Use of the Child Health Questionnaire in Children with Cerebral Palsy: A Systematic Review and Evaluation of the Psychometric Properties

Nichola McCullough, BSc, PhD and Jackie Parkes, BNurs, PhD

Queens University, Belfast

Objective   To review the psychometric performance of the Child Health Questionnaire (CHQ) in samples of children with cerebral palsy (CP).   **Method**   Four search terms were applied to five databases in a search for papers published between 1993 and January 2007.   **Results**   A total of 13 papers were identified, providing data on 1229 unique children aged 2–18 years old. Three studies reported on the reliability of the CHQ (internal consistency), whilst six studies provided evidence on various dimensions of validity (concurrent; discriminant and item discriminant validity).   **Conclusions**   This review identified a number of psychometric issues that need to be addressed. These include the assessment of additional types of reliability; an examination of the factor structure of the CHQ within the CP population; and the development of normative data using substantial representative samples, particularly in Europe. Until these issues are addressed, researchers utilizing the CHQ in children with CP should be cautious about its interpretation.

Key words   Child Health Questionnaire; cerebral palsy; psychometric properties; health status; pediatric psychology.

Cerebral palsy (CP) is a nonprogressive disorder of movement and posture caused by damage to the developing brain. Approximately 2–2.5 in 1000 children in the developed world have CP [Surveillance of Cerebral Palsy in Europe project (SCPE), 2000] making it the most common cause of motor impairment in childhood (Stanley, Blair, & Alberman, 2000). The condition varies by type, severity, and etiology with milder forms being more common. There are also a number of associated impairments related to intelligence, the special senses, feeding, pain, epilepsy, incontinence, sleeping, and behavioral difficulties. The World Health Organization (2001), through the International Classification of Functioning, Disability and Health (ICF), has recognized the importance of learning more about the adverse impact of CP upon both the child and family and ways of describing or quantifying this.

There are a number of disease-specific and generic measures available with which to assess the impact of CP. One of the most popular generic tools available is the Child Health Questionnaire or CHQ (Landgraf, Abetz,

& Ware, 1996), which evaluates the health status and well-being of 5–18 year olds. Generic instruments are appropriate for use in all children and normative values can be generated, providing a baseline with which to make comparisons of children with disabilities that is seen as an advantage over disease-specific measures. The aim of this article is to review the published studies that have applied the CHQ in children with CP and to evaluate the psychometric performance of the instrument in the CP population.

The Child Health Questionnaire

Developed by the Child Health Project in the early 1990s (see Landgraf et al., 1996 for full description), the conceptual underpinnings of the CHQ are consistent with the definition of WHO (1948) of health as a multidimensional construct. Landgraf and colleagues (1996) conceptualized children's health as comprising two dimensions: physical and psychosocial health. A disadvantage in either dimension is assumed to have

an adverse impact on an individual's ability to participate fully in social roles (Landgraf et al., 1996).

By reviewing the literature and instruments in use at that time, the Child Health Project team identified a number of dimensions that were perceived as being common to the health and well-being of all children (Landgraf et al., 1996). Exploratory factor analysis confirmed that these dimensions mapped onto one or the other of the two hypothesized, exclusive concepts of physical and psychosocial health. The multidimensional and comprehensive nature of the CHQ suggests that it may be an appropriate tool with which to assess the health and well-being of children with CP.

There are four versions of the CHQ: the self-report version for 10–18 year olds (CF87) and three parent forms designed for 5–18 year olds (PF98, PF50, and PF28). The PF50 and PF28 were derived from the original, longer form PF98 but despite being in the manual the PF98 has no norms or summary scores and is no longer supported (Landgraf, 2007, personal communication). The PF50 version is the most frequently used (Drotar, Schwartz, Palermo, & Burant, 2006). It has 13 single- and multi-item scales across a number of domains relating to "the last four weeks" with an additional global item assessing changes in health "over the last year." The physical domain includes the physical functioning scale (PF) that assesses the presence and level of physical limitations due to ill health, the role/social limitations—physical scale (RP) that measures limitations in school and friend related activities as a consequence of physical health problems, the general health perceptions scale (GH) that provides an overall subjective measure of health and illness, and the bodily pain scale (BP) which evaluates the intensity of general pain (Landgraf et al., 1996). The psychosocial domain of the CHQ includes the role/social limitations—emotional/behavioral scale (REB) that assesses restrictions in school and friend-related activities as a consequence of emotional/behavioral difficulties, the self-esteem scale (SE) that assesses satisfaction with school and athletic ability, looks/appearance, ability to get along with others and family and life overall, the mental health scale (MH) that assesses positive and negative states such as anxiety and depression, the general behavior scale (BE) that measures overt behavior e.g., aggression and hyperactivity, the parental impact-emotional (PE) and parental impact-time (PT) scales that assess parents level of distress and the reduction in personal time as a consequence of the child's illness, and the family activities scale (FA), which considers the extent to which the child's illness disrupts normal family activities. The family cohesion item looks at how well a family gets along together (Landgraf et al., 1996). Responses are scored for each domain producing a figure between 0–100, with lower scores indicating poorer health and well-being. Both the PF50 and PF28 have the advantage of two summary scores that represent physical (PhS) and psychosocial (PsS) health and are calculated to have a mean of 50 and SD of 10 in the normative population (Landgraf et al., 1996).

Normative data has been reported for children in the United States ($n = 391$; Landgraf et al., 1996) and Australia ($n = 5,414$; Waters et al., 2000), with extensive information available for school children in the Netherlands ($n = 2,474$; Raat, Botterweck, Landgraf, Hoogeveen, & Eissink-Bot, 2005). The CHQ has been used in over 30 countries and validated into at least 21 languages (Morales, et al., 2006; Raat et al., 2005; Ruperto et al., 2001). It has been applied to healthy children, those with a wide range of illnesses and chronic health conditions and to socially disadvantaged children.

