Reliability and Validity of the Child Health Questionnaire$^{PF\text{-}50}$ for European Children with Cerebral Palsy

Nichola McCullough,[1] PhD, Jackie Parkes,[1] PhD, BNurs, Melanie White-Koning,[2] PhD, Eva Beckung,[3] PT, PhD, and Allan Colver,[4] MD, FRCPH

[1]School of Nursing & Midwifery, Queen's University Belfast, [2]Institut National de la Santé et de la Recherche Médicale, Université Paul Sabatier, Faculté de Médecine, [3]Göteborg University, The Queen Silvia Children's Hospital, and [4]Sir James Spence Institute, Newcastle University, Royal Victoria Infirmary

Objective To evaluate the psychometric performance of the Child Health Questionnaire (CHQ) in children with cerebral palsy (CP). **Method** 818 parents of children with CP, aged 8–12 from nine regions of Europe completed the CHQ (parent form 50 items). Functional abilities were classified using the five-level Gross Motor Function Classification Scheme (Levels I–III as ambulant; Level IV–V as nonambulant CP). **Results** Ceiling effects were observed for a number of subscales and summary scores across all Gross Motor Function Classification System levels, whilst floor effects occurred only in the physical functioning scale (Level V CP). Reliability was satisfactory overall. Confirmatory factor analysis (CFA) revealed a seven-factor structure for the total sample of children with CP but with different factor structures for ambulant and nonambulant children. **Conclusion** The CHQ has limited applicability in children with CP, although with judicious use of certain domains for ambulant and nonambulant children can provide useful and comparable data about child health status for descriptive purposes.

Key words cerebral palsy; child health questionnaire; confirmatory factor analysis; exploratory factor analysis; reliability.

Background

Cerebral palsy (CP) is a term that refers to a collection of motor impairments resulting from damage to, or abnormal development of the immature brain. People with CP experience a range of movement, posture, and coordination disorders which are life-long, vary by type, severity, and etiology with milder forms being more common. Many people with CP also experience a number of associated impairments, affecting intelligence, the special senses, as well as other medical problems. Approximately two children in every 1,000 are affected by the condition in the developed world, but as opportunities for prevention remain limited, CP is likely to remain a leading cause of motor impairment in children for the foreseeable future.

Reliable and valid instruments for assessing the impact of CP on children and their families can provide important information for the planning and assessment of treatment programmes. For example, DISABKIDS (Simeoni et al., 2007) was developed to measure quality of life and well-being in children who have chronic illnesses. However, the use of generic instruments such as the Health Utilities Index (HUI; Furlong, Feeny, Torrance, & Barr, 2001), the Pediatric Quality of Life Inventory (PedsQL; Varni, Seid, & Rode, 1999) and the Child Health Questionnaire (CHQ; Landgraf, Abetz, & Ware, 1996) are advantageous in that they allow for a comparison of health status to be made with other populations of children, both healthy and chronically ill. Whilst the HUI and the pedsQL have been utilized to assess health and well-being in the CP population (Majnemer, Shevell, Rosenbaum, Law, & Poulin, 2007), this article will systematically evaluate the psychometric performance of the commonly used CHQ parent form (Landgraf et al., 1996). Verification of the psychometric properties of the instrument will have implications for child health research and clinical practice, particularly in relation to children with disabilities as we

All correspondence concerning this article should be addressed to Nichola McCullough, School of Nursing & Midwifery, Queen's University Belfast, 10 Malone Road, Belfast, BT9 5BN, UK.
E-mail: nichola.mccullough@qub.ac.uk

Journal of Pediatric Psychology 34(1) pp. 41–50, 2009
doi:10.1093/jpepsy/jsn048
Advance Access publication May 22, 2008
Journal of Pediatric Psychology vol. 34 no. 1 © The Author 2008. Published by Oxford University Press on behalf of the Society of Pediatric Psychology.
All rights reserved. For permissions, please e-mail: journals.permissions@oxfordjournals.org

have evaluated the instrument in a large sample of European children with CP.

The CHQ (Landgraf et al., 1996) assesses physical and psychosocial health and well-being, with three versions validated for 5- to 18-year-olds. A recent review (McCullough & Parkes, 2007) identified 13 studies published between January 1993 and January 2007, which had used the parent form CHQ (28 or 50 item) in children with CP. A further five papers have since been published (Aran, Shalev, Biran, & Gross-Tsur, 2007; Bjornson, Belza, Kartin, Logsdon, & McLaughlin, 2008; Kerr, Parkes, Stevenson, Cosgrove, & McDowell, 2008; Majnemer et al., 2007; Wren et al., 2007).

Whilst some authors have considered the CHQ to be suitable for use in this population (Morales et al., 2006; Wake, Salmon, & Reddihough, 2003), McCullough and Parkes (2007) identified psychometric and methodological shortcomings needing further investigation. For example, the reliability of the CHQ parent form 50 items (CHQ PF-50) in children with CP has not been examined across the range of gross motor ability. Floor and ceiling effects have been reported for a number of scales (Schneider, Gurucharri, Guitierrez, & Gaebler-Spira, 2001; Wake et al., 2003). Schneider and colleagues (2001) also found that some parents reported certain items to be inappropriate for their children given the severity of their condition.

