Quality of Life Research (2019) 28:1–12
https://doi.org/10.1007/s11136-018-1955-8

REVIEW

A systematic review of utility values in children with cerebral palsy

Utsana Tonmukayakul[1,3] · Long Khanh-Dao Le[1] · Shalika Bohingamu Mudiyanselage[1] · Lidia Engel[1] · Jessica Bucholc[1] · Brendan Mulhern[2] · Rob Carter[1] · Cathrine Mihalopoulos[1]

Accepted: 28 July 2018 / Published online: 2 August 2018

Abstract

Purpose Project aims include the following: (i) to identify reported utility values associated with CP in children aged ≤ 18 years; (ii) to explore utility value elicitation techniques in published studies; and (iii) to examine performance of the measures and/or elicitation approaches.

Methods Peer-reviewed studies published prior to March 2017 were identified from six electronic databases. Construct validity, convergent validity, responsiveness, and reliability of instruments were assessed.

Results Five studies met the inclusion criteria. Utility values of hypothetical general CP states obtained from a general population of parents ranged from 0.55 to 0.88 using time trade off (TTO) and 0.60–0.87 using standard gamble (SG) techniques. Utility values reported by clinicians of three hypothetical spastic quadriplegic CP states, using the Health Utility Index Mark 2 (HUI-2), ranged from 0.40 to 0.13. Other sources of utilities identified were based on both proxy and child ratings using Health Utility Index Mark 3 (HUI-3) (values ranged from − 0.013 to 0.84 depending on the valuation source) and the Assessment of Quality of Life 4 Dimension instrument, with values ranging from 0.01 to 0.58. Construct validity of the HUI-3 varied from moderate to strong, whereas mixed results were found for convergent validity. Responsiveness and reliability were not reported.

Conclusion There was substantial variation in reported utilities. Indirect techniques (i.e. via multi-attribute utility instruments) were more frequently used than direct techniques (e.g. TTO, SG). Further research is required to improve the robustness of utility valuation of health-related quality of life in children with CP for use in economic evaluation.

Keywords Utility value · Children · Adolescent · Cerebral palsy · Quality of life · Quality-adjusted life years · Utility weight

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s11136-018-1955-8) contains supplementary material, which is available to authorized users.

✉ Utsana Tonmukayakul
utsana.tonmukayakul@deakin.edu.au

Long Khanh-Dao Le
long.le@deakin.edu.au

Shalika Bohingamu Mudiyanselage
shalika.b@deakin.edu.au

Lidia Engel
lidia.engel@deakin.edu.au

Jessica Bucholc
j.bucholc@deakin.edu.au

Brendan Mulhern
Brendan.Mulhern@chere.uts.edu.au

Rob Carter
rob.carter@deakin.edu.au

Cathrine Mihalopoulos
cathy.mihalopoulos@deakin.edu.au

[1] Deakin Health Economics, Centre for Population Health Research, Deakin University, Locked Bag 20000, Geelong, VIC 3220, Australia

[2] Centre for Health Economics Research and Evaluation, University of Technology Sydney, 1-59 Quay Street, Hay Market, Sydney, NSW 2000, Australia

[3] Deakin Health Economics, Centre for Population Health Research, Deakin University, 3rd Floor BC Building 221 Burwood Highway, Burwood, Melbourne, VIC 3125, Australia

Introduction

Cerebral palsy (CP) is a term describing a group of permanent disorders affecting the development of movement and posture causing activity limitation and often involves motor disorders and disturbances of sensation, cognition, communication and behaviour [1]. CP is caused by damage to the brain in the early stages of life. Co-morbidities of CP include epilepsy, intellectual disability, hearing and vision impairments [2], bladder problems, poor saliva control [3] and sleep disorders, together with secondary musculoskeletal problems, such as progressive joint contractures of upper and lower extremities [2].

According to the Australian Cerebral Palsy Register, the point-prevalence for CP is 2.1 per 1000 live births (95% Confidence Interval 2.0–2.2) [4]. There are approximately 34,000 people with CP in Australia and 700 new cases are diagnosed each year [5]. CP is associated with substantial economic impact and health burden to families and carers [6]. The total health expenditure for CP in 2007 price was estimated at AUD $40.5 million or AUD$1,197 per person per annum [7]. Hence, there is a need to maximise benefits from healthcare spending. Economic evaluation is an approach that determines the value-for-money of different types of interventions.

Cost utility analysis (CUA) is a type of economic evaluation that compares both costs and outcomes of different alternatives. The most common outcome of CUA is the quality-adjusted life year (QALY). The QALY is a single, numeric index where time spent in a health state is weighted by the "utility value" of that health state. Utility values are anchored on a scale between 0 (being dead) and 1 (being in full health). Negative utility values are possible which are defined as health states valued as worse than death.

Utility values can be quantified using either direct or indirect methods [8]. Direct methods refer to the application of valuation techniques [such as the standard gamble (SG), time trade-off (TTO), visual analogue scale (VAS), and person trade-off (PTO)] to either individual health or to described vignettes of various health states [9]. SG and TTO techniques provide individuals with a choice between two options: (i) to remain in a reference health state or (ii) to move to full health for a shorter period of time in TTO, or to move to full health and accepting some risk of immediate death in SG. The VAS is a rating scale with defined end-points of 0 and 100, where 0 is the worst imaginable health state and 100 is the best imaginable health state. The VAS does not incorporate trade-offs, so the scores derived from VAS cannot be considered as utility values per se [8]. Besides these cardinal valuations of health states, other direct, ordinal utility elicitation methods have also been used, including discrete-choice experiments, ranking exercises, and best-worst scaling [10]. The indirect valuation method refers to utility values that are derived indirectly using mostly generic multi-attribute utility instruments (MAUIs). MAUIs have two main components: (i) a descriptive component [usually a questionnaire covering various domains of health-related quality of life (HRQoL)] and (ii) a valuation component, comprising a utility scoring algorithm that converts responses from the questionnaire to an index or utility value. The utility scoring algorithm is calibrated by using one of the direct valuation techniques described above [11].