Method

A literature search was carried out to identify studies that had utilized some or all domains of the CHQ in children with CP. The search terms used were "child health questionnaire," "CHQ," "quality of life," and "CP." The following databases were searched (January 1993–January 2007): Cumulative Index to Nursing & Allied Health Literature (CINAHL), Ovid MEDLINE, PsychINFO, and Science and Social Sciences Citation Index (Web of Science). Further, papers were identified by hand-searching the reference lists of published papers. To be included in the review, the papers had to be published in English and applied exclusively to children with CP. Studies that integrated children with a range of chronic conditions were excluded on the grounds that it would be difficult to reliably distinguish and report on the descriptive and psychometric data relating to children with CP. Where possible effect sizes have been calculated to determine the magnitude of differences in health between studies and samples according to the APA Guidelines (2001, Fifth Edition, p.25) and are presented descriptively in the text.

Results

Thirteen published papers were found that had used all or some of the scales and summary scores of the CHQ

Table I. Summary of Published Studies ($n = 13$)

Study & questionnaire	Recruitment method and sample size	Sample characteristics (gender, age, ethnicity and SES)	Clinical characteristics
Fung et al. (2002) PF28, (US).	Cross-sectional, multiple site, multiple methods (NAGCPP), $n = 230$.	59% male; 2.0–17.9 years (9.7 ± 4.6); 68% Caucasian, 47% African American, 8% other.	GMFCS III ($n = 54$), IV ($n = 55$), V ($n = 119$).
Houlihan, O'Donnell, Conaway, & Stevenson (2004) PF28, (US).	Prospective, multiple site, multiple methods (NAGCPP), $n = 198$.	58% male; 5–18 years (10.7 ± 3.11); 68% Caucasian, 47% African American, 8% other.	GMFCS III ($n = 44$), IV ($n = 44$), V no gastrostomy ($n = 119$), V with gastrostomy ($n = 44$).
Liptak et al. (2001) PF28, (US).	Cross-sectional, multiple site, multiple methods (NAGCPP), $n = 235$.	59% male; 2–18 years (9.6); 68% Caucasian, 47% African American, 8% other.	GMFCS III ($n = 56$), IV ($n = 55$), V ($n = 122$).
McCarthy et al. (2002) PF98/ITHQ, (US).	An orthopedic clinic, $n = 115$.	58% male; 3–10 years (5.8).	Diplegic ($n = 49$), Hemiplegic ($n = 28$), Quadriplegic ($n = 38$).
Morales et al. (2006) PF50, (Brazil).	Rehabilitation clinic, $n = 96$ (analysis based on 92).	56% male; 5–18 years; 71% Caucasian, 29% African Brazilian.	Diplegic ($n = 35$), Hemiplegic ($n = 24$), Quadriplegic ($n = 22$), Dyskinetic ($n = 11$).
Pipiris & Graham (2004) PF50, (US).	Consecutive cohort attending weekly clinic, $n = 300$.	49% male; (11 ± 3) years.	Diplegic ($n = 113$), Hemiplegic ($n = 115$), Quadriplegic ($n = 72$).
Samson-Fang et al. (2002) PF28, (US).	Cross-sectional, multiple site, multiple methods (NAGCPP), $n = 235$.	59% male; 2.0-17.9 years (9.7 ± 4.6); 68% Caucasian, 47% African American, 8% other.	GMFCS III (23%), IV (24%), V (52%).
Schneider, Gurucharri, Gutierrez, & Gaebler-Spira (2001) PF28, (US).	Convenience, clinic and out-patient pediatric physical therapy department, $n = 30$.	43% male; 5–15 years (8.5); Africa-American ($n = 18$), White ($n = 8$), Hispanic ($n = 3$), Middle Eastern ($n = 1$).	GMFCS I ($n = 3$), II ($n = 1$), III ($n = 10$), IV ($n = 12$), V ($n = 4$).
Vargus-Adams (2005) PF50, (US).	Out-patients clinic at 1 hospital (time 1), $n = 177$.	55% male; 3–18 years (8.6 ± 4.2); 84% White, 10% Black, 5% other; Full-time carer $n = 17$, employed $n = 60$, homemaker $n = 29$.	GMFCS I ($n = 70$), II ($n = 24$), III ($n = 24$), IV ($n = 29$), V ($n = 30$).
Vargus-Adams (2006) PF50, (US).	Follow-up, 3 time points over 1 year, $n = 177$.	As above	As above
Vitale et al. (2005) PF28, (US).	Prospective, hospital population. $n = 180$.	No sex distribution given; 5–18 years (10.7); States 'mostly Caucasian'.	Diplegic ($n = 46$), Hemipelgic ($n = 45$), Quadriplegic ($n = 87$).
Wake, Salmon, & Reddihough (2003) PF50, (Australia).	Out-patients clinic, $n = 80$.	56% male; 5–18 years (11.4 ± 3.6).	GMFCS I/2 ($n = 38$), III ($n = 7$), IV/V ($n = 33$).
Wallen, O'Flaherty, & Waugh (2004) PF50, (US).	Convenience, disabilities clinic, $n = 16$.	68% male; 1–12 years; (6.38 ± 3.24).	Triplegic ($n = 1$), Hemipelgic ($n = 7$), Quadriplegic ($n = 8$).

and met the inclusion criteria (Table I). Ten studies were based in the US, two in Australia, and one in Brazil. In summary, the CHQ was employed as a measurement tool to describe children's health status (Liptak et al., 2001; Vargus-Adams, 2005, 2006; Wake et al., 2003); to explore the nature of the relationship between characteristics of CP and health status (Fung et al., 2002; Houlihan et al., 2004; Samson-Fang et al., 2002); to assess the outcomes of interventions (Wallen et al., 2004); to validate alternative questionnaires (McCarthy et al., 2002; Pirpiris, & Graham, 2004; Schneider et al., 2001; Vitale et al., 2005); and to explore the

psychometric performance of the CHQ in a CP population (McCarthy et al., 2002; Morales et al., 2006; Wake et al., 2003).