There are few reported findings on the construct validity as demonstrated by exploratory or confirmatory factor analysis (CFA) (Drotar, Schwartz, Palermo, & Burant, 2006). Yet, an examination of the measurement structure of a generic questionnaire in children with chronic conditions like CP is important, both theoretically and practically. Furthermore, specific characteristics of the condition, for example gross motor function, may affect the applicability of the questionnaire or influence caregivers' interpretation. Consequently, one measurement model may not be applicable across the broad spectrum of abilities experienced by children with CP.

Most studies using the CHQ have included small samples of children with CP, recruited from clinics which are likely to result in systematic biases, limiting the extent to which findings can be generalized. Furthermore, the majority of findings to date have been based on North American (Bjornsen et al., 2008; Fung et al. 2002; Houlihan, O'Donnell, Conaway, & Stevenson, 2004; Liptak et al., 2001; Pipiris & Graham, 2004; Majnemer et al., 2007; Samson-Fang et al., 2002; Schneider et al., 2001; Vargus-Adams, 2005, 2006; Vitale et al., 2005; Wallen, O'Flaherty, & Waugh, 2004), Australian (Wake et al., 2003), Brazilian (Morales et al., 2006), and Israeli (Aran et al., 2007) CP samples, limiting its applicability to European children. It is important to establish the properties of a tool designed to assess subjective health and well-being in different cultures, as conceptualizations may vary.

The aims of this article are to evaluate the data quality, reliability (scale internal consistency), and factor structure of the CHQ (parent form 50 items; PF50) in a representative sample of children with CP living in Europe, with a particular focus on how its performance varies by gross motor function. The data were available from the SPARCLE Study (Colver, 2006), the aim of which is to establish the influence of environmental factors (social, attitudinal, and physical) on participation and quality of life in 8- to 12-year-old children with CP. A number of important, potential confounding factors had to be taken into consideration, including child health status. The performance of the instrument measuring child health status (the CHQ-PF50) is the subject of this report.

Method

Study Design

The SPARCLE study was a cross-sectional survey of 8- to 12-year-old children with CP and their families, in nine regions from seven countries of the European Union.

Participants

The children were identified from eight preexisting, population-based registers of children with CP, covering regions in north England, west Sweden, Northern Ireland, southeast France, southwest Ireland, east Denmark, central Italy, and southwest France. A further center in northwest Germany followed study protocols, but its sample was constructed with referrals from multiple sources in a defined geographic area.

The details of recruitment, sampling procedures, and inclusion criteria are described elsewhere (Dickinson et al., 2006). In summary, children with a diagnosis of CP, born July 31, 1991 to April 1, 1997, and who were resident in one of the geographical areas, were eligible to take part. The sampling strategy involved grouping children recorded on the registers into four categories of walking ability, with the ideal of recruiting 30 subjects per group with an overall target of 120 children and families per region. Where sufficient numbers were recorded some registry centers were able to select subjects randomly from all or some severity strata. This was undertaken to ensure that children with all severities of CP were equally represented in the SPARCLE study. A total of 1,174 children were identified

as potentially eligible; of these 993 (85%) were traced and approached, and 818 (70%) families participated. These families were then visited in their own homes by trained research assistants (RAs). The RAs administered the questionnaires in accordance with the authors' standardized instructions.

Definitions and Classifications

All participating regions used the same definition of CP in the compilation of their respective registers (SCPE, 2000). The definition used included the following key elements: an umbrella term used to describe a group of motor disorders; a permanent, but not unchanging, disorder in terms of clinical presentation; involving a disorder of movement and/or posture of motor function; due to a nonprogressive interference and/or lesion and/or abnormality of the developing brain (SCPE, 2000). All regions also used the SCPE definitions and decision tree to classify children by clinical subtypes in a standardized way (SCPE, 2000). The RAs further classified each child by their gross motor function using the Gross Motor Function Classification System (GMFCS), an age-specific scheme designed for children with CP based on five levels of ability ranging from Level I (most able) to Level V (least able) (Palisano et al., 1997). The psychometric properties of the GMFCS have been thoroughly tested and reported, and include evidence supporting its content validity (Palisano et al., 1997); construct validity (Palisano et al., 2000); concurrent validity (Beckung & Hagberg, 2000); inter-rater reliability (Palisano et al., 1997); and test–retest reliability (Wood & Rosenbaum, 2000). In this article, children in GMFCS levels I–III are considered to have ''ambulant CP'' and children in GMFCS levels IV and V to have ''nonambulant CP''.