MAUIs identified by Chen and Ratcliffe [12] that have been used in children/adolescents include: (i) the Quality of Well-Being Scale (QWB) [13]; (ii) the Health Utility Index Mark 2 (HUI-2) [14]; (iii) the Health Utility Index Mark 3 (HUI-3) [15]; (iv) the 16-dimensional measure of HRQoL (16D) [16]; (v) the 17-dimensional measure of HRQoL (17D) [17]; (vi) the Assessment of Quality of Life-6 Dimensions (AQoL-6D) Adolescent [18]; (vii) the Child Health Utility 9D (CHU-9D) [19]; (viii) the EuroQol-5 Dimensions Youth version (EQ-5D-Y) [20]; and (ix) the Adolescent Health Utility Measure (AHUM) [21]. Most of these MAUIs were initially developed for adults and subsequently adapted for use in younger populations [12].

Generic MAUIs are designed to be used across different diseases and health conditions. Thus, the sensitivity of these measures may be problematic in some conditions, particularly if they do not include dimensions that are important to people with the disease/condition being considered. Previous studies have documented that generic measures, such as the EQ-5D, often lack sufficient sensitivity regarding condition-specific changes in health status in adults for vision impairment [22], hearing loss [23], schizophrenia [24], epilepsy [25], and Alzheimer's disease [26]. Using an instrument that does not account for important disease-specific dimensions and lacking sensitivity can result in inaccurate utility values and subsequent QALYs that are used in CUA. It is because of this reason that some disease-specific MAUIs have been developed [27]. As mentioned above, CP is associated with high costs of care and economic evaluations can aid resource allocation decisions in this population. While evidence of MAUI application in some conditions is available, a comprehensive review of the application and assessment of MAUIs in the context of children with CP is lacking.

The aims of this systematic review are to (i) identify utility values associated with CP in children aged 18 years and younger, (ii) explore utility value elicitation techniques used in the literature, and (iii) examine the performance in terms of construct validity, convergent validity, responsiveness and reliability of generic MAUIs, particularly their ability to reflect health states in children with CP.

Methods

This systematic review adhered to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines [28] and was registered on the PROSPERO database [29]. Databases searches included MEDLINE (via Ovid, 1946 onwards), Embase (1947 onwards), EconLit (1886 onwards), Cochrane Library (including the Health Technology Assessment Database and NHS Economic Evaluation Database), Web of Science (1864 onwards) and School of Health and Related Research Health Utility Database. Searching was completed in March 2017 and grey literature was excluded. The search strategy (Table 1) was developed from previous systematic reviews of non-CP utility values and from a systematic review of CP [30–34]. The reference lists of included studies were also examined for other eligible studies.

Eligible studies were published in English and reported utility values of children with CP aged 18 years and younger. There was no restriction relating to the type of study (cross-sectional or longitudinal) or utility valuation methods. Both direct elicitation methods and indirect methods through the use of MAUIs were included in this review. Protocol papers or conference abstracts were excluded.

After removing all duplicates, the remaining studies were divided and screened by two groups of two independent reviewers (LL & SB and JB & UT). Titles, abstracts and full text were screened and the final included articles were extracted and entered into a table by all reviewers. Literature selection and data extraction were cross-checked between reviewers within and between groups. All discrepancies were resolved by discussion and final consensus between the four reviewers.

The data extracted included the following:

1. *Demographic information* study population, participant characteristics;
2. *Descriptive information about the study* publication year, country, sample size, respondents, study type (e.g. clinical trial, descriptive);
3. *Information about the utility valuation* (i) direct method (e.g. technique used (e.g. SG, TTO), health state description, reported utility values); and/or (ii) indirect method (e.g. instrument used, utility algorithm, reported utility values).

The conduct of a subsequent quantitative meta-analysis was only considered if the search identified at least two papers that reported utility values for the same population, with the same condition and using the same valuation methods, as previously recommended in the literature [35].

The performance of instruments was extracted from the included papers and assessed in terms of (1) *construct validity* (the ability to distinguish disease severity on the basis of clinical indicators); (2) *convergent validity* (the strength of correlation between an instrument and a second validated instrument); (3) *responsiveness* (the ability of an instrument to measure a change in health status); and (4) *reliability* (the repeatability of the values obtained when measured on an unchanged population and/or the comparability of scores across different assessors e.g. self-versus proxy-reports, inter-rater agreement) [36]. Using common "rule of thumb" strength of association classifications, the construct and convergent validities were classified into strong (correlation coefficient > 0.7), moderate (0.5–0.7) and weak (< 0.5) [37].

Results

Search results

A total of 3487 publications remained after removal of duplicates. Figure 1 shows the number of identified records, included and excluded studies, together with reasons for exclusion. Twenty publications were eligible for full-text review and of those, five were finally included.

Description of included studies

All five included studies employed a cross-sectional design. Table 2 presents details of each study. Only one study derived utility-values for the purpose of a health economic evaluation. This study evaluated the cost-effectiveness of intrathecal baclofen therapy for severe spastic quadriplegic CP [38]. The other four studies were largely descriptive in nature. Two were large population-based surveys determining the utility values of a broad spectrum of diseases including CP [39, 40]. The other two were studies measuring utility values from either self-reported surveys of children with CP [41] or proxy-reported surveys of parents/carers of children with CP [42]. Direct utility valuation methods (TTO and SG) were used in only one study [39], while the other four studies used generic MAUIs including the HUI-2, HUI-3 and AQoL-4D [38, 40–42].

Carroll and Downs [39] conducted a large survey of 4016 parents with at least one child under 18 years of age who may or may not have had a disability, in the United States (US). Three hypothetical health states, mild, moderate and severe CP, based on functional difficulties were included. The narrative scenarios described physical, social, and emotional characteristics of these three health states. Details of how each narrative scenario was developed were not provided. Health state descriptions for this survey are outlined