The 13 published papers relate to nine distinct datasets of children with CP. A total of 1,229 unique children, aged 2–18 years participated in the studies published between 2001–06. The majority of children were described as having ''moderate'' to ''severe'' CP and more than half the subjects were male (CP is more common in males). Most studies used the Gross Motor Function Classification System (GMFCS) (Palisano et al., 1997) to group children by severity (Fung et al., 2002;

Houlihan et al., 2004; Liptak et al., 2001; Samson-Fang et al., 2002; Schneider et al 2001; Vargus-Adams, 2005, 2006; Wake et al., 2003); five studies defined children by clinical subtype (McCarthy et al., 2002; Morales et al., 2006; Pipiris & Graham, 2004; Vitale et al., 2005; Wallen et al., 2004). All researchers used the parent form of the CHQ, with six administering the PF28 (Fung et al., 2002; Houlihan et al., 2004; Liptak et al., 2001; Samson-Fang et al., 2002; Schneider et al., 2001; Vitale et al., 2005) and six the PF50 (Morales et al., 2006; Pipiris & Graham, 2004; Vargus-Adams, 2005, 2006; Wake et al., 2003; Wallen et al., 2004). A further study had utilized the PF98 version of the CHQ in conjunction with the Infant Toddler Health Questionnaire (ITHQ) (McCarthy et al., 2002), designed for younger children.

Psychometric Performance of the CHQ

Of the 13 studies, only three provided information on the reliability of the CHQ within a CP population (McCarthy et al., 2002; Morales et al., 2006; Wake et al., 2003), (Table II). However, these three studies reported only on the internal consistency of the CHQ using Cronbach's alpha (α). McCarthy et al. (2002) reported that the internal consistency of the CHQ was satisfactory for a number of subscales and further that item homogeneity was found to be adequate for assessing group differences. Wake and colleagues (2003) also reported that internal consistency was satisfactory for all scales in the PF50. Morales and colleagues (2006) found that the reliability for the CHQ was satisfactory for all subscales with the exception of the general health perceptions scale.

There was relatively more information available on the validity of the CHQ within the CP population. Six studies provided evidence of concurrent validity by comparing the performance of the CHQ with other instruments assessing disability or functional ability that were designed for use in children with CP (McCarthy et al., 2002; Morales et al., 2006; Vargus-Adams, 2005); physicians' subjective ratings of children's physical and psychosocial health (Vitale et al., 2005); the presence of associated medical conditions (Vargus-Adams, 2005); a tool designed to assess the time children spent in an upright position (Pipiris & Graham, 2004); and with the Caregiver Questionnaire or CQ (Revivo, 2000; Schneider et al., 2001). Of these studies, four also reported on discriminant validity and two provided information on item discriminant validity of the CHQ (Table II).

Explicitly reporting on the concurrent validity of the PF98/ITHQ, McCarthy et al. (2002) found that scores on the physical functioning scale correlated strongly with scores on the Gross Motor Function Measure or GMFM (Russell et al., 1989), the associated physical/mobility scales of the Pediatric Evaluation of Disability Inventory or PEDI (Feldman, Haley, & Coryell, 1990), and the Pediatric Outcomes Data Collection Instrument or PODCI (Daltroy, Liang, Fossel, & Goldberg, 1998). Morales et al. (2006) also reported that the GMFM correlated strongly with the physical summary scores and the physical functioning scale, and moderately with the role/social limitations-physical scale of the CHQ. In contrast, the psychosocial summary score of the CHQ did not correlate with scores on the GMFM, although there was a moderate correlation with the role/social limitations-emotional scale. Vargus-Adams (2005) reported that scores in a number of scales in the CHQ were negatively associated with scores on the GMFCS (general health perceptions, physical functioning and parental impact-time, and role/social limitations-physical and the physical summary score). Conversely, scores on the behavior scale were positively associated with the GMFCS.

Two studies (Pipiris & Graham, 2004; Schneider et al., 2001) utilized the CHQ as a criterion measure, therefore, implicitly providing information on its concurrent validity. Pipiris & Graham (2004) used the CHQ as criteria with which to determine the psychometric properties of an instrument designed to assess time spent in an upright position. High positive correlations were observed between the amounts of time that children spent upright and scores on the physical functioning ($r = .81$) and the role/social limitations-physical ($r = .76$) scales of the CHQ, suggesting that the parents of children who were ambulatory perceived their physical well-being to be better than those whose mobility was limited. In the study assessing the performance of the CQ, Schneider and colleagues (2001) found moderate positive correlations between the parental impact-time scale and the total score and physical subscale score of the CQ ($r = .43$ and $r = .38$, respectively) and also between the family cohesion item of the CHQ and the interaction subscale of the CQ ($r = .48$).

The CHQ was designed to assess two distinct dimensions of health in children: physical and psychosocial well-being. Vitale and colleagues (2005) correlated the corresponding dimensions of the CHQ with subjective physical and psychosocial scores provided by a physician. They reported that positive associations were observed between the respective dimensions of health (ranging from 0.17 to 0.53 for the physical dimensions and 0.02 to 0.44 for the psychosocial dimensions),

Table II. Reliability and Validity of CHQ Scales and Summary Scores in Children with Cerebral Palsy

