Systematic review: Measurement properties of patient reported outcome measures evaluated with childhood brain tumor survivors or other acquired brain injury

Kim S. Bull[1], Samantha Hornsey[2], Colin R. Kennedy[1], Anne-Sophie E. Darlington[3], Martha A. Grootenhuis[4], Darren Hargrave[5,6], Christina Liossi[7,8] Jonathan P. Shepherd[9], David A. Walker[10], Christopher Morris[11]

[1]Clinical and Experimental Sciences, University of Southampton, Southampton, SO16 6YD, UK

[2]Primary Care, Population Sciences, and Medical Education, University of Southampton, Southampton, SO16 6YD, UK

[3]Health Sciences, University of Southampton, Southampton, SO16 6YD, UK

[4]Psychosocial Research and Healthcare Innovation, Princess Máxima Centre for Paediatric Oncology, 3584 CS Utrecht, NL

[5]UCL Great Ormond Street Institute of Child Health, University College London, London, WC1N 1EH, UK (DH)

[6]Paediatric Oncology, Great Ormond Street Hospital for Children NHS Foundation Trust, London, WC1N 3JH, UK (DH)

[7]Psychology, University of Southampton, Southampton, SO17 1BJ, UK (CL)

[8]Pain Control Service, Great Ormond Street Hospital for Children NHS Trust, WC1N 3JH, UK (CL)

[9]Southampton Health Technology Assessments Centre, University of Southampton, Southampton, SO16 7NS, UK

[10]Children's Brain Tumour Research Centre, University of Nottingham, Nottingham, NG7 2UH, UK

[11]The Peninsula Childhood Disability Research Unit, University of Exeter, Exeter, EX1 2LU, UK

Running title: Systematic review PROMs childhood ABI

Corresponding author: Dr Kim Bull, Department of University Child Health, MP803 G level Centre Block, Southampton General Hospital, Southampton SO16 6YD, UK. Email: k.s.bull@southampton.ac.uk Tel: 0044 2381203980

Co-first authorship will be shared between Kim Bull and Samantha Hornsey

Funding

This research was supported by The Brain Tumour Charity

Conflict of Interest

None declared by any of the authors

Total manuscript word count: 5462 (including words in references, all text sections, abstract, and figure legends, but not supplementary text)

Abstract

Background

Survivors of childhood brain tumors or other acquired brain injury (ABI) are at risk of poor health-related quality of life (HRQoL); its valid and reliable assessment is essential to evaluate the effect of their illness on their lives. The aim of this review was to critically appraise psychometric properties of patient-reported outcome measures (PROMs) of HRQoL for these children, to be able to make informed decisions around selection of the most suitable PROM for use in clinical practice.

Methods

We searched MEDLINE, EMBASE, and PsycINFO for studies evaluating measurement properties of HRQoL PROMs in children treated for brain tumors or other ABI. Methodological quality of relevant studies was evaluated using the COSMIN checklist.

Results

Eight papers reported measurement properties of four questionnaires: Health Utilities Index (HUI), PedsQL Core and Brain Tumor Modules, and Child and Family Follow-up Survey (CFFS). Only the CFFS had evidence of content and structural validity. It also demonstrated good internal consistency whereas both PedsQL modules had conflicting evidence regarding this. Conflicting evidence regarding test-retest reliability was reported for HUI and PedsQL Core Module only. Evidence of measurement error/precision was favorable for HUI and CFFS and absent for both PedsQL modules. All four PROMs had some evidence of construct validity/hypothesis testing but no evidence of responsiveness to change.

Conclusions

Valid and reliable assessment is essential to evaluate impact of ABI on young lives. However, measurement properties of PROMs evaluating HRQoL appropriate for this population require further evaluation, specifically construct validity, internal consistency, and responsiveness to change.

Keywords: systematic review, patient-reported outcomes, acquired brain injury, brain tumor, children

Introduction

One child in every 600 will develop some form of cancer by 16 years of age[1] and around 20% to 27% of these children will have a brain tumor[2]. Currently, 65.4% of children diagnosed with a brain tumor in Europe from 1999-2007 are reported to survive 5 or more years from diagnosis[3] and the majority should have prolonged survival and become adults. They often have multiple impairments and reduced health-related quality of life (HRQoL)[4-8]. Approximately 62% will be left with a life-altering long-term disability[9] comparable to the life-changing sequelae of severe traumatic or other acquired childhood brain injuries (ABI). ABI is post-natal injury to the brain that is sudden in onset and may be the result of head trauma or non-traumatic, following meningitis, stroke, metabolic derangement, sickle cell disease, or a brain tumor.

In children aged less than 16 years, the incidence of hospitalization for traumatic brain injury (TBI) has been reported to be between 280 and 500 per 100,000. This implies that the total number of children admitted to hospital for TBI per annum in the UK is at least 35,000. Of these, about 2,000 (5.7%) will have severe TBI, 3,000 (8.6%) moderate TBI, and 30,000 (85.7%) mild TBI. In addition, the total number of children who sustain non-traumatic coma associated with severe or moderate encephalopathy is around 4,000 per year[10]. Also, the Central Brain Tumor Registry of the United States reported an incidence rate of newly diagnosed cases of brain tumor in children to be 5.54 per 100,000, equating to 4,500 new cases annually[11] and the overall annual incidence of childhood stroke has been estimated to be around 1.2 to 13 cases per 100,000 children under the age of 18 years[12].

In the context of delivery of clinical care, doctors vary in their ability to explore, elicit, and respond to information about HRQoL[13] and discussion of the emotional, social, and cognitive issues impacting HRQoL after ABI or childhood cancer does not routinely take place in clinic consultations[14]. In addition, children and parents are often reluctant to raise psychosocial issues at clinic appointments[15,16] which they perceive to be more focused on medical issues such as monitoring tumor status and its response to anti-tumor treatments or complications of other types of ABI.

PROMs measure a patient's health status or health-related quality of life at a single point in time, and are collected through short, self-completed questionnaires[17] without any third party acting as an intermediary. In the context of clinical research, the use of PROMs, including those assessing HRQoL, has proved to be a practicable means of assessing quality of survival in multicenter treatment trials[18,19]. Individualized use of PROMs in the routine care of children with a long-term illness has the potential to add valuable information about the impact of the disease, inform treatment planning, provide clinicians with timely information about a patient's functional and emotional status and wellbeing[20], and enhance family-clinician communication[21]. This helps clinical staff to deliver care focused on the needs and choices of each individual child and family[22]. Such use of PROMs has been evaluated in large groups of typically developing children, adolescents and adults and in adult patients with cancer[23] and children with other long-term conditions[24-27] but not in child/adolescent survivors of brain tumor or other ABI.

When selecting PROMs for a specific purpose, it is necessary to examine how robust (valid and reliable) is the measurement of HRQoL produced by such questionnaires. A number of methodological approaches are available to determine aspects of reliability and validity[28]. The aim of the present systematic review was to critically appraise the psychometric properties of patient-reported outcome measures (PROMs) of HRQoL for these children, in order to be able to make informed decisions around selection of the most suitable PROM for use in clinical practice.