Table 1 MEDLINE Ovid search strategy

Theme/concepts	Search terms and strategy
(1) Utility values (outcomes)	Quality-adjusted life years [MeSH, exp] OR value of life [MeSH, exp] OR wellbeing OR qwb OR quality and wellbeing.mp OR utility OR utilities OR disutility OR disutilities OR "quality of life" OR ((preference-based or generic) adj8 (instrument? or measure?)).mp OR (health status adj3 (measure? or utility$)).mp OR health status indicator OR (health state adj8 preference?).mp OR health adj5 preference?).mp OR preference valu$.mp OR (utilit$ adj10 health state?).mp OR (utilit$ adj10 health status).mp OR (elicit$ adj5 utilit$).mp OR ((valuation? Or value?) adj5 state?).mp OR ((valuation? Or value?) adj8 health state?).mp OR ((preference? Or valu$ or utility$) adj8 state?).mp OR preference scale?.mp OR (preference? Adj3 (public or health or social or societal or society)).mp OR (health adj3 reference?).mp OR (valu$ adj2 (societal or social)).mp OR social values OR tariff?.tw OR value? set?.tw OR valuation set?.tw OR preference weight?.tw OR preference? elicitation?.mp OR value of life/ec [economics]
(2) Direct preference elicitation methods	Trade off?.mp OR standard gamble?.mp OR analogue scale?.mp OR visual analogue scale or VAS scale.tw OR discrete choice experiment or DCE.tw OR person trade off or ptt OR scaling method?.mp OR (equivalence or equivalent).tw OR magnitude estimation.mp OR contingency valuation? or CV?.mp OR best worst scaling? or bws?.mp OR (elicit$ adj8 (preference? or valu$ or view?)).mp OR valuation? stud$.mp OR direct valuation?.mp

Table 1 (continued)

Theme/concepts	Search terms and strategy
(3) Indirect preference elicitation methods	(Multiattribute or multi-attribute).mp OR (EQ 5D$ or EQ5D$).mp OR (EuroQol$ or Euro Qol$).mp OR SF-6D.mp OR (SF 12$ or SF12$ or Short-form 12$ or Short-form12$ or MOS 12 or MOS12).mp OR (SF 36$ or SF36$ or Short-form 36$ or Short-form36$ or MOS 36 or MOS36).mp OR (RAND36$ or RAND 36$).mp OR (VR36$ or VR 36$).mp OR (Health Utilities Index or Health Utility Index).mp OR (HUI-2 or HUI2 or HUI-3 or HUI3).mp OR quality of wellbeing scale?.mp OR QWB$.mp OR assessment of quality of life.mp OR AQoL$.mp OR 15d.tw OR 16d.tw OR 17d.tw OR (CHU 9D$ or CHU9D$ or CHU-9D).mp OR mapping OR valuation model.tw OR valuation model?.tw
(4) Population	Children or child
(5) Condition	Cerebral palsy
(6)	4 AND 5
(7)	1 AND 3 AND 6
(8)	1 AND 4 AND 6

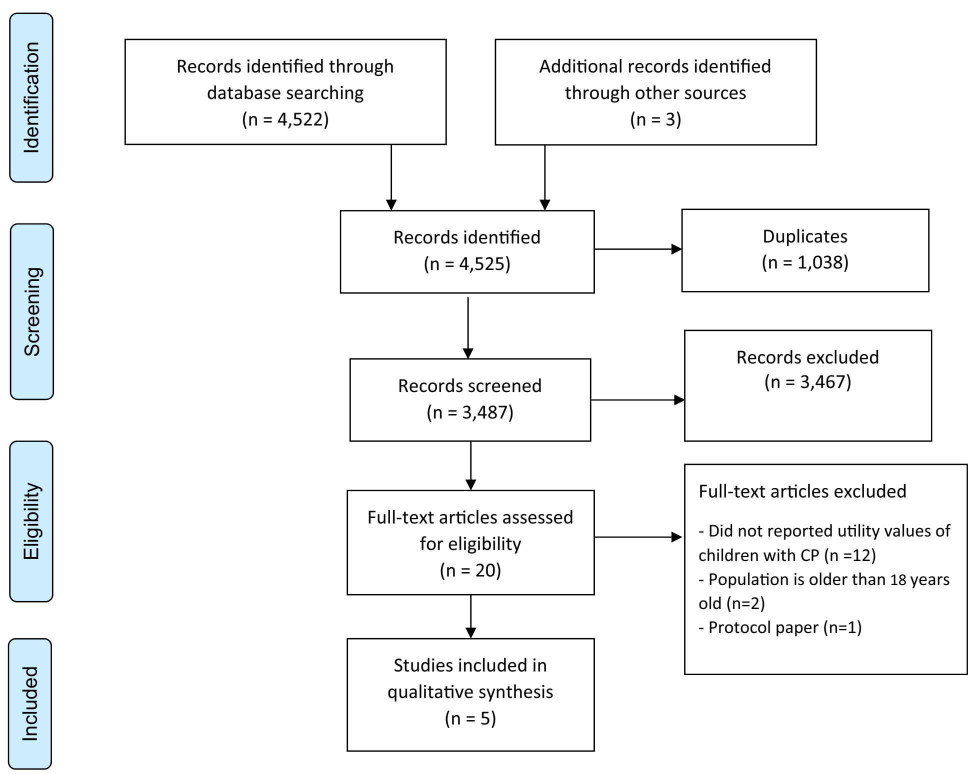

Fig. 1 PRISMA diagram

Table 2 Characteristics of included studies

Reference (country)	CP subtype	Severity	Sample size	Age of children with CP	Respondent	Sex	Utility valuation measurement	Study type
Direct method								
Carroll and Downs [39] (USA)	Not specified	Mild, moderate, severe	4016	Not reported	Parents of at least 1 child	3420 male 511 female	SG and TTO	Cross-sectional
Indirect method								
Lissavoy et al. [38] (USA)	Spastic quadriplegic CP[a]	Mild, moderate, severe	9	Not reported	Physicians and nurse	Not specified	HUI-2	Cross-sectional[b]
Petrou and Kupek [40] (England and Scotland)	Not specified	Not specified	178	Not reported	Children with disability	Not specified	HUI-3	Cross-sectional
Rosenbaum et al. [42] (Canada)	Not specified	GMFCS[c] levels I-V	203	13–20 years old	Parent/caregivers	111 male 92 female	HUI-3	Cross-sectional
Young et al. [41] (Canada)	Not specified	GMFCS levels I-V	129	13–17 years old	Children with CP parent/caregivers	106 male 86 female	HUI-3 and AQoL-4D	Cross-sectional

SG standard gamble, *TTO* time trade off, *HUI2* Health Utility Index Mark 2, *HUI3* Health Utility Index Mark 3, *AQoL-4D* the Assessment of Quality of Life

[a]Spastic quadriplegic: muscle weakness affects all four limbs, abnormal gait and posture

[b]Study type applied to utility values measurement

[c]GMFCS: The 'Gross Motor Function Classification System', which describes motor performance based on (i) functional ability; (ii) need for assistive technology; and (iii) need for wheelchair that ranges from level 1 (independent with few limitation) to level 5 (complete dependence for all activities)

in the supplementary file. The utility values were measured by TTO and SG techniques.