Study	Statistic	CHQ scales and summary scores													
Reliability		PF	RB	RP	BP	BE	MH	SE	GH	PE	PT	FA	FC	PhS	PsS
Internal consistency															
McCarthy et al. (2002)	Cronbachs α	0.94	–	–	0.92	0.80	0.81	–	0.84	–	–	–	–	–	–
Morales et al. (2006)	Cronbachs α	0.93	0.89	0.94	0.76	0.58	0.60	0.73	0.24	0.72	0.75	0.67	–	–	–
Wake et al. (2003)	Cronbachs α	←————————————————— α reported as ranging 0.75–0.97 across all scales ————————————————→													
Validity															
Concurrent															
Vargus Adams (2005)	Kendall's τ (GMFCS)	−0.51	−0.08	−0.27	−0.16	0.29	−0.01	−0.03	−0.29	−0.06	−0.25	–	–	−0.42	0.09
McCarthy et al. (2002)	Spearman partial (r) (GMFM)	0.74	–	–	−0.12	−0.30	−0.12	–	0.14	–	–	–	–	–	–
	(PEDI) Mobility	0.75	–	–	−0.03	−0.24	−0.03	–	0.16	–	–	–	–	–	–
	Self care	0.62	–	–	−0.11	−0.15	0.03	–	0.15	–	–	–	–	–	–
	Social func.	0.30	–	–	−0.04	0.01	0.17	–	0.17	–	–	–	–	–	–
	(PODCI) Mobility	0.81	–	–	0.01	−0.22	−0.02	–	0.20	–	–	–	–	–	–
	Arm func.	0.68	–	–	−0.05	−0.17	0.10	–	0.11	–	–	–	–	–	–
	Pain	−0.07	–	–	0.68	0.13	0.27	–	0.18	–	–	–	–	–	–
Morales et al. (2006)	Pearson's (r) (GMFM)	0.72**	0.53**	0.65**	0.10	−0.28**	−0.13	0.32**	–	−0.03	0.17	0.20	−0.06	0.74**	0.00
Discriminant															
McCarthy et al. (2002)	MANOVA (F) (Physical)	33.1***	–	–	0.1	9.9***	3.2*	0.6	0.8	–	–	–	–	–	–
	(Cognitive)	0.5	–	–	1.3	0.3	0.6	1.2	0.0	–	–	–	–	–	–
Morales et al (2006)	As above (physical)	21.8**	–	9.9*						–	–	–	–	22.0**	–
Wake et al. (2003)	Independent t-test (p)	0.92	0.47	0.82	0.16	0.77	0.64	0.95	0.97	0.45	0.32	0.49	0.30	0.60	0.52
	Epilepsy	0.82	0.21	0.54	0.88	0.85	0.14	0.02	0.64	0.11	0.95	0.83	0.04	0.91	0.15
	Severity	<0.001	0.19	0.001	0.04	<0.001	0.67	0.37	0.02	0.93	0.16	0.03	0.13	<0.001	0.33
Item discriminant															
Morales et al. (2006)	Pearson's (r) (%)	100	100	100	100	–	100	100	100	100	100	100	–	–	–
Wake et al. (2003)	Revised multitrait analysis (%)	100	100	100	100	100	100	100	89.4	93.9	93.9	100	100	100	100

*p < 0.05, **p < 0.01, ***p < 0.001.

PF, physical functioning; RB, role/social limitations-emotional-behavioral; RP, role physical; BP, bodily pain; BE, behavior; MH, mental health; SE, self esteem; GH, general health perceptions; PE, parental impact emotion; PT, parental impact time; FA, family activities; FC, family cohesion; PhS, physical summary score; PsS, psychosocial summary score. Func., Function.

Table III. Child Health Questionnaire Mean (SD) Scale Scores for Normative Samples and Cerebral Palsy Studies (Where Available)

Country	United States						Australia		Brazil
	Normative Landgraf et al. (1996) n = 391	CP data McCarthy et al. (2002) n = 115	CP data Vitale et al. (2005) n = 180	CP data Vargus-Adams (2005) n = 1770	CP data NAGCPP (2000–2004) n = 233		Normative data Waters et al. (2000) n = 5,414	CP data Wake et al. (2003) n = 80	CP data Morales et al. (2006) n = 96
Studies Scales	Mean (SD)	Mean (SD)	Mean	Mean (SD)	Mean (SD)	z-score	Mean (SD)	Mean (SD)	Mean (SD)
PF	96.1 (13.9)	49.0 (36.5)	39.6	46.1 (39.2)	-	-	94.6 (15.1)	47.9 (39.4)	45.9 (35.4)
REB	92.5 (18.6)	-	77.8	65.6 (36.6)	-	-	93.7(17.2)	64.0 (39.0)	54.3 (35.5)
RP	93.6 (18.6)	61.4 (36.0)	57.7	55.7 (40.1)	-	-	94.2 (17.2)	55.6 (41.5)	47.9 (37.9)
BP	81.7 (19.0)	73.9 (19.7)	67.8	66.3 (25.8)	-	−0.38	82.3 (18.8)	67.4 (27.6)	73.5 (21.8)
BE	75.6 (16.7)	71.9 (14.0)	69.9	71.6 (16.9)	-	-	77.5 (15.2)	68.4 (21.6)	68.1 (16.6)
MH	78.5 (13.2)	80.2 (12.2)	75.2	71.1 (13.8)	-	-	80.2 (12.4)	69.5 (17.3)	71.6 (13.6)
SE	79.8 (17.5)	77.8 (18.3)	76.6	68.9 (19.3)	-	-	79.7 (16.4)	68.3 (18.7)	75.5 (19.4)
GH	73.0 (17.3)	57.5 (18.2)	56.6	53.7 (23.8)	-	-	77.0 (16.0)	51.8 (24.2)	65.9 (15.0)
PE	80.3 (19.1)	-	51.6	49.4 (25.7)	-	−0.41	80.6 (20.1)	54.7 (27.7)	51.8 (28.3)
PT	87.8 (19.9)	-	62.9	61.5 (31.6)	-	−1.03	91.5 (16.4)	58.4 (33.4)	62.2 (27.0)
FA	-	-	-	-	-	-	85.5 (16.7)	59.3 (27.0)	79.0 (16.4)
FC	-	-	-	-	-	-	76.1 (20.7)	71.2 (21.5)	68.3 (25.8)
RSB[a]	-	83.9 (30.5)	-	-	-	-	-	-	-
RE[a]	-	87.2 (26.7)	-	-	-	-	-	-	-
GGH	-	74.2 (24.2)	-	-	-	−0.79	-	-	69.4 (22.2)
GBE	-	73.1 (27.6)	-	-	-	-	-	-	68.4 (26.9)
PhS	53.0 (8.8)	32.2 (16.6)	-	29.9 (19.0)	37.0 (17.0)	−1.29	50.0 (9.9)	31.7 (19.0)	34.1 (12.2)
PsS	51.2 (9.1)	50.7 (10.8)	-	45.0 (10.0)	-	-	49.9 (10.0)	43.4 (11.6)	41.0 (8.7)