Materials and Methods

Systematic Review

We undertook a systematic review of published evidence relating to the measurement properties of PROMs in children with brain tumors and other ABI and the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement[29]. A protocol was written which specified, a priori, the inclusion criteria and methods to be used. We also used methods recommended for appraising measurement properties and for assessing the methodological quality of papers that evaluate PROMs[30], including the consensus-based standards for the selection of health status measurement instruments (COSMIN) checklist for evaluation of publications [31].

Search strategy

The search strategy was designed by an experienced information specialist (see acknowledgements) in discussion with topic experts (KB, CK, and CM) and an experienced systematic reviewer (JS). Blocks of search terms were combined including variants of 'brain tumor/acquired brain injury', 'child/adolescent', 'patient reported outcome measure', 'psychometric' and the titles of generic PROMs suitable for use in all children or in all children with long-term health conditions, as listed in the most recent systematic review focusing on HRQoL in children with disabilities[32].

MEDLINE, EMBASE, and PsycINFO were searched for studies published from 1992 onwards in peer-reviewed journals whose purpose was to evaluate measurement properties of PROMs. An example from MEDLINE of this search strategy is shown in Appendix 1. The electronic searches were completed on 7th February 2017 and updated on 28th May 2019. Publication details were uploaded into an Endnote reference management database and duplicates removed. Backwards citation chasing (one generation) from the reference lists of included papers was conducted by CM. Forwards citation chasing for each included study using all databases in the Web of Science cited reference search resource was conducted by SH.

Inclusion and exclusion criteria

We sought published papers reporting evaluations of the measurement properties of multi-dimensional child (aged 5 to 18 years) self-report and/or parent-proxy report PROMs assessing health and wellbeing in children receiving care either for a brain tumor or other ABI of any kind (rather than for specific types of brain tumors or ABI). Evaluation of an English language version of the PROM was a requirement for inclusion. Studies in which only part of the sample were eligible for review were included only if psychometric analyses had been conducted on the eligible sub-groups within the sample. Instruments administered by an interviewer and single domain-specific questionnaires (e.g. to assess only depression, fatigue, or pain) were excluded.

Study selection

An inclusion/exclusion criteria decision chart was used to aid the selection of articles likely to yield relevant results from their titles and abstracts. The use of this chart was piloted by SH and KB who screened the first 10 articles together to test agreement over inclusion of articles. All remaining titles and abstracts were screened in batches of 40 by SH and, independently, by KB. The evaluations of each batch of 40 by the two reviewers were then compared and any disagreements discussed and resolved. Full texts were then retrieved from this list of potential studies by SH. KB then checked the list of included and excluded studies to confirm agreement. Disagreements were discussed and resolved between the reviewers.

Data extraction, appraisal, and synthesis of included studies

Descriptive characteristics of included studies and measurement properties of the PROMs were extracted by SH. These extracted data were checked by KB and the final extracted data set was agreed in discussion with CM. The criteria of Fitzpatrick et al. (1998)[33] were adopted for evaluation of the patient-based outcome measures within the extracted data set.

The COSMIN Risk of Bias checklist of Consensus-based standards for the selection of health status measurement instruments (COSMIN) was used to assess the methodological quality of the included studies. The checklist is comprised of 12 boxes which together cover three domains: content validity, internal structure, and remaining measurement properties – namely reliability, measurement error, criterion validity, hypothesis testing for construct validity, and responsiveness to change[31]. Ten of the 12 boxes can be used to assess whether a study meets standards for good methodological quality and 9 of them contain standards for the included measurement properties. These are each scored on a four-point rating scale of the way in which each measurement property was assessed.

All of the above properties were assessed (Table 1) excepting cross-cultural validity which was not relevant as our search only included English language reports. Criterion validity was not applicable as in the case of HRQoL there is no criterion against which HRQoL measures can be judged (except for the purpose of comparing long versions of an instrument and shortened forms of the same instrument).

An overall score for the methodological quality of a study was determined by CM for each measurement property separately as a single rating[34], arrived at by taking the lowest rating of any of the items in a box[35]. The review team then considered the evidence for each PROM and summarized in a single rating for each measurement property following methods commonly used for presentation of findings against the COSMIN criteria (Table 2). From these ratings conclusions were drawn on the extent to which each PROM could be considered robust for measuring HRQoL in children treated for brain tumors or other ABI.

Table 1 about here**.**

Table 2 about here

Results

The electronic searches resulted in 472 articles after the removal of duplicates. Of these, 374 were excluded leaving 98 potentially relevant studies whose full text articles were retrieved. Screening of these led to the exclusion of a further 90 papers leaving eight studies remaining for evaluation (Fig. 1). Backwards citation chasing identified two potentially relevant papers and forward citation chasing identified six potentially relevant papers, all of which were subsequently excluded due to inappropriate population (n=4), inappropriate instrument (n=3), or lack of relevant data (n=1).

Four self-report and/or parent-proxy report PROMs – the Health Utilities Index (HUI), the Pediatric Quality of Life Inventory Core Module (PedsQL), the PedsQL Brain Tumor Module, and the Child and Family Follow-Up Survey (CFFS) - were evaluated and appraised in the eight included studies (Tables 3&4) and these are briefly described here.

The HUI and PedsQL are generic measures of HRQoL whereas the PedsQL Brain Tumor Module and the CFFS are disease-specific. The HUI is a rating scale used to measure general health status with one question relating to HRQoL. Health utility values are commonly produced using HUI as a component

of the quality-adjusted life years (QALY) calculation used in population health and economics. Answers to 15 questions about health state, scored at 3 to 6 health status levels, can be grouped in two different ways to produce either HUI2 or HUI3 scores across 7 or 8 'attributes' of health. HUI3, for example, groups health status levels to create attribute scores for Vision, Hearing, Speech, Ambulation, Dexterity, Emotion, Cognition, and Pain.

The PedsQL is a measure of HRQoL with 23 questions across four core scales: Physical, Emotional, Social, and School. The 24-item PedsQL Brain Tumor Module was designed to measure HRQoL in children undergoing treatment for a brain tumor. The questions are divided between six subscales: Cognitive problems, pain and hurt, movement and balance, procedural anxiety, nausea, and worry.

The CFFS was developed as a parent report measure to monitor needs and outcomes of children and youth with acquired brain injury and their families. It consists of five sections with a total of 71 closed or open-ended questions. Section 1 asks about the child's physical and emotional health and well-being, primary way of moving around and communicating, and medical problems or hospitalizations within the last year or since leaving the rehabilitation program. Section 2 includes the Child and Adolescent Scale of Participation (CASP) and three subsequent open-ended questions about equipment, modifications or strategies that are used to promote the child's participation. Section 3 includes the Child and Adolescent Factors Inventory (CAFI) and Child and Adolescent Scale of Environment (CASE) and a question about health or medical restrictions on the child's daily activities. Section 4 enquires about the child's current educational placement, rehabilitation and health services, satisfaction with services, the family's quality of life, and current services and needs. Finally, Section 5 seeks suggestions to improve services at the program from where the child was discharged to better address the needs of the child and family and additional information that was not addressed in the CFFS.

