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

Hein Raat, Anita M Botterweck, Jeanne M Landgraf, W Christina Hoogeveen and Marie-Louise Essink-Bot

Updated information and services can be found at:
http://jech.bmjjournals.com/cgi/content/full/59/1/75

These include:

References
This article cites 30 articles, 7 of which can be accessed free at:
http://jech.bmjjournals.com/cgi/content/full/59/1/75#BIBL

1 online articles that cite this article can be accessed at:
http://jech.bmjjournals.com/cgi/content/full/59/1/75#otherarticles

Rapid responses
You can respond to this article at:
http://jech.bmjjournals.com/cgi/eletter-submit/59/1/75

Email alerting service
Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article

Topic collections
Articles on similar topics can be found in the following collections

Other Epidemiology (1668 articles)

Notes

To order reprints of this article go to:
http://www.bmjjournals.com/cgi/reprintform

To subscribe to *Journal of Epidemiology and Community Health* go to:
http://www.bmjjournals.com/subscriptions/

THEORY AND METHODS

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

Hein Raat, Anita M Botterweck, Jeanne M Landgraf, W Christina Hoogeveen, Marie-Louise Essink-Bot

....

See end of article for authors' affiliations
....

Correspondence to:
Dr H Raat, Department of
Public Health, Erasmus
MC, University Medical
Centre Rotterdam, PO Box
1738, 3000 DR
Rotterdam, Netherlands;
h.raat@erasmusmc.nl

Accepted for publication
13 May 2004
....

Study objectives: This study assessed the feasibility, reliability, and validity of the 28 item short child health questionnaire parent form (CHQ-PF28) containing the same 13 scales, but only a subset of the items in the widely used 50 item CHQ-PF50.
Design: Questionnaires were sent to a random regional sample of 2040 parents of schoolchildren (4–13 years); in a random subgroup test-retest reliability was assessed (n = 234). Additionally, the study assessed CHQ-PF28 score distributions and internal consistencies in a nationwide general population sample of (parents of) children aged 4–11 (n = 2474) from Statistics Netherlands.
Main results: Response was 70%. In the school and general population samples seven scales showed ceiling effects. Both CHQ summary measures and one multi-item scale showed adequate internal consistency in both samples (Cronbach's $\alpha > 0.70$). One summary measure and one scale showed excellent test-retest reliability (intraclass correlation coefficient > 0.70); seven scales showed moderate test-retest reliability (intraclass correlation coefficient 0.50–0.70). The CHQ could discriminate between a subgroup with no parent reported chronic conditions (n = 954) and subgroups with asthma (n = 134), frequent headaches (n = 42), and with problems with hearing (n = 38) (Cohen's effect sizes 0.12–0.92; $p < 0.05$ for 39 of 42 comparisons).
Conclusions: This study showed that the CHQ-PF28 resulted in score distributions, and discriminative validity that are comparable to its longer counterpart, but that the internal consistency of most individual scales was low. In community health applications, the CHQ-PF28 may be an acceptable alternative for the longer CHQ-PF50 if the summary measures suffice and reliable estimates of each separate CHQ scale are not required.

Reliable and validated generic health status measures are available to describe health and health related quality of life of children in evaluation studies of community health and clinical interventions,[1–4] burden of disease studies,[5 6] or in community health and clinical practice.[7 8] These measures are applied in addition to condition specific health related quality of life measures and clinical measures, and therefore should be as short as possible without losing precision and reliability.[9] This study is the first to evaluate the reliability, including test-retest reliability and validity of the shortest version of the widely used child health questionnaire (CHQ), the 28 item parent form (CHQ-PF28).[10–13]

The CHQ was developed in the USA, and has since been cross culturally validated into 21 languages (32 countries).[10–20] The CHQ uses the same structure and methodological approach as the SF-36,[21] but was developed specifically for children and therefore includes scales that consider the effects of the child's health on family functioning, as well as specific scales, such as behaviour and self esteem (table 1).[10–12] For adolescents, a self report CHQ child form is available (CHQ-CF87).[11 15 20 22] A CHQ form for pre-school children is in the process of development.[23] The 50 item parent form (CHQ-PF50) for school age children of about 4 or 5 years and older, is the most frequently applied version of the CHQ.[12 14] As young children up to around the age of 10 are considered unable to rate their own health consistently,[1–3] parents generally are used as surrogate responders. As reported in the CHQ user's manual,[11 23] in accordance with general guidelines,[9] regression techniques and item scaling analysis have been applied to derive a shortened CHQ-PF28 from the CHQ-PF50, which contains the same scales, but only a subset of the items that make up the CHQ-PF50. The objectives of this study were to assess, in a large random school based population:

(1) the feasibility of the child health questionnaire (CHQ-PF28) as a proxy measurement of child health and health related quality of life (indicators: response rate, missing/non-unique answers, presence of floor and ceiling effects);

(2) the reliability of the CHQ-PF28 scales (internal consistency and test-retest reliability);

(3) the validity of the CHQ-PF28 as judged by comparisons with 0–100 visual analogue scale (VAS) ratings of the child's health (concurrent validity), and the ability to discriminate between groups with and without specific self reported chronic conditions (discriminative validity).

Additionally, the presence of floor and ceiling effects and the internal consistency of scales were evaluated in a dataset from a nationwide general population sample from Statistics Netherlands (see Methods).[24] We compared the current results for the CHQ-PF28 with earlier findings on CHQ evaluations.[11–20]

....