PF, physical functioning; RB, role/social limitations-emotional-behavioral; RP, role physical; BP, bodily pain; BE, behavior; MH, mental health; SE, self esteem; GH, general health perceptions; PE, parental impact emotion; PT, parental impact time; FA, family activities; FC, family cohesion; GGH, global health item; GBH, global behaviour item; PhS, physical summary score; PsS, psychosocial summary score.
[a]One scale in the PF50 version.

with the strongest correlations being found in the hemiplegic group. Vitale et al. (2005) also reported that the internal validity of the CHQ was satisfactory, with strong correlations being present between the scales that represented physical health and also between the scales that represented psychosocial health. Further, evidence of the internal validity of the CHQ was provided by McCarthy and colleagues (2002) who noted that the subscales representative of psychosocial health on the CHQ/ITHQ were negatively or weakly correlated with those subscales that measured physical health. Morales and colleagues (2006) also reported that the correlation between the physical and psychosocial summary scores was weak.

The discriminative validity of the CHQ was demonstrated by those studies that identified differences when comparing scores in the CP population to scores in the normative population (Table III). In general, children with CP had lower mean scores (and thus poorer health) on both the physical and psychosocial summary scores of the CHQ, when compared to their country-specific norms. Within the studies carried out in the US, McCarthy et al. (2002), Vargus-Adams (2005), and the North American

Growth in Cerebral Palsy Project (NAGCPP) group reported that children with CP had lower scores on the physical summary scores than the normative population, with large effect sizes of 2.36, 2.62, and 1.81, respectively. However, McCarthy and colleagues and Vargus-Adams found smaller differences for the psychosocial summary scores with small and moderate effect sizes of 0.05 and 0.68, respectively. The Australian study detected a large effect size (1.84) between the physical summary scores of the CP sample and the normative population, where children with CP experienced poorer physical health (Wake et al., 2003). Differences in the psychosocial summary scores were moderate (0.65) and in the Brazilian study (Morales et al., 2006) children with CP were found to have a significantly lower physical summary scores than local healthy school children with a large effect size of 4.44, in addition to having significantly lower psychosocial summary scores (effect size = 1.84).

Evidence of discriminative validity was also observed at subscale level of the CHQ, with scores in the CP population being lower than those of the relative country-specific norms. In all domains, children with CP had the lowest scores (indicating poorer health; Table III).

In the US studies (McCarthy et al., 2002; Vargus-Adams, 2005; Vitale et al., 2005), large effect sizes (0.8–3.59) were found for general health perceptions, parental impact-emotion, physical functioning, parental impact-time, and role/social limitations-emotional/behavioral and role/social limitations-physical scales. In contrast, small effect sizes (0.1–0.4) were found for behavior, mental health, and self-esteem scales. There was one instance in which children with CP exhibited a higher score on the mental health scale than the US normative population (McCarthy et al., 2002), but this difference was diminutive (0.1). Within the Australian study (Wake et al., 2003), a small effect size was present in the family cohesion scale (0.23) with moderate effect sizes found in relation to behavior (0.5) and self-esteem (0.69). The remaining scales exhibited large effect sizes (range 0.86–3.09) and in the study by Morales et al., large effect sizes (range 0.82–6.65) were present for all domains with the exception of family cohesion that was moderate (0.51).

There was also some indication that the CHQ could discriminate within the CP population as a function of severity. Wake and colleagues (2003) found that the CHQ discriminated between children with severe and mild CP on a number of scales. Children with severe CP had significantly lower physical functioning, role/social limitations-physical, bodily pain, behavior, general health perceptions, family activities, and physical summary scores than those with mild CP. Vitale et al. (2005) reported that scores on a number of domains in the CHQ were lower for children in the quadriplegic subgroup in comparison to children in the hemiplegic and diplegic groups. In contrast, McCarthy et al. (2002) and Morales et al. (2006) found that the CHQ had limited discriminative properties in that it was only able to detect differences between either quadriplegic and hemiplegics or quadriplegic and diplegic groups, dependent upon the scale in question (Table II).

In examining the items within the subscales, Wake et al. (2003) assessed the item discriminant validity of the CHQ. Reporting scaling success they found it ranged from 89.4% to 100% with 8 of 11 scales achieving 100%, and with all but four items being found to correlate more strongly with their hypothesized scale than with other scales. Morales et al. (2006) reported similar findings with the exception of "general health perception" that had a success rate of 50%.

Determining the responsiveness of an instrument is an important consideration, particularly if it is to be used as an outcome measure to identify significant changes in health that may occur progressively over time, or as a consequence of therapeutic intervention (Guyatt, Feeny, & Patrick, 1993). Two of the studies provide some evidence regarding the responsiveness of the CHQ within CP populations (Vargus-Adams, 2006; Wallen et al., 2004). In describing the change and stability in children's health at 6 and 12 month's following a baseline assessment, Vargus-Adams (2006) found that it was only scores on the role/social limitations-physical scale of the PF50 that exhibited evidence of change across time. Wallen et al. (2004) used the CHQ as an outcome measure to assess the effectiveness of an intervention designed to improve upper limb functioning in children with CP. While other outcome measures did indicate evidence of an improvement following intervention, Wallen and colleagues found no significant difference in scores on the CHQ after baseline assessment.

Discussion

The psychometric properties of an instrument can fluctuate between various populations (McHorney, Ware, Lu, & Sherbourne, 1994) and researchers should attempt to verify and report some psychometric information when assessing the health and well-being of children. Yet, of the thirteen papers identified, less than half reported on one or more psychometric criteria of the CHQ when applied to a CP sample and even then limited information was available.