Completion time for the HUI and the PedsQL (core or brain tumor module) is about 5 minutes and for the CFFS about 30 minutes. Child self-report is available from 5 years old for the PedsQL

modules and from 12 years old for the HUI while the CFFS is available as parent-report only (Table 4). None of the studies had assessed all psychometric properties of the PROM in question.

Content validity: this had been assessed only for the CFFS and in this case the evidence for its validity was good.

Internal structure: only the CFFS had been assessed for evidence of structural validity and there was good evidence that it possessed this property. Internal consistency had been evaluated for the CFFS (good evidence) and for the PedsQL Core and PedsQL Brain Tumor Modules (equivocal evidence) but not for the HUI (Supplementary Table S1and Table 5).

Other measurement properties: evidence for test/re-test reliability and proxy reliability was available but conflicting for the HUI and PedsQL Core module and absent for the PedsQL Brain Tumor module and the CFFS. Favorable evidence of precision was available for the HUI but absent for the PedsQL Core and Brain Tumor Modules or the CFFS. Favorable evidence of hypothesis testing/construct validity was available for all measures. There was no evidence of responsiveness to change over time for any of the PROMs.

The methodological quality of the included studies varied from adequate to very good (Supplementary Table S2). The CFFS had had the most measurement properties evaluated and these studies were of high quality (Supplementary Table S2).

Table 3 about here

Table 4 about here

Table 5 about here

Discussion

This is the first systematic review of evaluations of the psychometric properties of PROMs in survivors of childhood brain tumors and other ABI of childhood. It identified only eight papers describing four

PROMs with relevant information about their measurement properties in children treated for brain tumors or ABI. Some evidence in favor of each instrument was found with respect to those properties that had been examined but caution is needed with respect to those properties that have not been evaluated: notably content and structural validity for the HUI and the PedsQL; test/retest reliability and precision / measurement error for the PedsQL; and responsiveness to change over time for all measures. In contrast to the HUI and the CFFS, the self-report versions of the two PedsQL modules had been specifically designed for the pediatric age group.

The PedsQL Core Module has previously been reported, in the setting of orthopedic, and rheumatology clinics, to be sensitive to increasing disease severity, responsive to clinical change over time, and to demonstrate impact on clinical decision-making resulting in increases in HRQoL[36]. The developer of the PedsQL has recommended it as a screening instrument to use in conjunction with disease-specific modules to target symptoms for interventions[37].

Our strict selection criteria did not reveal any longitudinal/follow-up studies in which responsiveness to change may have been assessed incidentally but the present study does not rule out their existence. Assessing the size of meaningful change above measurement error of the scores from PROMs is desperately needed from further research. It therefore behoves the user to design validation steps when adopting one of the questionnaires for clinical or research use to plug this evidence gap, for example when interpreting studies that have used these questionnaires to measure change.

The validity of the use of a PROM to communicate with families and better focus their care to improve their HRQoL depends on the method by which it was developed. This method of development of a PROM is to an extent separate from its measurement properties although may be reflected in measures of content validity. These methods have been highly variable and are often not clearly specified. Thus, there would be merit in discussing further with survivors of brain tumor or other ABI and their caregivers the salience and relevance of the individual questions within questionnaires and relying on responses to individual questions rather than questionnaire scores as a means to enhance communication between care-providers and service users about HRQoL. Such discussion with survivors of brain tumors or other ABI in

childhood would also help to identify whether or not there is a need to develop a condition-specific PROM for use in child and adult survivors of brain tumor or other ABI in childhood.

Two systematic reviews of HRQoL measures in children with long-term conditions other than ABI seem to have particular relevance to selection for use in child survivors of brain tumors or other ABI. The first conducted was a systematic review of the psychometric properties of measures for use in children with neurodisability[32,38]. It found evidence relating to measurement properties of seven generic PROMs (The Child Health and Illness Profile, The Child Health Questionnaire, the Child Quality of Life questionnaire, KIDSCREEN, the PedsQL, the Student Life Satisfaction Scales, and the Youth Quality of Life Instrument), two chronic-generic PROMs (the DISABKIDS and the Neurology Quality of Life Measurement System) and three preference-based measures (HUI, the EQ-5D-Y, and the Comprehensive Health Status Classification system – Preschool). In the instance of preference-based measures, they noted a dearth of evidence of face, content, and construct validity, or test-retest reliability and for all measures, a lack of evidence for responsiveness and measurement error.

The second systematic review was of PROMs of 'cancer-specific' HRQoL measures for use in children with cancer and identified nine measures for proxy completion, of which six had parallel measures for self-completion by children[39]. This review did not consider generic scales that had been applied in children with cancer (e.g. the PedsQL Core Module) but did note that the MMQL-UK child and parent versions have been validated as generic measures of QoL that can be used with healthy children and those with chronic conditions other than cancer. Adequate detail about how questionnaire items were generated from qualitative interviews was provided for only four questionnaires and most did not combine this with literature review or expert opinion. Some questionnaires required further psychometric evaluation before they could be recommended leaving just five recommendable measures: the Miami Pediatric Quality of Life Questionnaire (MPQS), the Minneapolis-Manchester Quality of Life Instrument (MMQL), the PedsQL Cancer Module, the Pediatric Functional Assessment of Cancer Therapy-Childhood: Brain Tumor Survivor (PFACT-BT), and the Pediatric Oncology Quality of Life Scale (POQOLS) (ibid.). These questionnaires may be suitable for clinical use in children receiving care

for a brain tumor or other ABI but, with the exception of the PFACT-BT, their measurement properties and performance have not been evaluated in either of these groups. The PFACT-BT is administered by an interviewer. This was an exclusion criterion for the present review and unfortunately also greatly limits the applicability of this measure.

Advantages of self-administered questionnaires include the reduction in burden associated with respondents of being able to answer at their own convenience and in their own time, the obviation of any need for a trained administrator, and, when done on line, the avoidance of transcription errors and greater efficiency and of data being entered at the moment that it is self-administered. However, the development of the questionnaires needs to be robust since measurement error may be made more likely by the absence of a trained administrator, if questions are poorly worded or formatted.

However, other considerations relating to the constraints of health care systems, including time and resources, need to be taken into account. Not all the PROMs we identified are suitable for systematic use in an outpatient clinical health care setting. PROMs with costly licensing fees are not feasible to use in public health care systems where funds are limited. Also PROMs which are lengthy to discuss will not be adopted due to clinical time constraints. PROMs also need to be relevant and suitable for follow-up consultations after treatment has ended. The CFFS appears to be the most thoroughly developed and comprehensive measure in this population but it is lengthy, at 71 questions, and the absence of any self-report version is a limitation of its use as a measure of quality of life. For these reasons, the PedsQL – Core Module, which is being widely used in childhood cancer research, may be the most suitable PROM for use in a clinical setting, notwithstanding the gaps in evidence regarding some of its psychometric properties.