Instruments

The CHQ

The CHQ-PF50 (Landgraf et al., 1996) has 13 single and multi-item scales that assess child health status over ''the last four weeks,'' and a further global item assessing change in health ''over the last year.'' The CHQ is designed to assess both physical and psychosocial well-being. Scales in the physical domain include physical functioning, role/physical–social limitations, general health perceptions, and bodily pain. Scales in the psychosocial domain include role/social–emotional/behavioral, self-esteem, mental health, general behavior, parental impact–emotion, parental impact–time, and the family activities scale. Also included is a single item that assesses family cohesion. Responses are scored for each domain, producing a figure between 0 and 100, with higher scores indicating better health and well-being. The scales in the PF50 generate two summary scores representing 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). The CHQ-PF50 is available in 60 languages, having been translated by an independent company for the authors of the CHQ. This was done using rigorous forward–backward translations. Multi-trait item scaling analysis was carried out posttranslation to affirm that appropriate translations had been made. Normative data are available for children in Australia (Waters, Salmon, & Wake, 2000) and US (Landgraf et al., 1996). The children in this study were administered with the appropriate version in accordance with their country of residence.

Ethics Approval

Ethics approval to conduct the study was obtained in each of the countries and complied with the local requirements (Colver, 2006).

Analysis

Data quality (missing item response, floor, and ceiling effects) was examined for both individual items and subscales by GMFCS level and for the sample overall. Floor and ceiling effects were defined as the percentage scoring the highest and lowest absolute values on the domain scores only (Langraf et al., 1996). This is inappropriate for summary scores (Langraf et al., 1996); therefore, percentile values were used to describe the extent of children with below and above average (25th and 75th percentile, respectively) physical and psychosocial summary scores.

Scale internal consistency was evaluated for all multi-item subscales of the CHQ by each level of the GMFCS, and for the total sample using Cronbachs α coefficient, with an α-value of .70 or higher defined as an acceptable level (Eiser & Morse, 2001).

An exploratory factor analysis (EFA) was conducted to identify a measurement model of the CHQ-PF50 in children with CP. A total of 48 items, representing 11 scales of the CHQ-PF50, were entered into a principle axis factor analysis using varimax rotation. Orthogonal rotation was chosen as in the original EFA by Landgraf and colleagues (1996). Factor analyses were run for models with 6–13 factors, in order to determine if the 11 factor model as hypothesized by Landgraf et al. (1996) had the cleanest factor structure. The item loadings were then examined to identify problem items that were common across each of these eight models. Items with primary factor loadings $< .40$, and secondary factor loadings $> .30$ were removed one at a time, with the factor analysis being

rerun for 6–13 factors after each item removal. This procedure was carried out until a clean solution with primary loadings ≥ .40 and secondary loadings ≤ .30 was found.

The factors identified by the EFA were then tested using CFA, firstly testing the stability of the final model across the total sample of children using multigroup nested models. This was used to test whether the factor structure was invariant across the ambulant and nonambulant groups. In this procedure, the initial CFA model (M0) was run without constraining any of the parameters to be equal. We then fitted a model (M1) constraining factor loadings to be equal across the two groups (ambulant and nonambulant) and tested for differences in chi-square between M0 and M1. As there was measurement variance between ambulant and nonambulant children, this suggested the factor structure found for the pooled group was not valid in each of the subgroups, and so separate EFAS for ambulant and nonambulant children were undertaken. CFA with nested models testing was again performed, this time on each separate model to assess invariance between the ambulant and nonambulant subgroups. Should an alternative model fit both the ambulant and nonambulant subgroups (variance), this model could be used to assess health status in both groups collectively. However, the finding of a unique fit (invariance) will indicate that separate models are required to assess health status of ambulant and nonambulant subgroups. All models were evaluated for goodness-of-fit using root mean square error of approximation (RMSEA), Tucker Lewis Index (TLI) and comparative fit index (CFI) indices. A TLI and CFI score of .90 was considered acceptable, but a score of > .95 was considered an excellent fit (Bentler & Chou, 1987); and RMSEA score of < .08 was considered acceptable, with < .05 considered to be good (Browne & Cudeck, 1993). All analyses were conducted using SPSS/PC Version 15, with the exception of the CFAs (single and multigroup testing), which were carried out using a structural equation modeling program (MPlus, Muthen & Muthen, 2000).

Results

Child and family characteristics of the SPARCLE sample are shown in Table I. Of the 818 children and families who took part, 808 had CHQ data and are the subject of the following results.

Data Quality

Forty items on the CHQ had < 5% of missing responses, and ten items had missing responses that ranged from 5% to 10%. Table II shows a summary of data quality at scale level for the total sample, although analysis was also carried out by GMFCS levels (not shown). In general, the percentage of missing data was small, the largest percentage being for the two summary scores—largely due to the scoring procedure, whereby if one or more subscale scores are missing then summary scores cannot

Table I. Demographic Data for Children with Cerebral Palsy and their Families (*n*=818)

Characteristics		*n*	%
Centres	North England	116	14
	West Sweden	83	10
	Northern Ireland	102	13
	South East France	67	8
	South West Ireland	98	12
	East Denmark	115	14
	Central Italy	85	10
	South West France	77	9
	North West Germany	75	9
	Missing	0	0
Gender	Boys	484	59
	Girls	334	41
	Missing	0	0
Age (years)	7/8	184	23
	9	158	19
	10	166	20
	11	159	19
	12/13	151	19
	Missing	0	0
GMFCS	I (no limitations)	257	31
	II	164	20
	III	139	17
	IV	113	14
	V (total assistance)	145	18
	Missing	0	0
Intellectual Impairment	None-mild (IQ > 70)	385	47
	Moderate (IQ 50-70)	186	23
	Severe (IQ < 50)	242	30
	Missing	5	0.6
Family Structure	Married, living with partner	576	71
	Living with partner	84	10
	Single/separated, living with parents	18	2
	Single alone	138	17
	Missing	1	0.1
Parent Occupation	Full-time professional	226	28
	Full-time trade trade	402	15
	Part-time	48	35
	Neither partner working	141	24
	Missing	1	0.1

