from the follow up score and dividing by the baseline SD. Effect sizes can be graded as small (<0.3), medium (~0.5), or large (>0.8).
Floor/ceiling effects	The ability of an instrument to measure accurately across full spectrum of a construct	Floor/ceiling effects for summary scores <15%
Practical Properties		
Acceptability	Acceptability of an instrument reflects' respondents' willingness to complete it and impacts on quality of data	Low levels of incomplete data or non-response
Feasibility/burden	The time, energy, financial resources, personnel or other resources required of respondents or those administering the instrument	Reasonable time and resources to collect, process and analyse the data.