Strengths of the present review include a comprehensive and systematic search strategy, use of standard criteria for the evaluation of the measurement properties of each PROM, and use of defined criteria to measure the quality of the studies that had been undertaken to assess these properties in participants with brain tumors or acquired brain injury in childhood. Synthesis of the findings of this review with the findings of previous reviews relating to children with other long-term conditions is also a

strength. The restriction of the systematic review to evaluations of questionnaires in the English language is both a limitation of this study, in that it restricts its relevance to English speaking service users, and a strength in that issues of cross-cultural validity apply to a much smaller extent than would be the case for an evaluation of instruments in more than one language[40].

In summary, both the present systematic review of measurement properties of PROMs when used in child survivors of brain tumors or other ABI and the preceding systematic reviews of PROMs when used in survivors of childhood cancer and in children with neurodisability indicate lack of evidence regarding measurement error or responsiveness to change and, in the case of preference-based measures, lack of evidence of content or construct validity, or test-retest reliability. Factors contributing to this lack of evidence may include the assumption by investigators that psychometric properties shown in healthy populations also apply to survivors of brain tumors, difficulty of accessing study populations of sufficient size to reach reliable conclusions about the validity of measures used, and/or limited awareness of investigators about the importance of validating psychometric properties of those measures.

To conclude, the four PROMs that were identified in our systematic review and a handful of other PROMs identified in previous systematic reviews of child survivors of non-CNS cancers and of children with neuro-disability had some evidence of favorable measurement properties but this was limited and insufficient to enable selection of PROMs suitable for use in survivors of childhood brain tumors or other ABI, particularly for the measurement of change. For communication about HRQoL, the paucity of evidence of content validity in these groups suggests the need for further discussion with these patient groups to inform selection of questions that address their concerns and we are, to that end, currently engaged in a qualitative study of the expressed views of brain tumor survivors. In the meantime there is clearly a need for studies that evaluate the measurement properties of those generic PROMs of HRQoL when used in these patients whether the purpose is to inform the care of individuals or to describe the HRQoL of groups of patients.

Funding

The Brain Tumour Charity Quality of Life Award (GN-000366).

Acknowledgements

We would like to acknowledge the help of Karen Welch who conducted the systematic search on our behalf. Also thanks goes to Sasja Schepers who commented on a draft of the systematic review protocol on which the methods used were based.

References

1. Stiller CA. *Childhood cancer in Britain: Incidence, survival, mortality*. Oxford: Oxford University Press; 2007.

2. Baade PD, Youlden DR, Valery PC, et al. Trends in incidence of childhood cancer in Australia, 1983-2006. *Br J Cancer* 2010;102(3):620-626.

3. Gatta G, Botta L, Rossi S, et al. Childhood cancer survival in Europe 1999-2007: results of EUROCARE-5--a population-based study. *Lancet Oncol* 2014;15(1):35-47.

4. Packer RJ, Gurney JG, Punyko JA, et al. Long-term neurologic and neurosensory sequelae in adult survivors of a childhood brain tumor: childhood cancer survivor study. *J Clin Oncol* 2003;21(17):3255-3261.

5. Boman KK, Hoven E, Anclair M, et al. Health and persistent functional late effects in adult survivors of childhood CNS tumours: a population-based cohort study. *Eur J Cancer* 2009;45(14):2552-2561.

6. Cardarelli C, Cereda C, Masiero L, et al. Evaluation of health status and health-related quality of life in a cohort of Italian children following treatment for a primary brain tumor. *Pediatr Blood Cancer* 2006;46(5):637-644.

7. Boman KK, Lindblad F, Hjern A. Long-term outcomes of childhood cancer survivors in Sweden: A population-based study of education, employment, and income. *Cancer* 2010;116(5):1385-1391.

8. Bull KS, Liossi C, Culliford D, et al. Child-related characteristics predicting subsequent health-related quality of life in 8- to 14-year-old children with and without cerebellar tumors: a prospective longitudinal study. *Neuro-Oncol Pract* 2014;1(3):114-122.

9. Macedoni-Luksic M, Jereb B, Todorovski L. Long-term sequelae in children treated for brain tumors: Impairments, disability, and handicap. *Pediatr Hematol Oncol* 2003;20(2):89-101.

10. NHS England. *2013/14 NHS standard contract for paediatric neurorehabilitation*. https://www.england.nhs.uk/wp-content/uploads/2018/09/Paediatric-Neurorehabilitation.pdf. Accessed May 5, 2019

11. Ostrom QT, Gittleman H, Liao P, et al. CBTRUS Statistical Report: Primary brain and other central nervous system tumors diagnosed in the United States in 2010-2014. *Neuro-Oncol* 2017;19:V1-V88.

12. Tsze DS, Valente JH. Pediatric stroke: a review. *Emerg Med Int* 2011;2011:734506.

13. NICE. *Children and young people cancer survivorship initiative: Improving quality and the patient experience*. http://www.nice.org.uk/savingsAndProductivityAndLocalPracticeResource?ci=http%3a%2f%2farms.evidence.nhs.uk%2fresources%2fQIPP%2f844812%3fniceorg%3dtrue. Accessed May 5, 2019.

14. Engelen V, van Zwieten M, Koopman H, et al. The influence of patient reported outcomes on the discussion of psychosocial issues in children with cancer. *Pediatr Blood Cancer* 2012;59(1):161-166.

15. Sibelli A. *Quality of Life following Treatment for a Brain Tumour: child and parent perspectives*. MSc. Southampton, 2011.

16. King T. *A Qualitative Exploration of Longitudinal Quality of Life and Illness Experience in Childhood LGCA Brain Tumour Survivors: From a Child's Perspective*. MSc. Southampton, 2012.

17. . *Patient Reported Outcome Measures (PROMs)*. https://www.england.nhs.uk/statistics/statistical-work-areas/proms/. Accessed May 5, 2019.

18. Bull KS, Spoudeas HA, Yadegarfar G, et al. Reduction of health status 7 years after addition of chemotherapy to craniospinal irradiation for medulloblastoma: A follow-up study in PNET 3 trial survivors - On behalf of the CCLG (formerly UKCCSG). *J Clin Oncol* 2007;25(27):4239-4245.

19. Limond JA, Bull KS, Calaminus G, et al. Quality of survival assessment in European childhood brain tumour trials, for children aged 5 years and over. *Eur J Paediatr Neurol* 2015;19(2):202-210.

20. Santana MJ, Feeny D. Framework to assess the effects of using patient-reported outcome measures in chronic care management. *Qual Life Res* 2014;23(5):1505-1513.

21. Ireland P, Horridge KA. The Health, Functioning and Wellbeing Summary Traffic Light Communication Tool: a survey of families' views. *Dev Med Child Neurol* 2017;59(6):661-664.

22. NHS England. *Putting Patients First: The NHS business plan for 2013/2014-2015/2016*. http://www.england.nhs.uk/pp-1314-1516/. Accessed May 5, 2019.