Information on reliability was provided by three studies with the consensus being that the CHQ exhibited satisfactory levels of internal consistency across all domains and summary scores when used with children with CP. The one exception was the reliability of the general health perceptions scale in Morales et al. (2006) study that was notably low. Morales and colleagues reported that this may be a consequence of cultural adaptation of the CHQ, which was exacerbated further by the subjective nature of the subscale. Despite the availability of several types of reliability estimates, analysis was confined exclusively to internal consistency. Reliability may be measured in a number of ways including internal consistency, test–retest, and inter-rater reliability (Eiser & Morse, 2001) and may vary depending on how it is measured, because each type of estimate may be influenced by a different source of measurement error. Given the issues identified earlier and the paucity of data on the reliability of the CHQ within the CP population, there is a need for further studies to

be carried out, with a focus on testing additional forms of reliability, e.g., test–retest reliability (Waters et al., 2000).

More evidence was provided on the validity of the CHQ. In general, the concurrent validity of the CHQ in CP samples was found to be satisfactory. However, it must be noted that the instruments with which the CHQ was compared were primarily assessing functional ability. Given that the purpose of the CHQ is to assess both the physical and psychosocial well-being of children, the failure to validate the CHQ with a suitable psychosocial alternative instrument in the CP population suggests the need for more work in this area. For example, the Strengths and Difficulties Questionnaire is a generic, behavioral screening measure for children whose psychometric performance has been reported extensively (Goodman, 1997, 2001) and functions well at detecting symptoms (emotion, conduct, hyperactivity, and peer problems), and the impact on the child and family.

There was some evidence that the CHQ did assess two distinct dimensions of health, namely physical and psychosocial, in children with CP. However, none of the studies had explicitly considered the factor structure of the CHQ in the CP population. This is one type of validity that remains very much understudied (Drotar et al., 2006) in all populations that have utilized the CHQ. Besides the exploratory factor analysis by Landgraf et al. (1996) and Waters et al. (2000), only two others (Drotar et al., 2006; Hepner & Sechrest, 2002) have conducted a confirmatory factor analysis on the PF50. Different models were found for healthy children and those with chronic illnesses (Drotar et al., 2006), so it is imperative from a theoretical and applied perspective that the factor structure of the CHQ within children with CP is explored. Otherwise, these findings have the potential to undermine the confidence with which the physical and psychosocial summary scores (and indeed the subscales) can be used to make reliable comparisons across studies of different childhood populations.

Some concerns were also raised regarding the appropriateness of a number of items within the CHQ in relation to children with severe CP, particularly those relevant to physical functioning (Schneider et al., 2001; Wake et al., 2003). Items such as "bike riding" and "soccer" were perceived by some parents as being inappropriate as a consequence of the severity of their child's condition (Schneider et al., 2001). Schneider and colleagues also pointed out that the CHQ did not address issues related to transferring or handling and as such was of limited value for children with CP. The presence of

floor and ceiling effects reported by some also indicates that the CHQ may not have provided an accurate profile of children's health. Ceiling effects within the psychosocial domains were more common, with floor effects observed more so in the physical domains. Overall this suggests poor face validity of the parent report CHQ when applied to children with severe CP. When occasions such as this arise, it may be more appropriate to apply a disease-specific tool in conjunction with the CHQ to elicit a more comprehensive profile of children's health.

Overall, the CHQ demonstrated that children with CP exhibited poorer health and well-being compared to able-bodied children, suggesting that it had good discriminative properties with large effect sizes found in relation to the physical summary score. Given that CP is primarily a physical disability this finding was not unexpected. Psychosocial health was found to be better than physical health, with children in the CP populations exhibiting scores representative of psychosocial well-being similar to the normative population. The CHQ also showed that in some instances as severity of CP increased, levels of psychosocial functioning also increased, indicating that the parents of children with more severe CP perceived their children to have better psychological health than children with less severe CP. It is possible that children with severe CP have lower levels of psychological difficulties as for example, conduct disorders, where as a result of their physical impairments they are unable to take part in hitting, stealing, or fighting. It is also possible that intellectual and communication problems, which are more common in children with severe CP, mean the child's inner world remains relatively remote and inaccessible to those individuals making proxy assessments about the child's psychological adjustment and emotional well-being.

The extent to which these relationships can be generalized to the CP population may be somewhat limited. None of the studies identified presented results across the spectrum of CP. The studies conducted by the NAGCPP chose to exclude children who had coexisting medical conditions that may have had an impact on their growth, for example, congenital heart disease. While they rationalize this decision with regard to their research hypothesis, it does limit the extent to which their sample may be viewed as being representative. Indeed, few of the studies included children with a full range of motor abilities with a disproportionate emphasis on moderately and severely affected children. Future research should consider the inclusion of children with milder forms of CP. In addition to providing valuable information on the

health status of these children, such studies could also contribute to the construct validity of the CHQ in the CP. In particular, as significantly fewer of these children have cognitive impairments the rarely used self-completed version of the CHQ could be employed.

Within the CP population there were variations in scores on the CHQ within- and between-countries. Some of these differences may reflect real variation in child health, parental expectations, the level of support, and help that families receive in different countries, but they may also reflect a number of important methodological issues. For example, none of the studies reported on a standardized approach to the validation of the diagnosis "CP" or on the method used to classify children by clinical subtype. Validating a diagnosis of CP is a particular concern, when including children <5 years because young children can outgrow the early signs of motor impairment (Nelson & Ellenberg, 1982). Seven studies included young children and five studies classified their sample by clinical subtype using the ambiguous terms "diplegia," "triplegia," and "quadriplegia" thus limiting the extent to which comparisons can be made between studies. On a more positive note, nine studies classified the children using the GMFCS, a valid and reliable way of differentiating children with CP by functional abilities (Palisano et al., 1997). Inconsistencies in confirming the diagnosis and defining clinical subgroups highlights the importance for future research to adopt a more systematic and standardized approach as recommended by the Surveillance of Cerebral Palsy in Europe project (SCPE, 2000).