Table II. Data Quality for Sample Overall ($n = 808$) and Internal Consistency (Cronbach's α) by GMFCS Levels and Overall

	Items	Data Quality (%)			Scale reliability (by GMFCS levels)					
CHQ domains	(*n*)	Missing	Floor effects	Ceiling effects	I	II	III	IV	V	Total sample
Behavior	6	2.2	0.1	1.8	.747	.746	.663	.623	.324	.681
Bodily pain	2	0.5	1.4	28.3	.862	.891	.870	.831	.916	.872
General health	6	0.1	1.0	2.0	.606	.653	.519	.646	.606	.645
Mental health	5	1.2	0.4	3.9	.777	.635	.705	.761	.697	.727
Parent impact—emotional	3	0.4	1.6	13.4	.629	.722	.708	.759	.713	.685
Physical functioning	6	4.0	6.7	43.0	.840	.916	.947	.974	.980	.940
Parent impact-time	3	0.9	3.4	39.1	.841	.833	.795	.740	.773	.796
Role—emotional/behavioral	3	3.1	3.8	58.7	.905	.952	.945	.954	.956	.941
Role physical	2	2.3	8.1	60.2	.884	.929	.932	.923	.973	.935
Self-esteem	6	3.3	0.1	8.6	.758	.809	.703	.840	.845	.788
Family activities	6	0.6	0.5	19.8	.890	.885	.872	.893	.893	.888
Family cohesion	1	1.0	0.6	16.1	_	_	_	_	_	_
Global health	1	0.6	2.6	23.6	_	_	_	_	_	_
Global behavior	1	2.5	4.3	16.5	_	_	_	_	_	_
Physical summary score	_	9.0	32.76	54.02	_	_	_	_	_	_
Psychosocial sum. score	_	9.0	43.05	55.64	_	_	_	_	_	_

be reliably calculated (Landgraf et al., 1996). The proportion of missing data for the summary scores increased by GMFCS level and was lowest for children in Level I and highest for children in Level V (χ^2 for trend, $p \leq .001$).

For the total sample there was little evidence of floor effects. However, by GMFCS levels, floor effects were observed for children in Level V, with 27% and 22% of children scoring the lowest possible score in the ''*physical functioning*'' and ''*role-physical*'' scales, respectively. Ceiling effects were present in a number of scales for the total sample (Table II). A consistently high proportion of the study sample exhibited floor and ceiling effects for the summary scales, not only evident among the total sample but also by GMFCS levels. For the physical summary score, the proportion of children exhibiting floor effects decreased as GMFCS levels increased; there was no evidence of a similar trend for ceiling effects. On the psychosocial summary score, remarkably similar proportions of children exhibited floor and ceiling effects (around 40 and 55%, respectively) for the overall sample and by GMFCS levels.

Reliability

Scale internal consistency

Table II shows the scale internal consistency of the CHQ-PF50 for the total sample and by GMFCS levels. For the total sample, three scales had an α-value below the .70 threshold. In relation to ''*behavior*,'' internal consistency declined by GMFCS levels, being adequate for children in Levels I and II, but decreasing to 0.32 for children in Level V. Five scales had α-values .80 or higher. These scales were relatively stable across all levels of the GMFCS.

Construct validity

Exploratory Factor Analysis for total CP sample

The exploratory factor analysis, based on the total sample, revealed a 32-item, seven-factor solution. Original factors that remained included ''*physical functioning*;'' ''*role emotional behavior*;'' ''*bodily pain*;'' ''*behavior*;'' ''*self-esteem*;'' ''*general health*;'' and ''*family activities*''. Whilst the items that loaded onto these factors were consistent with Landgraf et al. (1996), certain factors gained additional items whilst others lost items. The physical functioning scale gained an additional item ''*limited in the kind of activity*'' from the original ''*role social physical*'' scale. The ''*family activities*'' factor also included a new item that originated in the ''*parental impact time*'' scale ''*your child's emotional well-being or behavior*.'' The ''*family activities*'' factor also lost two items ''*caused tension and conflict*'' and ''*source of disagreements or arguments*.'' The ''*behavior*'' factor lost two original items ''*concentrate*'' and ''*stole*'' and the ''*general health*'' factor lost one original item ''*never seriously ill*.'' Factors that failed to emerge included: ''*role physical*,'' ''*mental health*,'' ''*parental-impact emotion*,'' and ''*parental-impact time*.''