23. Velikova G, Booth L, Smith AB, et al. Measuring quality of life in routine oncology practice improves communication and patient well-being: A randomized controlled trial. *J Clin Oncol* 2004;22(4):714-724.

24. Tavernor L, Barron E, Rodgers J, et al. Finding out what matters: validity of quality of life measurement in young people with ASD. *Child Care Health Dev* 2013;39(4):592-601.

25. Haverman L, van Rossum MA, van Veenendaal M, et al. Effectiveness of a web-based application to monitor health-related quality of life. *Pediatr* 2013;131(2):e533-543.

26. Haverman L, Engelen V, van Rossum MA, et al. Monitoring health-related quality of life in paediatric practice: development of an innovative web-based application. *BMC Pediatr* 2011;11:3.

27. Snyder CF. Using Patient-Reported Outcomes in Clinical Practice: A Promising Approach? *J Clin Oncol* 2014;32(11):1099-+.

28. Mokkink LB, Terwee CB, Patrick DL, et al. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *J Clin Epidem* 2010;63(7):737-745.

29. Liberati A, Altman DG, Tetzlaff J, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration. *Brit Med J* 2009;339.

30. Dissemination CfRa. *Systematic reviews: CRD's guidance for undertaking reviews in health care*. 3rd ed. York, UK York Publishing Services Ltd.; 2009.

31. Mokkink LB, de Vet HCW, Prinsen CAC, et al. COSMIN Risk of Bias checklist for systematic reviews of Patient-Reported Outcome Measures. *Qual Life Res* 2018;27(5):1171-1179.

32. Janssens A, Rogers M, Gumm R, et al. Measurement properties of multidimensional patient-reported outcome measures in neurodisability: a systematic review of evaluation studies. *Dev Med Child Neurol* 2016;58(5):437-451.

33. Fitzpatrick R, Davey C, Buxton MJ, et al. Evaluating patient-based outcome measures for use in clinical trials. *Health Technol Assess* 1998;2(14):i-iv,1-74.

34. Uijen AA, Heinst CW, Schellevis FG, et al. Measurement Properties of Questionnaires Measuring Continuity of Care: A Systematic Review. *PloS one* 2012;7(7).

35. Terwee CB, Mokkink LB, Knol DL, et al. Rating the methodological quality in systematic reviews of studies on measurement properties: a scoring system for the COSMIN checklist. *Qual Life Res* 2012;21(4):651-657.

36. Varni JW, Seid M, Knight TS, et al. The PedsQL (TM) 4.0 Generic Core Scales: Sensitivity, responsiveness, and impact on clinical decision-making. *J Beh Med* 2002;25(2):175-193.

37. Varni JW, Burwinkle TM, Seid M. The PedsQL as a pediatric patient-reported outcome: Reliability and validity of the PedsQL Measurement Model in 25,000 children. *Expert Rev Pharmacoecon Outcomes Res* 2005;5(6):705-719.

38. Janssens A, Rogers M, Coon JT, et al. A Systematic Review of Generic Multidimensional Patient-Reported Outcome Measures for Children, Part II: Evaluation of Psychometric Performance of English-Language Versions in a General Population. *Val Health* 2015;18(2):334-345.

39. Klassen AF, Strohm SJ, Maurice-Stam H, et al. Quality of life questionnaires for children with cancer and childhood cancer survivors: A review of the development of available measures. *Support Care Cancer* 2010;18(9):1207-1217.

40. Wild D, Grove A, Martin M, et al. Principles of good practice for the translation and cultural adaptation process for patient-reported outcomes (PRO) measures: Report of the ISPOR Task Force for Translation and Cultural Adaptation. *Val Health* 2005;8(2):94-104.

41. Barr RD, Simpson T, Whitton A, et al. Health-related quality of life in survivors of tumours of the central nervous system in childhood - a preference-based approach to measurement in a cross-sectional study. *Euro J Cancer* 1999;35(2):248-55.

42. Glaser AW, Davies K, Walker D, et al. Influence of proxy respondents and mode of administration on health status assessment following central nervous system tumours in childhood. *Qual Life Res* 1997;6(1):43-53.

43. Glaser AW, Furlong W, Walker DA, et al. Applicability of the Health Utilities Index to a population of childhood survivors of central nervous system tumours in the U.K. *Euro J Cancer* 1999;35(2):256-61.

44. Bhat SR, Goodwin TL, Burwinkle TM, et al. Profile of daily life in children with brain tumors: An assessment of health-related quality of life. *J Clin Oncol* 2005;23(24):5493-500.

45. Eiser C, Vance YH, Horne B, et al. The value of the PedsQLTM in assessing quality of life in survivors of childhood cancer. *Child Care Health Dev* 2003;29(2):95-102.

46. Palmer SN, Meeske KA, Katz ER, et al. The PedsQL (TM) brain tumor module: Initial reliability and validity. *Pediatr Blood Cancer* 2007;49(3):287-293.

47. Bedell GM. Developing a follow-up survey focused on participation of children and youth with acquired brain injuries after discharge from inpatient rehabilitation. *NeuroRehab* 2004;19(3):191-205.

48. Bedell G. Further validation of the Child and Adolescent Scale of Participation (CASP). *Dev Neurorehab* 2009;12(5):342-351.

49. Torrance GW, Feeny DH, Furlong WJ, et al. Multiattribute utility function for a comprehensive health status classification system. Health Utilities Index Mark 2. *Med Care* 1996;34(7):702-22.

50. Feeny D, Furlong W, Torrance GW, et al. Multiattribute and single-attribute utility functions for the health utilities index mark 3 system. *Med Care* 2002;40(2):113-28.

51. Varni JW, Seid M, Kurtin PS. PedsQL (TM) 4.0: Reliability and validity of the pediatric quality of life Inventory (TM) Version 4.0 generic core scales in healthy and patient populations. *Med Care* 2001;39(8):800-12.

Figure Legend

Figure 1. PRISMA flowchart for the identification and selection of studies evaluating psychometric properties of PROMs in children treated for brain tumors or acquired brain injury.