Furthermore, the means by which CP samples were identified and recruited limits the conclusions that can be drawn from the studies conducted thus far. The majority of studies recruited from hospital clinics and yet a previous analysis has shown this source to be incomplete and biased in terms of the clinical characteristics of attendees (Parkes, Kerr, McDowell, & Cosgrove, 2006). Multiple ascertainments of children with CP is important as these children tend to see many professionals and have no "one point" of contact with services. There are a number of geographically-defined population registers across Europe (SCPE, 2000) and elsewhere which could facilitate researchers in recruitment of children with CP. One study in particular—SPARCLE (Colver, 2006)—has just adopted such an approach to investigating participation and quality of life in 8–12 year-old children with CP in seven European countries. The SPARCLE study has included measures like the CHQ and will provide, for the first time, data from western European countries, across the severity spectrum of CP in children systematically recruited from population-based registers whose methods comply with SCPE guidelines.

It is important that comparative studies should attempt to use, where available, norms relative to the country or region within which the research is being conducted, ideally collated, and systematically matched to the clinical sample on a number of possible confounding factors such as socio-economic status. Whilst norms for the CHQ are available for a range of populations there is no normative data published for children in Western Europe or more specifically the United Kingdom and this needs to be addressed. Where possible future studies should systematically match samples on characteristics such as gender, age, ethnicity, and socio-economic status, yet the studies presented here provide few details. These sorts of variables have been shown to influence children's health status and perceptions of what constitutes "good health" (Landgraf & Abetz, 1997), yet none of the studies included here examined the data to determine if there were differences as a function of these variables.

Within the studies identified, attention also needs to be drawn to the age of the children included in the samples. The CHQ was designed for use in children aged 5–18 years, yet six studies included children that were below 5 years of age. McCarthy and colleagues (2002) did attempt to address this issue by including a questionnaire relevant to younger children. However, the remainder of the studies did not justify the administration of the CHQ to children younger than five years. Variability in the scores across the studies may in part have been a consequence of the inclusion of younger children. Children in this stage of life are subject to rapid growth and dynamic development. Furthermore, younger children may show greater dependency on adults for care, making it difficult to determine the extent to which the condition of CP resulted in greater demands in areas such as family activities.

Finally, generic questionnaires have been reported as having lower responsiveness to change and less sensitivity than disease-specific measures (Eiser & Morse, 2001; Guyatt et al., 1993). The CHQ did not prove to be a suitable tool with which to assess the outcome of an intervention in children with CP (Wallen et al., 2004) again indicating that in certain instances the CHQ may not be suited to this population, although this is based on evidence from one study. It may be that the CHQ could demonstrate adequate responsiveness

to change when applied in relation to different interventions.

Whilst a number of studies concluded that the CHQ was a reliable, valid, and acceptable generic questionnaire for use in children with CP (Morales et al., 2006; Wake et al., 2003), the authors of this article conclude that further evaluative work needs to be carried out. The paucity of information available about its psychometric viability and indeed some problem areas identified in its performance are concerns that need to be addressed. In particular, future work involving confirmatory factor analysis is essential. The scarcity of normative data on substantial representative samples also needs to be addressed. Until such times, researchers and clinicians using the CHQ in children with CP should exercise sensitivity in its administration and caution in its interpretation.

Conflict of interest: None declared.

References

American Psychological Association. (2001). *Publication manual of the American Psychological Association* (5th ed.). Washington, DC: Author.

Colver, A. (The SPARCLE group) (2006). Study protocol: SPARCLE- a multi-centre European study of the relationship of environment to participation and quality of life in children with cerebral palsy. *BMC Public Health*, 6,105. Retrieved August 23, 2007, from http://www.biomedcentral.com/1471-2458/6/105.

Daltroy, L. H., Liang, M. H., Fossel, A. H., & Goldberg, M. J. (1998). The POSNA Pediatric Musculoskeletal Functional Health Questionnaire: Report on reliability, validity and sensitivity to change. *Journal of Pediatric Orthopedics, 18*, 561–571.

Drotar, D., Schwartz, L., Palermo, T. M., & Burant, C. (2006). Factor structure of the child health questionnaire-parent form in pediatric populations. *Journal of Pediatric Psychology, 31*(2), 127–138.

Eiser, M., & Morse, R. (2001). Quality of life measures in chronic diseases of childhood. *Health Technology Assessment, 5*, 4.

Feldman, A. B., Haley, S. M., & Coryell, J. (1990). Concurrent and construct validity of the Pediatric Evaluation and Disability Inventory. *Physical Therapy, 70*, 602–610.

Fung, E. B., Samson-Fang, L., Stallings, V. A., Conaway, M., Liptak, G., Henderson, R. C., et al. (2002). Feeding dysfunction is associated with poor growth and health status in children with cerebral palsy. *Journal of the American Dietetic Association, 102*(3), 361–373.

Goodman, R. (1997). The Strengths and Difficulties Questionnaire: A research note. *Journal of Child Psychology and Psychiatry, 38*, 581–586.

Goodman, R. (2001). Psychometric properties of the Strengths and Difficulties Questionnaire. *Journal of the American Academy of Child and Adolescent Psychiatry, 40*, 1337–1345.

Guyatt, G. H., Feeny, D. H., & Patrick, D. L. (1993). Measuring health-related quality of life. *Annals of Internal Medicine, 118*(8), 622–629.

Hepner, K. A., & Sechrest, L. (2002). Confirmatory factor analysis of the Child Health Questionnaire-Parent Form 50 in a predominantly minority sample. *Quality of Life Research, 11*(8), 763–773.

Houlihan, C. M., O'Donnell, M., Conaway, M., & Stevenson, R. D. (2004). Bodily pain and health-related quality of life in children with cerebral palsy. *Developmental Medicine and Child Neurology, 46*(5), 305–310.

Landgraf, J. M., & Abetz, L. (1997). Influences of sociodemographic characteristics on parental reports of children's health and well-being: Early experiences with the Child Health Questionnaire. In D. Drotar (Ed.), *Quality of life assessment in children.* New Jersey: Lawrence-Erlbaum Press.