CFA and Subgroup Comparisons

CFA showed that the initial model identified in the EFA was an excellent fit across the total sample ($\chi^2 = 705.024$, df = 121, $p < .0001$; CFI = 0.966, TLI = 0.986, RMSEA = 0.077), and confirmed a seven-factor structure. Fitting

Table III. Final Confirmatory Factor Analysis Model on AMBULANT Children with CP

Domains	Items		Estimate	*SE*	Two-tailed *p*-value
Physical functioning	PFWALK	CHQ2.1D	0.951	0.013	<0.001
	PFBEND	CHQ2.1E	0.906	0.017	<0.001
	PFNEIGH	CHQ2.1C	0.894	0.018	<0.001
	PFCARESELF	CHQ2.1F	0.852	0.022	<0.001
	PFSOME	CHQ2.1B	0.810	0.024	<0.001
	RPKIND	CHQ3.2A	0.791	0.030	<0.001
Family activities	FADROP	CHQ9.3A	0.904	0.015	<0.001
	FALIMIT	CHQ9.3C	0.873	0.016	<0.001
	FAINTERUP	CHQ9.3B	0.824	0.021	<0.001
	FACANCEL	CHQ9.3F	0.800	0.022	<0.001
	PTEMOTBEH	CHQ9.2B	0.780	0.030	<0.001
Self esteem	SEOVERALL	CHQ7.1F	0.865	0.027	<0.001
	SEFAMILY	CHQ7.1E	0.718	0.032	<0.001
	SEFRIEND	CHQ7.1C	0.711	0.035	<0.001
	SELOOKS	CHQ7.1D	0.649	0.032	<0.001
	SESCHOOL	CHQ7.1A	0.622	0.039	<0.001
	SEATHLETIC	CHQ7.1B	0.562	0.040	<0.001
General health	GHLESSHLTH	CHQ8.1A	0.830	0.033	<0.001
	GHGLOBAL	CHQ1.1	0.733	0.032	<0.001
	GHEXPECT	CHQ8.1D	0.683	0.038	<0.001
	GHWORRIES	CHQ8.1E	0.573	0.040	<0.001
	GHCATCH	CHQ8.1C	0.484	0.042	<0.001
Role emotion behavior	REBPERFORM	CHQ3.1C	0.977	0.008	<0.001
	REBAMOUNT	CHQ3.1B	0.947	0.009	<0.001
	REBKIND	CHQ3.1A	0.935	0.010	<0.001
Behavior	BETANTRUMS	CHQ5.1E	0.879	0.055	<0.001
	BEARGUE	CHQ5.1A	0.652	0.048	<0.001
Bodily pain	BPGLOBAL	CHQ4.2	0.933	0.031	<0.001
	BPFREQ	CHQ4.1	0.873	0.031	<0.001

χ^2-value = 316.984; df = 108, $p < 0.0001$; CFI = 0.970; TLI = 0.987; RMSEA = 0.059.

Table IV. Final Confirmatory Factor Analysis Model on NON AMBULANT Children with CP

Domains	Items		Estimate	*SE*	Two-tailed *p*-value
Physical functioning	PFWALK	CHQ2.1D	0.965	0.006	<0.001
	PFBEND	CHQ2.1E	0.959	0.006	<0.001
	PFALOT	CHQ2.1A	0.956	0.006	<0.001
	PFSOME	CHQ2.1B	0.947	0.008	<0.001
	PFNEIGH	CHQ2.1C	0.930	0.009	<0.001
	PFCARESELF	CHQ2.1F	0.921	0.010	<0.001
	RPKIND	CHQ3.2A	0.795	0.020	<0.001
Family activities	FALIMIT	CHQ9.3A	0.918	0.012	<0.001
	FADROP	CHQ9.3C	0.893	0.012	<0.001
	FAINTERUP	CHQ9.3B	0.848	0.015	<0.001
	FACANCEL	CHQ9.3F	0.814	0.017	<0.001
Self esteem	SEOVERALL	CHQ7.1F	0.854	0.020	<0.001
	SEFAMILY	CHQ7.1E	0.735	0.026	<0.001
	SEFRIEND	CHQ7.1C	0.706	0.027	<0.001
	SEAPPEAR	CHQ7.1D	0.683	0.025	<0.001
	SESCHOOL	CHQ7.1A	0.638	0.030	<0.001
	SEATHLETIC	CHQ7.1B	0.561	0.033	<0.001
	MHCHEER	CHQ6.1E	0.558	0.035	<0.001
General health	GHLESSHLTH	CHQ8.1A	0.847	0.030	<0.001
	GHEXPECT	CHQ8.1D	0.760	0.029	<0.001
	GHCATCH	CHQ8.1C	0.497	0.036	<0.001
Role emotion behavior	REBPERFORM	CHQ3.1C	0.975	0.005	<0.001
	REBAMOUNT	CHQ3.1B	0.958	0.006	<0.001
	REBKIND	CHQ3.1A	0.948	0.007	<0.001
Parental impact-time	PTEMBEHAV	CHQ9.2B	0.984	0.022	<0.001
	PTATTENT	CHQ9.2C	0.799	0.024	<0.001
Bodily pain	BPFREQ	CHQ4.1	0.919	0.022	<0.001
	BPGLOBAL	CHQ4.2	0.910	0.021	<0.001