PRISMA=Preferred Reporting Items for Systematic Reviews and Meta-Analyses; PROMs=Patient-reported outcome measures

NOP-D-19-00078R1

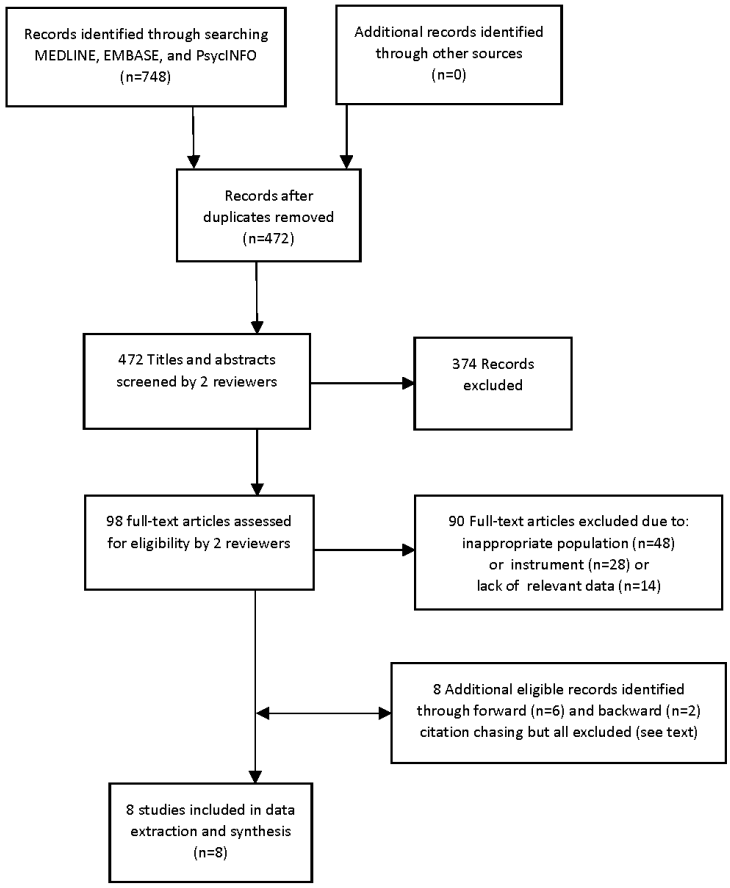

Figure 1

Table 1 Appraisal of measurement properties and indicative criteria (COSMIN checklist)

Psychometric property	Indicative criteria
Content validity	• Clear conceptual framework consistent with stated purpose of measurement • Qualitative research with potential respondents
Internal structure	• Structural validity Factor analysis & post-hoc tests of uni-dimensionality by Rasch analysis • Internal consistency: Cronbach's alpha coefficient >0.7 and <0.9 • Differential item and scale functioning between different sexes, ages, and diagnoses
Reliability/ Reproducibility	• Test-retest reliability: ICC >0.7 adequate, >0.9 excellent • Proxy-reliability: Child and parent-reported reliability ICC >0.7
Measurement error/ Precision	• Assessment of measurement error; floor or ceiling effects <15%; evidence provided by Rasch analysis and/or interval level scaling
Hypothesis testing/ Construct validity	• Hypothesis testing, with a priori hypotheses about direction and magnitude of expected effect sizes
Criterion validity	• Comparison of a shortened PROM to the original long version
(Cross-cultural validity)	• (Not assessed in this systematic review of English language PROMs)
Responsiveness	• Longitudinal data about change in scores with reference to hypotheses, measurement error, and minimal important difference

Table 2. Indices for appraising psychometric properties of patient-reported outcome measures (PROMs) (COSMIN checklist)

Rating	Definition	Description
?	Not clearly determined	studies were rated poor methodological quality; results not considered robust
-	Evidence not in favor	studies were rated good or excellent methodological quality; results did not meet standard criteria for this property
+/-	Conflicting evidence	studies were rated fair, good, or excellent methodological quality; results did not consistently meet standard criteria for this property, e.g. not for all domain scales
+	Some evidence in favor	studies were rated fair or good methodological quality; standard criteria were met for the property
++	Some good evidence in favor	studies were rated good or excellent methodological quality; standard criteria were met or exceeded
+++	Good evidence in favor	studies were rated good or excellent methodological quality; standard criteria were exceeded; results have been replicated

Table 3. Studies identified in the systematic review as reporting psychometric properties of PROMs in children with brain tumors or acquired brain injury up to 18 years old

Acronym of PROM	Author (year)	Aim/Purpose	Study population	N	Age range years	Mean age years (SD)	Country
HUI2/HUI3	Barr et al. (1999)[41]	To assess inter-rater agreement/reliability and construct validity	Brain tumors	44 families	1.7-17.9	9.5	Canada
HUI2/HUI3	Glaser et al. (1997)[42]	To assess test-retest reliability when HUI completed at home and within 2 weeks, in clinic, and compare agreement between patients and parents	CNS tumors	33 families	5-16	10.7 (3.3)	England
HUI2/HUI3	Glaser et al. (1999)[43]	To assess the acceptability, inter-observer reliability and interpretability of HUI2 & HUI3 in UK survivors of childhood cancer	CNS tumors	30 families	6-16	10.5	UK
PedsQL™ (Generic Core Scales)	Bhat et al. (2005)[44]	To assess reliability and validity	Brain tumors	108 families, 17 parents only, 9 children only	NR	11.8 (5.4)	USA
PedsQL™ (Generic Core Scales)	Eiser et al. (2003)[45]	To assess reliability and validity	CNS tumors Other cancers (not included in this review)	23 families 45 families	NR NR	13.7 (3.1) 13.5 (3.2)	England
PedsQL™ (Brain Tumor Module)	Palmer et al. (2007)[46]	To assess validity and internal consistency reliability	Brain tumors	99 families	2-18	9.8	USA

Acronym of PROM	Author (year)	Aim/Purpose	Study population	N	Age range years	Mean age years (SD)	Country
CFFS	Bedell (2004)[47]	To assess preliminary findings of reliability, internal consistency, and criterion validity.	ABI	60 parents	3-27	13.2 (5.2)	USA
CASP (section of the CFFS)	Bedell (2009)[48]	To validate the CASP for young people/children with ABI	ABI, developmental disability, no identified disability, and learning/attention/sensory disability	313 parents ABI=176 (56%)	3-22	12.8 (4.6)	USA, Canada, Australia, Israel

PROM, patient-reported outcome measure; N, sample size; SD, standard deviation; HUI2/HUI3, Health Utilities Index 2/3; PedsQL, Pediatric Quality of Life Inventory; CFFS, The Childhood and Family Follow-Up Survey; CASP, Child and Adolescent Scale of Participation; NR, not reported; ABI, acquired brain injury; CNS, central nervous system

Table 4. Characteristics of the PROMs described in studies of children with brain tumors or acquired brain injury up to 18 years old identified by the systematic review

Acronym of PROM	Original publication (year)	Description	No. of items (type)	Scoring	Domains/scales	Recall period	Time to complete (mins)	Responder	Age range (years)
HUI2	Torrance et al. (1996)[49]	Generic preference-based system for measuring health status and HRQoL	15 (multiple choice)	-0.03 (most disabled) – 1.00 (perfect health)	sensation, mobility, emotion (distress, anxiety), cognition (learning), self-care, pain (frequency and type of control), fertility*	1-, 2-, 4-weeks; usual health status	5 – 10	Proxy Self	≥5 ≥12
HUI3	Feeny et al. (2002)[50]	Generic preference-based system for measuring health status and HRQoL	15 (multiple choice)	-0.36 (most disabled) – 1.00 (perfect health)	vision, hearing, speech, ambulation, dexterity, emotion (happiness vs depression), cognition (ability to solve day-to-day problems), pain (severity)	1-, 2-, 4-weeks; usual health status	5 – 10	Proxy Self	≥5 ≥12
PedsQL™ 4.0 (Generic Core Scales)	Varni et al. (2001)[51]	Generic measure of HRQoL	23 (Likert scale)	0 – 100, higher scores, better functioning	Physical health, Psychosocial health (comprising emotional, social, and school scales)	1 month	5	Child Parent	5-18 2-18
PedsQL™ 4.0	Palmer et al. (2007)[46]	Brain tumor specific measure of HRQoL	24 (Likert scale)	0 – 100, higher scores, better functioning	Cognitive problems, pain and hurt, movement and balance,	7 days	5	Child parent	5-18 2-18