Lissavoy et al. [38], measured three possible treatment outcomes of 5-year intrathecal baclofen therapy in severe spastic quadriplegic CP in the US. The health state valuation task was part of the CUA. These health states were initially developed from a literature review of typical symptoms, potential functional benefits, and complications associated with intrathecal baclofen treatment in children aged 11 years with severe spastic quadriplegic CP. Narration of severity of health states was then added based on the scores of three instruments: Ashworth scores; the Pediatric Evaluation of Disability Inventory Mobility; and the Pediatric Evaluation of Disability Inventory Self-Care. The first draft of each health state description was revised by one clinical author. The final draft (see supplementary files) was presented to eight physicians and one nurse practitioner, who all had extensive experience treating young, spastic and quadriplegic CP patients. The nine respondents were asked to rate the utility of hypothetical health states on the HUI-2.

Petrou and Kupek [40] analysed utility values of children with CP using data from the Disability Survey 2000: Survey of Young People with Disability and Sport in the United Kingdom [43]. The HUI-3 proxy was administered in this national survey and completed by 178 principal caregivers of children with CP based on the International Classification of Disease codes. The severity of CP amongst survey respondents was not specified. Both utility and disutility (an inversion of utility scores or scores away from perfect health) values were reported.

Rosenbaum et al. [42] obtained utility values using the HUI-3 proxy version from 203 parents and/or caregivers of youth with CP aged 13–20 years in Canada. The spectrum of CP severity was stratified according to the Gross Motor Function Classification System (GMFCS) where level 1 is the least severe while level 5 is the most severe stage [44, 45].

Young et al. [41] used HUI-3 and the AQoL-4D to assess the impact of CP severity, age and sex on the quality of life for 129 children with CP aged 13–17 years and for adults with CP aged 24–29 years, in Canada. Self-report was recommended, but assistance to complete the questionnaires and/or proxy report were allowed if necessary. This review reported only the utility values of children with CP who were younger than 18 years of age. It is worth noting that the AQoL-4D is a generic MAUI designed for use in adults [46].

Reported utility values

Table 3 shows the reported utility values. The CI of utility values derived from TTO and SG was overlapping across each CP severity level and statistically significant differences were not reported [39]. Mean proxy-reported utility values measured by TTO and SG for mild CP were 0.88 (SD 0.19) and 0.87 (SD 0.20), respectively. Moderate CP vales were 0.76 (SD 0.26) by TTO and 0.76 (SD 0.23) by SG, while severe CP vales were 0.55 (SD 0.33) by TTO and 0.60 (SD 0.28) by SG.

Clinician-reported utility values for severe spasticity, derived from the HUI-2, were 0.40 (SD 0.11) at 1-year post-intrathecal baclofen therapy pump implantation, 0.16 (SD 0.05) at 6-week post-implantation and 0.13 (SD 0.6) in those who did not receive the treatment [38].

The mean proxy-reported utility value of children with CP was 0.28, while the mean disutility value was −0.73 [40]. The mean proxy-reported utility values, varied by CP severity stratified by the GMFCS were 0.84 (SD 0.20) for *level I*; 0.50 (SD 0.31) for *level II*; 0.39 (SD 0.21) for *level III*; 0.16 (SD 0.26) for *level IV*; and −0.08 (SD 0.23) for *level V* [42].

Mean self-reported utility values derived from the HUI-3 were 0.67 (SD 0.32) for *GMFCS level I*; 0.59 (SD 0.35) for *level II*; 0.43 (SD 0.39) for *level III*, 0.08 (SD 0.25) for *level IV*; and −0.13 (SD 0.19) for *level V*. Mean utility values obtained from the AQoL-4D were slightly lower than the HUI-3 scores. The Australian-based AQoL-4D utility scoring algorithm was used, producing mean utility values of 0.58 (SD 0.31) for *level I*; 0.53 (SD 0.34) for *level II*; 0.31 (SD 0.32) for *level III*, 0.06 (SD 0.12) for *level IV*; and −0.01 (0.07) for *level V* [41].

Performance of generic MAUIs in measuring utility values in children with CP

Only the studies by Rosenbaum et al. [42] and Young et al. [41] contained adequate information to assess *construct and convergent validity* of the HUI-3 and AQoL-4D (Table 3). Construct validity in both studies referred to statistically significant different utility values across the spectrum of the GMFCS. Rosenbaum et al. [42] found that parent/caregiver-reported utility values varied significantly by GMFCS ($p < .05$). This implied that the HUI-3 had strong construct validity.

Young et al. [41], however, reported that the HUI-3 had moderate construct validity, as the GMFCS levels could explain only half (53%) of the variance in the HUI-3 utility scores. Moderate construct validity was also found for the AQoL-4D, with GMFCS levels explaining 45% of the variance in utility scores.

In terms of convergent validity, Rosenbaum and colleagues analysed the correlation between the utility scores measured by the HUI-3 and the HRQoL score measured by the Quality of Life Instrument for people with Developmental Disabilities. Weak correlation between the two measures was found, with an overall correlation coefficient of $r = .28$. The HUI-3 was also weakly correlated with some of the quality of life domains, including (i) 'Being' ($r = .37$); (ii) 'Belonging' ($r = .17$); and (iii) 'Becoming' ($r = .20$). The coefficient of variation (r^2) suggested that the HUI-3 utility values explained between 3% for 'Belonging' and 14% for 'Being' in terms of score variance, indicating weak convergent validity of HUI-3 against the Quality of Life Instrument for People with Developmental Disabilities.

The convergent validity in Young et al. [41] was assessed based on the correlation between the HUI-3 and AQoL-4D. Results showed that they were statistically highly correlated (Pearson $r = .87$, $p < .001$), suggesting strong convergent validity between these two generic MAUIs.