χ^2-value = 431.463, df = 95; $p < 0.0001$; CFI = 0.982; TLI = 0.992; RMSEA = 0.066.

this initial model (M0) across the total sample without constraining any of the parameters to be equal was undertaken ($\chi^2 = 647.288$, df =201, $p < .0001$; CFI = 0.979, TLI = 0.990, RMSEA = 0.074) followed by a nested other model (M1), but this time constraining factor loadings to be equal. This revealed that the model was not the same across ambulant and nonambulant groups (χ^2-test for difference = 52.812, df = 17, $p < .0001$). Separate EFAs for both the ambulant and nonambulant children showed that both groups did indeed have different factor structures, and subsequent CFAs confirmed the separate factor structures for children in the two groups (Tables III and IV). Both final CFA models showed an excellent fit as indicated by the TLI and CFI scores and an acceptable fit based on RMSEA indices (ambulant $\chi^2 = 316.984$ df = 108, $p < .0001$, CFI = 0.970, TLI = 0.987, RMSEA = 0.059; nonambulant $\chi^2 = 431.463$, df = 95, $p < .0001$, CFI = 0.982, TLI = 0.992, RMSEA = 0.066). Six factors were consistently identified across both groups, with the additional factors ''*behavior*'' emerging uniquely among ambulant children and ''*parent-impact time*'' among nonambulant children.

Finally, to determine whether the final model found for the ambulant group might fit the nonambulant group and whether the final model found for the nonambulant group might fit the ambulant group, nested models were used to test for measurement invariance. First, the ambulant model (MA0 given in Table IV) was fit across the total sample unconstrained. The χ^2 difference test between the model constraining factor loadings to be equal across groups and MA0 was statistically significant (χ^2-test for difference = 74.254, df = 17, $p < .0001$) demonstrating measurement variance across groups. A similar conclusion was reached concerning the nonambulant model MNA0

(Table IV), with a statistically significant χ^2 difference test between the unconstrained and the constrained models (χ^2-test for difference = 45.805, df = 15, $p < .0001$). Hence, neither the ambulant nor the nonambulant model can be used across both groups.

Discussion

This is the first report of the psychometric performance of the CHQ-PF50 in a population-based sample of children with CP in Europe. It is also the largest study of the CHQ-PF50 in children with CP reported in the literature to date. The extent of missing data reported here was low overall, suggesting that the instrument was acceptable to parents. However, when stratified by GMFCS levels, significant numbers of scores were missing from scales in the physical domain for children in GMFCS level V. This is not an unexpected finding, as children with severe CP clearly have reduced functional abilities frequently associated with intellectual and communication impairments.

The large number of ceiling effects and to a lesser extent floor effects found here is consistent with previous reports (McCarthy et al., 2002; Morales et al., 2006; Vitale et al., 2005). When examined by GMFCS levels, floor effects for the ''*physical functioning*'' and ''*role/social-physical*'' scale were present in children with Level V CP. Paradoxically, these scales also had extreme ceiling effects in both nonambulant and ambulant children. Given that the instrument was designed for use in clinical outcome studies, the presence of so many ceiling effects may not be unusual. The heterogeneous nature of type and severity of impairment in children with CP (Morales et al., 2006), and the willingness of parents to report that their children are healthy may also have been a factor. The existence of floor and ceiling effects does however call into question the utility of certain scales of the CHQ for the population of children with CP because they were unable to detect variability. It is desirable that scores on instruments are both variable and evenly distributed in order to identify evidence of effects, for example, when assessing the efficacy of an intervention. Only one previous study has utilized the CHQ as an outcome measure in an intervention trial (Wallen et al., 2004), with no significant difference being reported between baseline and three and 6-month follow-up scores on the CHQ. These findings in relation to data quality suggest poor sensitivity and/or clinical relevance of certain items and domains of the CHQ-PF50 for children with severe forms of CP.

The reliability of the CHQ-PF50 was found to be satisfactory for most scales, and was comparable to that reported by other studies of children with CP in America, Australia, and Brazil. However, three scales had reliability coefficients of < 0.7. Others have also reported lower reliability coefficients for the same scales in CP samples (Morales et al., 2006; Wake et al., 2003). When examined by GMFCS levels there was some variability for the ''*behavior*'' scale, with the lowest coefficients being found for GMFCS levels III–V, the latter being the lowest. The ''*general health perceptions*'' scale was consistently lower across all subgroups. Both Landgraf et al. (1996) and Waters et al. (2000) reported that the ''*general health perceptions*'' scale had lower reliability coefficients in normative populations. One possible explanation is that this scale was purposefully designed to be both subjective and heterogeneous in order to avoid response acquiesce (Landgraf et al., 1996). However, the key message remains that there are inherent problems with the ''*general health perceptions*'' scale in both normative and CP samples. This suggests that it may need revision; alternatively, it could be said that an α-value $> .5$ is acceptable when the instrument is used for descriptive purposes rather than as an outcome measure to detect differences between groups, where the preferred minimum is .7 (Eiser & Morse, 2001).