Acronym of PROM	Original publication (year)	Description	No. of items (type)	Scoring	Domains/scales	Recall period	Time to complete (mins)	Responder	Age range (years)
(Brain Tumor Module)					procedural anxiety, nausea, worry				
CFFS (includes CASP, CAFI, and CASE)	Bedell (2004)[47]	To monitor needs and outcomes of children and adolescents with ABI and their families after discharge from inpatient rehabilitation	5 sections: I. 6 (multiple choice) II. 20 in the CASP (4-point scale) and 3 open-ended III. 15 in the CAFI, 18 in the CASE (both 3-point scales) and 1 open-ended IV. 6 (multiple choice) V. 2 (open-ended) Total 71	I. categorical II. CASP 0 – 100, higher scores, greater age-expected participation III. CAFI 0 – 100, higher scores, greater extent of problem; CASE 0 – 100, higher scores greater extent of environment problem	5 sections: I. Physical and emotional health and well-being, primary way of moving around and communicating, and medical problems or hospitalizations II. CASP including equipment, modifications or strategies to promote participation. III. CAFI and CASE IV. Educational placement, rehabilitation and health services, satisfaction with services; family's quality of life, services, and needs V. Suggestions to improve services	within the last year or since leaving the program	30	Parent	5-18

Acronym of PROM	Original publication (year)	Description	No. of items (type)	Scoring	Domains/scales	Recall period	Time to complete (mins)	Responder	Age range (years)
					and additional information not already addressed				

PROM, patient-reported outcome measure; HUI2/HUI3, Health Utilities Index 2/3; PedsQL, Pediatric Quality of Life Inventory; CFFS, The Childhood and Family Follow-Up Survey; CASP, Child and Adolescent Scale of Participation; HRQoL, Health-related quality of life; ABI, acquired brain injury; *The fertility question is not integral to the questionnaire but can be added if relevant to the population (Furlong et al. 2001); CAFI, Child and Adolescent Factors Inventory; CASE, Child and Adolescent Scale of Environment

Table: 5 Summary appraisal of measurement properties of each PROM identified in the systematic review

Instrument version	Content validity	Structural validity	Internal consistency	Test-retest reliability/ Reproducibility	Proxy reliability/ Reproducibility	Measurement error/ Precision	Hypothesis testing/ Construct validity	Responsiveness
HUI				+/-	+/-	+	+	
PedsQL			+/-	+/-	+/-		+	
PedsQL Brain Tumor Module			+/-				+	
CFFS (including CASP section)	+	++	++			+	+	

PROM, patient-reported outcome measure; HUI, Health Utilities Index; PedsQL, Pediatric Quality of Life Inventory; CASP, Child and Adolescent Scale of Participation; CFFS, Childhood and Family Follow-Up Survey; +, some evidence in favor; ++, some good evidence in favor; +/-, conflicting evidence

Table: 6 Summary appraisal of measurement properties of each PROM identified in the systematic review

Instrument version	Content validity	Structural validity	Internal consistency	Test-retest reliability/ Reproducibility	Proxy reliability/ Reproducibility	Measurement error/ Precision	Hypothesis testing/ Construct validity	Responsiveness
HUI				+/-	+/-	+	+	
PedsQL			+/-	+/-	+/-		+	
PedsQL Brain Tumor Module			+/-				+	
CFFS (including CASP section)	+	++	++			+	+	

PROM, patient-reported outcome measure; HUI, Health Utilities Index; PedsQL, Pediatric Quality of Life Inventory; CASP, Child and Adolescent Scale of Participation; CFFS, Childhood and Family Follow-Up Survey; +, some evidence in favor; ++, some good evidence in favor; +/-, conflicting evidence

Supplementary Table S1. Psychometric properties of PROMs validated in children treated for brain tumors or acquired brain injury up to 18 years old

Acronym of PROM	Author (year)	Content Validity	Structural Validity	Internal Consistency	Reliability/ Reproducibility	Measurement Error/ Precision	Hypothesis Testing / Construct Validity	Responsiveness
HUI2/HUI3	Barr et al. (1999)[41]				No interrater reliability reported for children vs parents ICC between raters for global scores ranged from 0.54-0.95 Attribute/tem agreement varied poor/fair/good		Impact of (i) radiotherapy, (ii) disease status on HUI2/HUI3 (mean) scores (i) HUI2: None (0.9), posterior fossa (0.92), supratentorium (0.82), craniospinal (0.71) HUI3: None (0.82), posterior fossa (0.79), supratentorium (0.71), craniospinal (0.42) (ii) HUI2: None (0.89), residual tumor (0.81), recurrent tumor (0.56) HUI3: None (0.78), residual tumor (0.56),	

Acronym of PROM	Author (year)	Content Validity	Structural Validity	Internal Consistency	Reliability/ Reproducibility	Measurement Error/ Precision	Hypothesis Testing / Construct Validity	Responsiveness
							recurrent tumor (0.32)	
HUI2/HUI3	Glaser et al. (1997)[42]				Test-retest reliability of HUI single attributes Kappa values 0.02–1.0; Percentages 12-100 Patient-parent interrater reliability of single attributes Kappa values 0.28–1.0; Percentages 53-100			
HUI2/HUI3	Glaser et al. (1999)[43]				Test-retest reliability of HUI emotion utility score: Parent Kappa 0.84, 90%, ICC 0.87; Children Kappa 0.54, 68%, ICC 0.62. HUI2 patient-parent interrater		Similarities in health status between patients in Canada (Barr et al., 1999) and the U.K.	