Discussion

Five studies measuring utility values in children with CP were reviewed. Measurement techniques varied, with direct methods used in one study and indirect methods in the remaining four studies. Utility values in children with CP ranged from −0.08 for severe CP using the proxy-reported HUI-3 [42] through to 0.88 for mild CP using the SG in the general population with at least one child (the child may or may not have had a disability) [39].

Calculation of pooled utility values using meta-analysis was not appropriate, largely because none of the included studies met criteria for meta-analysis outlined in the "Methods" section [35]. Different utility valuation techniques and scoring algorithms can result in different utility values for a similar health state. A clear example of this was the Young et al. study, where the HUI-3 and AQoL-4D were compared [41]. While these two questionnaires had good convergent validity, there were differences between them. The HUI-3 includes eight dimensions: vision, hearing, speech, ambulation, dexterity, emotion, cognition and pain with four to six response levels [47], while the AQoL-4D contains four domains: independent living, relationships, mental health, and senses with four response levels [46]. Furthermore, the HUI-3 uses a scoring algorithm derived from the Canadian general population aged 16 years or older [48], while the AQoL-4D algorithm is derived from Australian adults aged 18 years or older [49]. Hence, difference in reported utility values measured by these two MAUIs is likely due to their varying descriptive system and different cultural and demographic characteristics of the people who participated in the weighting exercise upon which the scoring algorithms are

Table 3 Estimated utility values and overview performance of TTO, SG, HUI2, HUI3 and AQoL-4D

Reference	Health state	Direct valuation		Indirect valuation			Measures of clinical severity	Utility/disutility values [mean (SD)]	p-Value if reported	Construct validity (known group severity)	Convergent validity
		TTO	SG	HUI2	HUI3	AQoL-4D				Significant	Significant
Carroll and Downs [39]	Mild cerebral palsy	✓					Unclear	0.88 (0.19)		Not reported	Not reported
	Moderate cerebral palsy							0.76 (0.26)			
	Severe cerebral palsy							0.55 (0.33)			
	Mild cerebral palsy		✓				Unclear	0.87 (0.20)			
	Moderate cerebral palsy							0.76 (0.23)			
	Severe cerebral palsy							0.60 (0.28)			
Lissavoy et al. [38]	Health state A			✓			Unclear	0.13 (0.06)		Not reported	Not reported
	Health state B							0.16 (0.05)			
	Health state C							0.40 (0.11)			
Petrou and Kupek [40]	Own health state				✓		ICD subdivision codes	0.28 (−0.73)		Not reported	Not reported
Rosenbaum et al. [42]	GMFCS level I				✓		GMFCS quality of life instrument for people with developmental disabilities (self-report)	0.84 (0.20)	< .01	Strong	Weak
	GMFCS level II							0.50 (0.31)			
	GMFCS level III							0.39 (0.21)			
	GMFCS level IV							0.16 (0.26)			
	GMFCS level V							−0.08 (0.23)			
Young et al. [41]	GMFCS level I				✓		GMFCS	0.67 (0.32)		Moderate for both HUI3 and AQoL-4D	Strong between HUI3 and AQoL-4D
	GMFCS level II							0.59 (0.35)			
	GMFCS level III							0.43 (0.39)			
	GMFCS level IV							0.08 (0.25)			
	GMFCS level V							−0.13 (0.19)			
	GMFCS level I					✓		0.58 (0.31)			
	GMFCS level II							0.53 (0.34)			
	GMFCS level III							0.31 (0.32)			
	GMFCS level IV							0.06 (0.12)			
	GMFCS level V							−0.01 (0.07)			

TTO time trade off, *SG* standard gamble, *HUI2* Human Utility Index Mark 2, *HUI3* Human Utility Index Mark 3, *AQoL-4D* Assessment of Quality of Life, *ICD* international classification of diseases, *GMFCS* Gross motor function classification system based. The fact that various MAUIs produce different utility values in the same person is also well demonstrated in other published research [50].

The commonly used MAUIs may not capture important HRQoL domains for children with CP. Results from Rosenbaum et al. [42] found that the HUI-3 weakly correlated with the Quality of Life Instrument for People with Development Disability. This weak convergent validity may be the result of two theoretically different constructs—*functioning* (e.g. vision, hearing, speech, ambulation, dexterity) in HUI-3 [47] and *capabilities* (e.g. Being, Belonging, Becoming) in the Quality of Life Instrument for People with Developmental Disability [51]. This is confirmed by a small coefficient of variation between the HUI-3 and the quality of life instrument, particularly in the variance of "capability" domains. This suggests, when capability is the main outcome of interest, the HUI-3 might not be a sensitive measure.

Different sources of utility valuation can lead to different utility values, which relates to an important question, "Who should value health states?" This question underpins two important issues: (i) who should complete the questionnaires and (ii) whose values should be used in the utility scoring algorithm, particularly when MAUIs are administered. Concerning the first issue, a child with CP may provide a more accurate valuation of their own health state if they are able to describe how CP affects their HRQoL and comprehend utility valuation methods. Varni et al. found that children as young as 5 years of age can accurately report their own HRQoL [52], but this might be challenging in children with CP who often have cognitive and/or language impairments. The validity of child self-report utility scoring remains questionable and proxy-report has been widely used in the literature [53, 54]. However, recent evidence found that parent-report scores are systematically lower than self-reports of children with mobility impairment [55]. This finding highlights a form of bias associated with response type effect i.e. parents might be more pessimistic and do not allow for adaption to health states that children may exhibit, while the child with the disability may be more optimistic because of their ability to cope with the chronic health state. The parent-reported outcome appears to be valid for physical impairments and symptoms that can be observed, but might not be the case for subjective aspects e.g. pain [56].

Clinicians can also be useful proxies, as they recognise children's conditions, symptoms and function, but they do not have contact with the child outside of the clinics [57]. Despite advantages and disadvantages of self- and proxy-report utility values, there is no consensus as to who should report such values when the child may not be capable of self-report. None of the included studies in this review investigated the agreement nor difference of utility values rated by different types of respondent e.g. child with CP versus parents, parents versus clinicians.