The factor analysis of the CHQ-PF50 in the overall the sample of children with CP showed that a number of factors, predominantly those that represented the child's psychosocial health status, were consistent with Landgraf et al.'s (1996) model. This indicates that certain scales of the CHQ-PF50 can be utilized by researchers and clinicians to make reliable comparisons of the health status of children with CP and the general population. However, the identification of fewer factors than those originally hypothesized by Landgraf et al. (1996) in this study is similar to the findings of Drotar et al. (2006). As with Drotar et al., we found that factors assessing the burden of child health status on parental time and well-being were not present (based on the total sample of CP children). However, whilst Drotar and colleagues identified mental health as a primary factor, our study did not. One reason for this may be that many of the items in this domain are not relevant to children with CP. Also, Drotar et al. studied children with chronic illnesses, which were somewhat more diverse, a fact that may have ensured that a greater number of items were responded to.

Further variation in the factor structure of the CHQ was also observed between ambulant and nonambulant children. The models for each of these two subgroups suggested that not all domains were relevant across the spectrum of CP. Within the nonambulant group the factor

for ''*parental-impact time*'' was present, whilst it was missing for the ambulant children, with the item ''*emotional well-being or behaviour*'' loading onto the ''*family activities*'' scale. Given the nature of CP, especially for those children at the more severe end of the spectrum, it is not unexpected that the factor of parental impact time appears only in the nonambulant group. These children have limited functional abilities and some will require total care. In addition, their condition may be made more complex by the presence of associated and sometimes severe impairments hampering their independence. Parents/carers will have to invest significantly more time in caring for these children, limiting the time that they have to invest in their own needs. In contrast, the factor ''*behavior*'' did not emerge in the nonambulant group, whereas it did for the ambulant children. Interestingly, this factor also exhibited lower reliability coefficients in the nonambulant group. As before, the poor performance of this factor in nonambulant children may be due to the interaction between the items and the functional capacity of children with severe CP. Items in this domain largely represent externalising behaviors (e.g., stealing, cheating, lying) that require children to have a certain level of physical, intellectual, and communication ability. Children with more severe CP are likely to be more restricted in these areas and therefore these items are not as relevant. This may also explain differences in the item loadings for other factors between the two groups. For example, in the ''*general health*'' factor the ambulant group had five items, whereas among nonambulant children it only had three.

Taken together, this study has revealed a number of important findings that have both clinical and scientific relevance for those intent on utilizing the CHQ-PF50 as a tool with which to research or assess children's health status. Primarily, this study found a model with fewer items than hypothesised by Landgraf et al. (1996). This may be viewed as beneficial to researchers, clinicians, and respondents in that fewer items are required to obtain a measure of well-being, a point previously highlighted by Drotar et al. (2006). However, the fact that fewer scales have been identified in this population does suggest that the use of summary scores as a reliable indication of physical and psychosocial health and well-being is not suitable in populations of children with CP. Furthermore, as this study found different factor structure models for both ambulant and nonambulant children, the assumption of equivalence across the spectrum of children with CP cannot be made for certain scales of the CHQ. Rather than administering all scales of the CHQ-PF50 universally to all children with CP, consideration of both the needs of the researcher and the characteristics of the children should direct more judicious use of the CHQ. The CHQ-PF50 can provide a useful descriptive, ''snapshot'' of the health and well-being of children with CP that can be compared to the profiles of other children, but can only do so reliably in relation to specific domains with some caveats: ''*physical functioning*'' (although a CP specific measure may be preferable), ''*family activities*'', ''*self-esteem*'', ''*general health*'' (although we found lower reliability coefficients suggesting an inherent problem with this scale), ''*role emotional behavior*'' and ''*bodily pain*.''

It is important that the findings of this study are interpreted in the context of certain conceptual and methodological limitations. Firstly, this study included parent report alone. Child self-report (where possible) may have produced different findings. Future research should explore the extent to which the factor structure of the CHQ may vary when reported by either parents or children. Whilst this may be difficult to achieve across all levels of severity in a CP population, it could be investigated in subgroups of individuals less severely affected. Second, we do not know the extent to which some parents may have responded to the CHQ interpreting questions about their child's ''health'' to mean the same as their child's ''disability,'' whereas others may have perceived and reported on these concepts separately leading to lower response rates, inconsistencies in reporting, and ultimately factors failing to emerge in the final analysis. Finally, in terms of our response rates, 37% of families traced did not take part in the study (a total of 24% actually refused), although these rates are comparable to other, similar surveys (Dickinson et al., 2006). There was significant heterogeneity between regions in terms of response rates but we were unable to record other individual or societal factors associated with refusal to take part in the research. These families are likely to have been systematically different in some way, although not necessarily in terms of their child's health status. However, families were originally sampled from population-based registers (as opposed to health clinics/outpatient departments), and this will have reduced selection bias in the first instance (Dickinson et al., 2006). Few surveys in child health research have been able to identify potential participants in the same systematic way due to lack of geographically defined sampling frames.