Landgraf, J. M., Abetz, L., & Ware, J. E. (1996). *The Child Health Questionnaires Users' Manual* (1st ed.). Boston: The Health Institute: New England Medical Centre.

Liptak, G. S., O'Donnell, M., Conaway, M., Chumlea, W. C., Worley, G., Henderson, R. C., et al. (2001). Health status of children with moderate to severe cerebral palsy. *Developmental Medicine and Child Neurology, 43*(6), 364–370.

McCarthy, M. L., Silberstein, C. E., Atkins, E. A., Harryman, S. E., Sponseller, M. D., & Hadley-Miller, N. A. (2002). Comparing reliability and validity of pediatric instruments for measuring health and well-being of children with spastic cerebral palsy. *Developmental Medicine and Child Neurology, 44*(7), 468–476.

McHorney, C. A., Ware, J. E., Lu, J. F. R., & Sherbourne, C. D. (1994). The MOS 36-Item Short-Form Health Survey (SF-36): III. Tests of data

quality, scaling assumptions and reliability across diverse patient groups. *Medical Care, 32*, 40–66.

Morales, N. M. O., Silva, C. H. M., Frontarolli, A. C., Araújo, R. R. H., Rangel, V. O., Pinto, R. M. C., et al. (2006). Psychometric properties of the initial Brazilian version of the CHQ-PF50 applied to the caregivers of children and adolescents with cerebral palsy. *Quality of Life Research*. Published online DOI 10.1007/s11136-006-9136-6.

Nelson, K. B., & Ellenberg, J. H. (1982). Children who 'outgrew' cerebral palsy. *Pediatrics, 69*(5), 529–536.

Palisano, R., Rosenbaum, P., Walter, S., Russell, D., Wood, E., & Galuppi, B. (1997). Development and reliability of a system to classify gross motor function in children with cerebral palsy. *Developmental Medicine and Child Neurology, 39*, 214–223.

Parkes, J., Kerr, C., McDowell, B., & Cosgrove, A. (2006). Recruitment bias in a population-based study of children with cerebral palsy. *Pediatrics, 118*, 1616–1622.

Pirpiris, M., & Graham, H. K. (2004). Uptime in children with cerebral palsy. *Journal of Pediatric Orthopaedics, 24*(5), 521–528.

Raat, H., Botterweck, A. M., Landgraf, J. M., Hoogeveen, W. C., & Essink-Bot, M. (2005). Reliability and validity of the short form of the child health questionnaire for parents (CHQ-PF28) in large random school based and general population samples. *Journal of Epidemiology and Community Health, 59*(1), 75–82.

Revivo, G. (2000). *The use of Tizanidine in a pediatric population: Clinical outcome*. Free paper presented at the American Academy of Cerebral Palsy and Developmental Medicine, Toronto, Canada.

Ruperto, N., Ravelli, A., Pistorio, A., Malattia, C., Cavoto, S., Gado-West, L., et al. (2001). Cross-cultural adaptation and psychometric evaluation of the Childhood Health Assessment Questionnaire (CHAQ) and the Child Health Questionnaire (CHQ) in 32 countries. Review of the general methodology. *Clinical and Experimental Rheumatology, 4*(Suppl 23), S1–S9.

Russell, D. J., Rosenbaum, P. L., Gowland, C., Cadman, D. T., Gowland, C., Hardy, S., et al. (1989). The gross motor function measure: A means to evaluate the effects of physical therapy. *Developmental Medicine and Child Neurology, 31*, 341–352.

Samson-Fang, L., Fung, E., Stallings, V. A., Conaway, M., Worley, G., Rosenbaum, P., et al. (2002). Relationship of nutritional status to health and

societal participation in children with cerebral palsy. *Journal of Pediatrics, 141*(5), 637–643.

Schneider, J. W., Gurucharri, L. M., Gutierrez, A. L., & Gaebler-Spira, D. H. (2001). Health-related quality of life and functional outcome measures for children with cerebral palsy. *Developmental Medicine and Child Neurology, 43*(9), 601–608.

Stanley, F., Blair, E., & Alberman, E. (2000). *Cerebral palsies: Epidemiology and causal pathways. (Clinics in developmental medicine No. 151)*, London: Mac Keith Press.

Surveillance of Cerebral Palsy in Europe Collaborative Group (SCPE). (2000). Why a European collaboration of cerebral palsy surveys and registers? *Developmental Medicine and Child Neurology, 42*(12), 816–824.

Vargus-Adams, J. (2005). Health-related quality of life in childhood cerebral palsy. *Archives of Physical Medicine and Rehabilitation, 86*(5), 940–945.

Vargus-Adams, J. (2006). Longitudinal use of the Child Health Questionnaire in childhood cerebral palsy. *Developmental Medicine and Child Neurology, 48*, 343–347.

Vitale, M. G., Roye, E. A., Choe, J. C., Hyman, J. E., Lee, F. Y., & Roye, D. P. (2005). Assessment of health status in patients with cerebral palsy: What is the role of quality of life measures. *Journal of Pediatric Orthopaedics, 25*(6), 792–797.

Wake, M., Salmon, L., & Reddihough, D. (2003). Health status of Australian children with mild to severe cerebral palsy: Cross-sectional survey using the Child Health Questionnaire. *Developmental Medicine and Child Neurology, 45*(3), 194–199.

Wallen, M. A., O'Flaherty, S. J., & Waugh, M. C. A. (2004). Functional outcomes of intramuscular boulinum toxin type A in the upper limbs of children with cerebral palsy. *Archives of Physical Medicine and Rehabilitation, 85*(2), 192–200.

Waters, E., Salmon, L., & Wake, M. (2000). The parent-form Child Health Questionnaire in Australia: Comparison of reliability, validity, structure, and norms. *Journal of Pediatric Psychology, 25*(6), 381–391.

World Health Organisation (1948) Preamble to the Constitution. International Health Conference, New York, 19–22 June, World Health Organisation; 1946.

World Health Organisation (2001) *International Classification of Functioning, Disability and Health*. Geneva: World Health Organisation.