Acronym of PROM	Author (year)	Content Validity	Structural Validity	Internal Consistency	Reliability/ Reproducibility	Measurement Error/ Precision	Hypothesis Testing / Construct Validity	Responsiveness
					reliability of single attributes Kappa values 0.37–0.72, percentages 67-88; global HRQoL ICC 0.57			
PedsQL™ 4.0 (Generic Core Scales)	Bhat et al. (2005)[44]			Proxy and self-reports total HRQoL and psychosocial summary scores Cronbach's α coefficients >0.70; Self-report domain scores ranged 0.49-0.68).	Parent-child interrater reliability Pearson's correlation coefficients for subscales and total HRQoL scores 0.34-0.73		Significant difference between patient and parent responses and healthy controls for all subscales (P<0.001). Age, time from diagnosis, and sex no significant association with total HRQoL or psychosocial health. Lower social functioning (P <.05) associated with longer time from diagnosis (parent-report). Females reported higher school	

Acronym of PROM	Author (year)	Content Validity	Structural Validity	Internal Consistency	Reliability/ Reproducibility	Measurement Error/ Precision	Hypothesis Testing / Construct Validity	Responsiveness
							functioning than males (P<.05). No impact of tumor location and for child-report no impact of pathology. Parents showed significant differences by pathology for the physical (P<0.01) and emotional (P<0.01) scales. Parents of children with low-grade gliomas reported higher total HRQOL (P<0.05) and better emotional and physical functioning (P<0.05). Presence of a shunt lower total HRQoL (P<0.01 self-reports; P<0.05 parent-reports) and decreased social	

Acronym of PROM	Author (year)	Content Validity	Structural Validity	Internal Consistency	Reliability/ Reproducibility	Measurement Error/ Precision	Hypothesis Testing / Construct Validity	Responsiveness
							functioning (P<0.05); Self-report lower physical (P<0.05) and psychosocial functioning (P<0.01). Treatment type revealed significant differences in total HRQOL for child (P<0.03) and parent (P<0.001) reports. Magnitude of expected effect sizes not reported.	
PedsQL™ 4.0 (Generic Core Scales)	Eiser et al. (2003)[45]			Cronbach's α >0.7 Item-total correlations broadly acceptable	Interrater reliability between mother and child: physical health 0.29, psychosocial health 0.51, total HRQoL scores 0.50	Ceiling effects for proxy- (4.35%) and self-report (4.34%) for physical functioning. No floor effects.	As predicted, ALL better physical and psychosocial scores than CNS: Child means 86.16 vs 70.65, 79.07 vs 69.19; mother means 84.74 vs 60.05, 73.71 vs 57.74; mothers' lower	

Acronym of PROM	Author (year)	Content Validity	Structural Validity	Internal Consistency	Reliability/ Reproducibility	Measurement Error/ Precision	Hypothesis Testing / Construct Validity	Responsiveness
							scores than survivors.	
PedsQL Brain Tumour Module	Palmer et al. (2007)[46]			Parent proxy-report for all scales and age groups Cronbach alpha range 0.66-0.95; child-report 0.30-0.93	Interrater reliability between parent and child reports are 0.39 to 0.53		Inter-correlations between child-report PedsQL Brain Tumor Module Scales and the Generic Core Scales range 0.03-0.73; Multidimensional Fatigue Scales range 0.05-0.77. Parent-report PedsQL Brain Tumor Module Scales and the Generic Core Scales range 0.12-0.68; Multidimensional Fatigue Scales range 0.08-0.77.Correlations predicted between Brain Tumor Module Scales and PedsQL Core HRQoL total score (range 0.11-0.60, child-report;	

Acronym of PROM	Author (year)	Content Validity	Structural Validity	Internal Consistency	Reliability/ Reproducibility	Measurement Error/ Precision	Hypothesis Testing / Construct Validity	Responsiveness
							0.24-0.60 parent-report). High correlations predicted between Brain Tumor Module Cognitive Problems Scale and PedsQL Core School Functioning Scale (0.56, child-report; 0.65 parent-report). Correlations predicted between Brain Tumor Module Cognitive Problems Scale and Multidimensional Fatigue Cognitive Fatigue Scale (0.77, both child- and parent-report)	
CFFS including the CASP, CAFI, and CASE	Bedell (2004)[47]	Developed from the ICF, rehabilitation outcomes, feedback from stakeholders	Factor, Principal Components, and Rasch analysis for the CASP	Cronbach's alpha = 0.98 and 0.95 for the CASP; item – total score	Test retest ICC: CASP = 0.94 CAFI = 0.67 CASE = 0.75	CASP floor and ceiling effects: 1.5% scored at floor 12% at the ceiling	CASP summary scores positively correlated with PEDI scores (greater participation,	

Acronym of PROM	Author (year)	Content Validity	Structural Validity	Internal Consistency	Reliability/ Reproducibility	Measurement Error/ Precision	Hypothesis Testing / Construct Validity	Responsiveness
		including rehabilitation service providers and administrators, family caregivers of children with ABI	section (other sections not reported)	correlations 0.67 to 0.81		(other sections not reported)	higher functional activity ability) (r = 0.72, self-care; r = 0.65, social function; r = 0.51, mobility, P<0.01). CASP summary scores negatively correlated with CASE (less participation, more impact of environment (r = −0.57, P<0.01) and CAFI summary scores (more child-related problems (r = −0.58, P<0.01). (other sections not reported)	
CASP (section of the CFFS)	Bedell (2009)[48]		Factor, Principal Components, and Rasch analyses.	Cronbach's alpha = 0.96		Ceiling effects 10%; no floor effects	Significant differences in scores between disability groups (p<.001). CASP scores significantly negatively correlated with	

Acronym of PROM	Author (year)	Content Validity	Structural Validity	Internal Consistency	Reliability/ Reproducibility	Measurement Error/ Precision	Hypothesis Testing / Construct Validity	Responsiveness
							CAFI and CASE scores (r=-0.66 and -0.43, p<0.001, respectively).	

PROM, patient-reported outcome measures; HUI, Health Utilities Index; PedsQL, Pediatric Quality of Life Inventory; CASP, Child and Adolescent Scale of Participation; CFFS, Childhood and Family Follow-Up Survey; ICF, International Classification of Functioning, Disability and Health; PEDI, Pediatric Evaluation of Disability Inventory; CASE, Child and Adolescent Scale of Environment; CAFI, Child and Adolescent Factors Inventory; ALL, acute lymphoblastic leukemia; ABI, acquired brain injury; HRQoL, health-related quality of life

Supplementary Table S2: COSMIN risk of bias checklist of methodological quality of studies evaluating measurement properties of candidate PROMs of included studies

Instrument version	Author	Content validity	Internal structure		Reliability/ Reproducibility		Measurement error/ Precision	Hypothesis testing/ Construct validity	Responsiveness
			Structural validity	Internal consistency	Test-retest reliability	Proxy reliability			
HUI2/HUI 3	Barr et al. (1999)[41]					very good			
HUI2/HUI 3	Glaser et al. (1997)[42]				adequate	adequate			
HUI2/HUI 3	Glaser et al. (1999)[43]					adequate			
PedsQL™ (Generic Core Scales)	Bhat et al. (2005)[44]			very good		adequate			
PedsQL™ (Generic Core Scales)	Eiser et al. (2003)[45]			very good		adequate	adequate		

Instrument version	Author	Content validity	Internal structure		Reliability/ Reproducibility		Measurement error/ Precision	Hypothesis testing/ Construct validity	Responsiveness
			Structural validity	Internal consistency	Test-retest reliability	Proxy reliability			
PedsQL™ (Brain Tumor Module)	Palmer et al. (2007)[46]			very good	very good				
CFFS	Bedell (2004)[47]	very good	very good	very good	very good		very good		
CASP (section of the CFFS)	Bedell (2009)[48]		very good						

COSMIN= COnsensus-based Standards for the selection of health status Measurement INstruments;

PROM=Patient reported outcome measures.