To reiterate Drotar et al.'s (2006) recommendations, future research should continue to examine the factor structure of the CHQ in other population-based samples of children with specific chronic illnesses, and whether factor

structures vary as a function of severity of illness. This study also highlights the need to establish new and specifically designed instruments to monitor the health of children with severe CP, especially among children unable to self-report. Such instruments could then be used in *conjunction* with specific scales of the CHQ, providing both generic- and disease-specific reliable information on the health and well-being of children with CP.

Acknowledgments

The study was funded by the European Union Research Framework 5 Programme—Grant number QLG5-CT-2002-00636. The German region joined later, funded by Bundesministerium für Gesundheit/German Ministry of Health (GRR-58640-2/14) and Stiftung für das Behinderte Kind/Foundation for the Disabled Child.

Conflicts of interest: Dr Melanie White-Koning was privately engaged as a statistical consultant and received payment from The School of Nursing and Midwifery Research Unit Queens University Belfast.

Received November 23, 2007; revisions received and accepted April 25, 2008

References

Aran, A., Shalev, R. S., Biran, G., & Gross-Tsur, V. (2007). Parenting style impacts on quality of life in children with cerebral palsy. *Journal of Pediatrics, 151*, 470–475.

Beckung, E., & Hagberg, G. (2000). Correlation between ICIDH handicap code and Gross Motor Function Classificatin System in children with cerebral palsy. *Developmental Medicine and Child Neurology, 42*, 669–673.

Bentler, P. M., & Chou, C. (1987). Practical issues in structural modelling. *Sociological Methods and Research, 16*, 78–117.

Bjornson, K. F., Belza, P. T., Kartin, D., Logson, R. G., & McLaughlin, J. (2008). Self-reported health status and quality of life in youth with cerebral palsy and typically developing youth. *Archives of Physical Medicine and Rehabilitation, 89*, 121–127.

Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen, & J. S. Long (Eds.), *Testing structural equation models* (pp. 136–162). Newbury Park, CA: Sage.

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 April 30, 2006, from http://WWW.biomedcentral.com/1471-2458/6/105

Dickinson, H., Parkinson, K., McManus, V., Arnaud, C., Beckung, E., Fauconnier, J., et al. (2006). Assessment of data quality in a multi-centre cross-sectional study of participation and quality of life of children with cerebral palsy. *BMC Public Health, 6*, 273.

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.

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.

Furlong, W. J., Feeny, D. H., Torrance, G. W., & Barr, R. D. (2001). The Health Utilities Index (HUI®) System for assessing health-related quality of life in clinical studies. *Annals of Medicine, 33*(5), 375–384.

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.

Kerr, C., Parkes, J., Stevenson, M., Cosgrove, A. P., & McDowell, B. C. (2008). Energy efficiency, activity, participation and health status in a population of children with cerebral palsy. *Developmental Medicine and Child Neurology, 50*, 204–210.

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.

Majnemer, A., Shevell, M., Rosenbaum, P., Law, M., & Poulin, C. (2007). Determinants of life quality in school-age children with cerebral palsy. *Journal of Pediatrics, 151*, 470–475.

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.

McCullough, N., & Parkes, J. (August 28, 2007). Use of the child health questionnaire in children with cerebral palsy: A systematic review and evaluation of the psychometric properties. *Journal of Pediatric Psychology*. Epub ahead of print; doi: 10.1093/jpepsy/jsm070.

Morales, N. M. O., Silva, C. H. M., Frontarolli, A. C., Araújo, R. R. H., Rangel, V. O., Pinto, R. M. C., et al. (November 25, 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,* Epub ahead of print; doi: 10.1007/s11136-006-9136-6.

Muthén, L. K., & Muthén, B. O. (2000). *Mplus: The comprehensive modeling program for applied researchers user's guide*. Los Angeles: Muthén & Muthén.

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.

Palisano, R. J., Hanna, S. E., Rosenbaum, P. L., Russell, D. J., Walter, S. D., Wood, E. P., et al. (2000). Validation of a model of gross motor function of children with cerebral palsy. *Physical Therapy, 80*, 974–985.

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

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.

Simeoni, M. C., Schmidt, S., Mühlan, H., Debensason, D., Bullinger, M., & DISABKIDS Group. (2007). Field testing of a European quality of life instrument for children and adolescents with chronic conditions: The 37-item DISBKIDS Chronic Generic Module. *Quality of Life Research, 16*(5), 881–893.

Surveillance of Cerebral Palsy in Europe Collaborative Group (SCPE) (2000). Why a European collaboration of cerebral palsy surveys and registers? *Developmental Medicine andChild 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.

Varni, J. W., Seid, M., & Rode, C. A. (1999). The PedsQoL: Measurement model for the Pediatric Quality of Life Inventory. *Medical Care, 37*, 126–139.

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 botulinum 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.

Wood, E., & Rosenbaum, P. (2000). The gross motor function classification system for cerebral palsy: A study of reliability and stability over time. *Developmental Medicine and Child Neurology, 42*, 292–296.

Wren, T. A., Sheng, M., Hara, R., Otsuka, N. Y., Bowen, R. E., Scaduto, A. A., et al. (2007). Agreement among three instruments for measuring functional health status and quality of life in pediatric orthopaedics. *Journal of Pediatric Orthopedics, 27*(2), 233.