In regard to the second issue of whose utility values should be used, this remains a debatable topic. One argument is that utility values should be derived directly from individuals who have the condition of interest, as they will be the recipient of the intervention being assessed. However, asking children to place value on their health state (using direct elicitation techniques) can be challenging. Ratcliffe et al. found many adolescents indicated that the direct utility valuation was difficult to do [58]. The opposing argument suggests that utility values should come from the general population because public healthcare services are mainly tax-funded [10]. Obtaining utility values from general population, as employed in the study by Carroll and Downs [39], have been suggested as preferential by a number of reimbursement agencies i.e. the National Institute for Health and Care Excellence [59] and the Canadian Agency for Drugs and Technologies in Health [60]. Where the national guidelines are not available or not appropriate, then it is up to the evaluator to ensure that it is clearly specify whose value have been included in the valuation.

According to information reported in the included studies, the HUI-3 appeared to be a promising generic MAUI in children with CP with good discriminative properties when compared to the GMFCS. However, Livingston et al. argued that the correlation between the health-related quality of life measure and gross motor function was often contradictory, particularly when mobility domains did not consider mobility with the use of an assistive device such as a wheelchair [61]. Whitehurst et al. also reported the poor performance of the HUI-3 in wheelchair-bound individuals [62]. The performance of generic MAUIs in CP in terms of *responsiveness* and *reliability* could not be assessed by the available information, largely because of the cross-sectional design employed in all five studies. Further research is required for robust application of the HUI-3 and other generic MAUIs in the children with CP population.

There are a number of challenges in deriving utility values in children that have been previously discussed in the literature, including (i) the absence of a health state classification system that takes into account the dynamics of child growth and development; (ii) the lack of a health state classification system for use in children younger than 5 years; (iii) children's limited cognitive ability to complete SG, TTO or MAUI questions; and (iv) the accuracy and reliability of proxy report on subjective domains such as emotion [57, 63]. Although the HUI-3 was originally designed for adults [48], our review has shown that it is promising utility value measure in children with CP. However, it would be interesting to explore whether a generic MAUI exclusively developed for children, such as the CHU-9D, is more appropriate in children with CP when compared with the HUI-3.

For certain diseases or conditions, where there is limited information available about the appropriateness of the

MAUIs in measuring utility values, Brazier et al. proposed two alternatives to enable utility valuation for CUA [64]. The first option refers to mapping scores from a condition-specific measure onto a generic MAUI scale using regression techniques. If mapping cannot be conducted because of a weak correlation between the two measures, then the development of a condition-specific MAUI is the alternate solution. The development of a condition-specific MAUI can use either: (i) a sound psychometric methodology where a new instrument is developed "from scratch" or; (ii) items can be systematically selected from the available CP specific, non-MAUI measures to create a CP-specific health state classification system which can then be valued using one of the valuation techniques. To the best of our knowledge, neither mapping nor development of a condition-specific MAUI for CP has been reported in the scientific literature.

There are a few limitations associated with our review. This literature review did not consider grey literature and non-English studies; some utility valuations in children with CP could have been missed and may have led to selection bias. However, the exclusion of the grey literature can be offset by the fact that the published literature had gone through the peer-review process, which can facilitate methodological rigour and overall quality of the studies [63].

Conclusions

The number of studies measuring utility values of children with CP was small. Different utility values for similar health states were reported. These utility value discrepancies were affected by variation in methods and measures, the different conceptual frameworks applied in each MAUI, the application of different utility score algorithms, the variations in the health state descriptions, and the types of respondents. Both direct and indirect valuation techniques were used, but indirect techniques are more commonly used in the CP context. Based on available evidence, the HUI-3 is a promising generic MAUI for use in children with CP, but can be insensitive to aspects of capability. Due to the cross-sectional research approach employed in the past studies, performance of generic MAUIs i.e. reliability and responsiveness in children with CP remains unknown. These findings highlight that the measurement of utility values in children with CP for economic evaluation requires further research to confirm the reliability, validity, and sensitivity of MAUIs in this population as well as exploration of other alternatives in measuring utility values for use in CUA.

Funding This study was funded by the Centre of Research Excellence in Cerebral Palsy (NHMRC APP 1057997) for supporting the conduct of this systematic review. The author UT has received PhD scholarship from the Centre of Research Excellence in Cerebral Palsy.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Human and animal participants This article does not contain any studies with human participants performed by any of the authors.

References

1. Rosenbaum, P., Paneth, N., Leviton, A., Goldstein, M., & Bax, M. (2007). A report: The definition and classification of cerebral Palsy. *Developmental Medicine & Child Neurology, 49*(Suppl s2), 9.
2. Novak, I., Hines, M., Goldsmith, S., & Barclay, R. (2012). Clinical prognostic messages from a systematic review on cerebral palsy. *Pediatrics, 130*(5), e1285–e1312.
3. Reid, S. M., McCutcheon, J., Reddihough, D. S., & Johnson, H. (2012). Prevalence and predictors of drooling in 7- to 14-year-old children with cerebral palsy: A population study. *Developmental Medicine & Child Neurology, 54*(11), 1032–1036.
4. Australian Cerebral Palsy Register Group. (2017). Australian Cerebral Palsy Register Report 2016, Australian Cerebral Palsy Register.
5. Cerebral Palsy Alliance. (2013) Key facts and statistics. Retrieved from https://www.cerebralpalsy.org.au/what-is-cerebral-palsy/facts-about-cerebral-palsy.
6. Bourke-Taylor, H., Howie, L., & Law, M. (2011). Barriers to maternal workforce participation and relationship between paid work and health. *Journal of Intellectual Disability Research, 55*, 511–520.
7. Access Economics Pty. (2008). *The economic impact of cerebral Palsy in Australia in 2007*. Cerebral Palsy Australia.
8. Drummond, M. F., Sculpher, M. J., Claxton, K., Stoddart, G. L., & Torrance, G. W. (2015). *Methods for the economic evaluation of health care programmes* (4th ed.). Oxford: Oxford University Press.
9. Brazier, J., Ratcliffe, J., Salomon, J. A., & Tsuchiya, A. (2017). *Measuring and valuing health benefits for economic evaluation* (2nd ed.). Oxford: Oxford University Press.
10. Brazier, J., Ratcliffe, J., Tsuchiya, A., & Solomon, J. A. (2017). *Measuring and valuing health benefits for economic evaluation*. New York: Oxford University Press.
11. Arnold, D., Girling, A., Stevens, A., & Lilford, R. (2009) Comparison of direct and indirect methods of estimating health state utilities for resource allocation: Review and empirical analysis. *BMJ*. https://doi.org/10.1136/bmj.b2688.
12. Chen, G., & Ratcliffe, J. (2015). A review of the development and application of generic multi-attribute utility instruments for paediatric populations. *Pharmacoeconomics, 33*(10), 1013–1028.
13. Seiber, W. J., Groessl, E. J., David, K. M., & Ganiats, T. G., Kaplan, R. M. (2008). *Quality of Well Being Self-Administered (QWB-SA) Scale user's manual*. Sandiego: Health services Research Center, University of California.
14. Torrance, G., Feeny, D., Furlong, W., Barr, R., Zhang, Y., & Wang, Q. (1996). Multiattribute utility function for a comprehensive health status classification system: Health Utilities Index Mark 2. *Medical Care, 34*, 702–722.
15. Feeny, D., Furlong, W., & Torrance, G. (2002). Multi-attribute and single-attribute utility functions for Health utilities Index Mark 3 system. *Medical Care, 40*(2), 113–128.
16. Apajasalo, M., Sintonen, H., holmberg, C., Sinkkonen, J., Aalberg, V., Pihko, H., et al. (1996). Quality of life in early

adolescence; a sixteen-dimensional health-related measure (16D). *Quality of Life Research, 5*(2), 205–211.
17. Apajassalo, M., Rautonen, J., holmberg, C., Sinkkonen, J., Aalberg, V., Pihko, H., et al. (1996). Quality of life in pre-adolescence: A 17-dimensional ehalth-related measure (17D). *Quality of Life Research, 5*(6), 532–538.
18. Moodie, M., Richardson, J., Rankin, B., & Iezzi, A., Sinha, K. (2010). Prdicting time trade-off health state valuations of adolescents in four Pacific countries using the Assessment of Quality-of-Life (AQoL-6D) instrument. *Value Health, 13*(8), 1014–1027.
19. Stevens, K. (2009). Developing a descriptive system for a new preference-based measure of health-related quality of life for children. *Quality of Life Research, 18*, 1105–1113.
20. Wille, N., Badia, X., Bonsel, G., et al. (2010). Development of the EQ-5D-Y: A child friendly version of the EQ-5D. *Quality of Life Research, 19*(6), 875–886.
21. Beusterien, K., Yeung, J.-E., Pang, F., & Brazier, J. (2012). Development of the multi-attribute Adolescent Health Utility Measure (AHUM). *Health and Quality of Life Outcomes, 10*, 102.
22. Tosh, J., Brazier, J., Evans, P., & Longworth, L. (2012). A review of generic preference-based measures of health-related quality of life in visual disorders. *Value In Health: The Journal of the International Society for Pharmacoeconomics and Outcomes Research, 15*, 118–127.
23. Longworth, L., Yang, Y., Young, T., Mulhern, B., Hernandez Alava, M., Mukuria, C., et al. (2014). Use of generic and condition-specific measures of health-related quality of life in NICE decision-making: A systematic reveiw, statistical modelling and survey. *Health Technology Assessment* (Winchester, England), *18*(9), 1–224
24. Papaioannou, D., Brazier, J., & Parry, G. (2011). How valid and responsive are generic health status measures, such as the EQ-5D and SF-36, in schizophrenia? A systematic review. *Value Health: The Journal of the International Society for Pharmacoeconomics and Outcomes Research, 14*, 907–920.
25. Selai, C. E., Trimble, M. R., Price, M. J., & Remak, E. (2005). Evaluation of health status in epilepsy using the EQ-5D questionnaire: A prospective, observational, 6-month study of adjunctive therapy with anti-epileptic drugs. *Current Medical Research and Opinion, 21*(5), 733–739.
26. Neumann, P. J. (2005). Health utilities in Alzheimer's disease and implications for cost-effectiveness analysis. *Pharmacoeconomics, 23*(6), 537–541.
27. Rowen, D., Brazier, J., Ara, R., & Azzabi, Z. I. (2017). The role of condition-specific preference-based measures in health technology assessment. *Pharmacoeconomics, 35*(Suppl 1), 33–41.
28. Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2009) Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *BMJ*. https://doi.org/10.1136/bmj.b2535.
29. Tonmukayakul, U., Le, L., Bohingamu, S., Engel, L., Bucholc, J., & Mihalopoulos, C. A systematic review of utility values in children with cerebral palsy. PROSPERO 2017 CRD42017058643. http://www.crd.york.ac.uk/PROSPERO/display_record.php?ID=CRD42017058643.
30. Engel, L., Bansback, N., Bryan, S., Doyle-Waters, M. M., & Whitehurst, D. G. (2016). Exclusion Criteria in National Health State valuation studies: A systematic review. *Medical Decision Making, 36*(7), 798–810.
31. Kennedy-Martin, T., Paczkowski, R., & Rayner, S. (2015). Utility values in diabetic kidney disease: A literature review. *Current Medical Research and Opinion, 31*(7), 1271–1282.
32. Liem, Y. S., Bosch, J. L., & Hunink, M. G. (2008). Preference-based quality of life of patients on renal replacement therapy: A systematic review and meta-analysis. *Value Health, 11*(4), 733–741.
33. Wyld, M., Morton, R. L., Hayen, A., Howard, K., & Webster, A. C. (2012). A systematic review and meta-analysis of utility-based quality of life in chronic kidney disease treatments. *PLoS Medicine, 9*(9), e1001307.
34. Novak, I., McIntyre, S., Morgan, C., Campbell, L., Dark, L., Morton, N., Stumbles, E., Wilson, S. A., & Goldsmith, S. (2013). A systematic review of interventions for children with cerebral palsy: State of the evidence. *Developmental Medicine & Child Neurology, 55*(10), 855–910.
35. Peasgood, T., & Brazier, J. (2015). Is meta-analysis for utility values appropriate given the potential impact different elicitation methods have on values? *Pharmacoeconomics, 33*(11), 1101–1105.
36. Brazier, J., & Deverill, M. (1999). A checklist for judging preference-based measures of health related quality of life: Learning from psychometrics. *Health economics, 8*(1), 41–51.
37. Hinkle, D. E., Wiersma, W., & Jurs, S. G. (2003). *Applied statistics for the behavioural sciences* (5th ed.). Boston: Houghton Mifflin.
38. Lissovoy, G. M., L. S.; Green, H.; Werner, M.; Edgar, T (2007). Cost-effectiveness of intrathecal baclofen therapy for the treatment of severe spasticity associated with cerebral palsy. *Journal of Child Neurology, 22*(1), 49–59.
39. Carroll, A. E., & Downs, S. M. (2009). Improving decision analyses: Parent preferences (utility values) for pediatric health outcomes. *The Journal of Pediatrics, 155*(1), 21.e1–25.e5.
40. Petrou, S., & Kupek, E. (2009). Estimating preference-based health utilities index mark 3 utility scores for childhood conditions in England and Scotland. *Medical Decision Making, 29*(3), 291–303.
41. Young, N. L., Rochon, T. G., McCormick, A., Law, M., Wedge, J. H., & Fehlings, D. (2010). The health and quality of life outcomes among youth and young adults with cerebral palsy. *Archives of Physical Medicine and Rehabilitation, 91*(1), 143–148.
42. Rosenbaum, P. L., Livingston, M. H., Palisano, R. J., Galuppi, B. E., & Russell, D. J. (2007). Quality of life and health-related quality of life of adolescents with cerebral palsy. *Developmental Medicine & Child Neurology, 49*(7), 516–521.
43. Sport England. (2000). *Disability Survey: Survey of young people with a disability and sport*. London, UK: Sport England.
44. Palisano, R. J., Rosenbaum, P., Bartlett, D., & Livingston, M. H. (2008). Content validity of the expanded and revised gross motor function classification system. *Developmental Medicine & Child Neurology, 50*(10), 744–750.
45. Palisano, R., Rosenbaum, P., Walter, S., Russel, D., Wood, E., & Galuppi, B. (1997). Development and reliability of a system to classify gross motor function in children with cerebral palsy. *Developmental Medicine & Child Neurology, 39*, 214–223.
46. Hawthorne, G., Richardson, J., & Osborne, R. (1999). The Assessment of Quality of Life (AQoL) instrument: A psychometric measure of health-related quality of life. *Quality of Life Research, 8*(3), 209–224.
47. Horsman, J., Furlong, W., Feeny, D., & Torrance, G. (2003) The Health Utilities Index (HUI(®)): Concepts, measurement properties and applications. *Health and Quality of Life Outcomes, 1*(1), 54.
48. Feeny, D., Furlong, W., Torrance, G. W., Goldsmith, C. H., Zhu, Z., DePauw, S., et al. (2002). Multiattribute and single-attribute utility functions for the health utilities index mark 3 system. *Medical Care, 40*(2), 113–128.
49. Hawthorne, G., & Richardson, J. (1997) *The Assessment of Quality of Life (AQoL) instrument construction, initial validation and utility scaling*. Melbourne: Centre for Health Program Evaluation.
50. Mihalopoulos, C., Chen, G., Iezzi, A., Khan, M. A., & Richardson, J. (2014). Assessing outcomes for cost-utility analysis in depression: Comparison of five multi-attribute utility instruments

with two depression-specific outcome measures. *British Journal of Psychiatry, 205*(5), 390–397.
51. Renwick, R., Fudge Schormans, A., & Zekovic, B. (2003). Quality of life: A new conceptual framework for children with disabilities. *Journal on Developmental Disabilities, 10*, 107–114.
52. Varni, J. W., Limbers, C. A., & Burwinkle, T. M. (2007). How young can children reliably and validly self-report their health-related quality of life? An analysis of 8,591 children across age subgroups with the PedsQL™ 4.0 Generic Core Scales. *Health and Quality of Life Outcomes, 5*(1), 1.
53. Eiser, C. (1995). Choices in measuring quality of life in children with cancer: A comment. *Psycho-Oncology, 4*(2), 121–131.
54. Thorrington, D., & Eames, K. (2015). Measuring health utilities in children and adolescents: A systematic reveiw of the literature. *PLoS ONE, 10*(8), e0135672.
55. Bray, N., Noyes, J., Harris, N., & Edwards, R. T. (2017). Measuring the health-related quality of life of children with impaired mobility: Examining correlation and agreement between children and parent proxies. *BMC Research Notes, 10*(1), 377.
56. Ungar, W. J. (2011). Challenges in health state valuation in paediatric economic evaluation: Are QALYs contraindicated? *Pharmacoeconomics, 29*(8), 641–652.
57. Petrou, S. (2003). Methodological issues raised by preference-based approaches to measuring the health status of children. *Health Economics, 12*(8), 697–702.
58. Ratcliffe, J., Huynh, E., Stevens, K., Brazier, J., Sawyer, M., & Flynn, T. (2015). Nothing about us without us? A compariosn of adolescent and adult health-state values for the child health utility-9D using profile case bast-worst scaling. *Health Economics, 25*(4), 486–496.
59. NICE. (2013). *Guide to the methods of technology appraisal*. London: NICE.
60. Canadian Agency for Drugs and Technologies in Health. (2006). *Guidelines for the economic evaluation of health technologies* (3rd ed.). Ottawa: Canadian Agency for Drugs and Technologies in Health.
61. Livingston, M. H., Rosenbaum, P. L., Russell, D. J., & Palisano, R. J. (2007). Quality of life among adolescents with cerebral palsy: What does the literature tell us? *Developmental Medicine & Child Neurology, 49*(3), 225–231.
62. Whitehurst, D. G. T., Mittmann, N., Noonan, V. K., Dvorak, M. F., & Bryan, S. (2016). Health state descriptions, valuations and individuals' capacity to walk: A comparative evaluation of preference-based instruments in the context of spinal cord injury. *Quality of Life Research, 25*(10), 2481–2496.
63. Griebsch, I., Coast, J., & Brown, J. (2005). Quality-adjusted life-years lack quality in pediatric care: A critical review of published cost-utility studies in child health. *Pediatrics, 115*(5), e600–e614.
64. Brazier, J. E., Rowen, D., Mavranezouli, I., Tsuchiya, A., Young, T., Yang, Y., et al. (2012). Developing and testing methods for deriving preference-based measures of health from condition-specific measures (and other patient-based measures of outcome). *Health Technology Assessment ((Winchester, England)), 16*(32), 1–114. https://doi.org/10.3310/hta16320.