Abbreviations: CHQ-PF28, child health questionnaire parent form 28 items; CHQ-CF87, child health questionnaire child form 87 items; VAS, visual analogue scale

Table 1 CHQ-PF28 scales, number of items per scale, and score interpretation*

Scale	Number of items	Description low score	Description high score
Physical functioning (PF)	3	Child is limited a lot in performing all physical activities, including self care, because of health	Child performs all types of physical activities, including the most vigorous, without limitations attributable to health
Role functioning: emotional/behaviour REB)	1	Child is limited a lot in school work or activities with friends as a result of emotional or behaviour problems	Child has no limitations in schoolwork or activities with friends as a result of emotional or behaviour problems
Role functioning: physical (RF)	1	Child is limited a lot in school work or activities with friends as a result of physical health	Child has no limitations in schoolwork or activities with friends as a result of physical health
Bodily pain (BP)	1	Child has extremely severe, frequent, and limiting bodily pain	Child has no pain or limitations because of pain
General behaviour (BE)	4	Child very often exhibits aggressive, immature, delinquent behaviour	Child never exhibits aggressive, immature, delinquent behaviour
Mental health (MH)	3	Child has feelings of anxiety and depression all of the time	Child feels peaceful, happy, and calm all of the time
Self esteem (SE)	3	Child is very dissatisfied with abilities, looks, family/peer relationships, and life overall	Child is very satisfied with abilities, looks, family/peer relationships' and life overall
General health perceptions (GH)	4	Parent believes child's health is poor and likely to get worse	Parent believes child's health is excellent and will continue to be so
Parental impact: emotional (PE)	2	Parent experiences a great deal of emotional worry/concern as a result of child's physical and/or psychosocial health	Parent doesn't experience feelings of emotional worry/concern as a result of child's physical and/or psychosocial health
Parental impact: time (PT)	2	Parent experiences a lot of limitations in time available for personal needs because of child's physical and/or psychosocial health	Parent doesn't experience limitations in time available for personal needs because of child's physical and/or psychosocial health
Family activities (FA)	2	The child's health very often limits and interrupts family activities or is a source of family tension.	The child's health never limits or interrupts family activities or is a source of family tension
Family cohesion (FC)	1	Family's ability to get along is rated ''poor''	Family's ability to get along is rated ''excellent''
Change in health (CH)	1	Child's health is much worse now than one year ago	Child's health is much better now than one year ago

*Reproduced with permission from the principal author J M Landgraf (page 38–39).[11]

METHODS

Study population and data collection

In 2001, parents of a random sample of 2040 children (aged 4–13 years) attending any one of 28 elementary schools in two middle sized Dutch cities (Krimpen and Ridderkerk, chosen for practical reasons) were mailed a questionnaire. The 2040 children were selected from a computerised database with all clients (school children) of the municipal health service by applying computer generated random numbers; their parents were invited to participate by the medical director of the municipal health service. The parents themselves decided if either the father or the mother should complete the questionnaire. Up to two reminders were sent; no incentives applied. After three weeks, a random subgroup of 480 parents of these 2040 children who had returned the first questionnaire were mailed the same questionnaire again to assess test-retest reliability.

Questionnaires

The CHQ adaptation into Dutch that included three independent forward and two backward translations was made following international guidelines.[12 14–16 25 26] CHQ-PF28 items have four, five, or six response options, divided over eight multi-item scales and five single item concepts (table 1). Per scale, the items are summed up (some recoded/recalibrated) and transformed into a 0 (worst possible score) to 100 (best possible score) scale.[11] ''Physical'' and ''psychosocial'' CHQ summary scores, which are based on a factor analytical model of a US child population sample, were calculated in a manner analogous to the construction of the summary scores in the SF-36; summary scores of 50 represent the mean in the US reference population children; 10 points above/below 50 reflect one standard deviation difference in either direction.[11 26]

In addition, the questionnaires consisted of items on standard sociodemographic variables and the presence of chronic conditions. Respondents were asked to indicate on a 0–100 VAS, labelled from worst to best imaginable health state, how good or how bad they felt their child's current health state to be.[27 28]

Analysis

Feasibility

We assessed response and missing/non-unique answers for the regional sample of schoolchildren as well as score distributions to evaluate floor and ceiling effects (>25% of the respondents having the lowest/highest score).[15] Floor and ceiling effects were also evaluated in subgroups with a chronic condition (see below).

Additionally, the CHQ-PF28 score distributions (and the internal consistency of multi-item scales; see below) were assessed in a dataset from the 2000 and 2001 random nationwide general population samples from Statistics Netherlands.[24] These data were collected as health measure regarding children aged 4–11 years as a subgroup in the Statistics Netherlands continuous survey of living conditions that included, on a yearly basis, 10 000 persons from all ages in total. (Please see the addendum (available on line http//:www.jech.com/supplemental)) for a table showing the comparison of mean scores and standard deviations of the CHQ-PF28 in a subgroup of children aged 4–11 years of the regional sample of schoolchildren and the whole nationwide general population sample of Statistics Netherlands.) The randomisation of the general population sample from Statistics Netherlands was conducted in two steps: firstly, municipalities were randomly selected from 40 so called COROP areas (strata) in the Netherlands and from every COROP area one or more municipality/municipalities; secondly, persons were randomly selected from the selected municipalities; the response rate was 57%.[24] The data were gathered by parent interviews at home by trained assistants with the use of portable computers for direct data entry: n = 2474; age of the children 4–11 years, mean age 7.5 years (SD 2.3); 49.2% were girls.[24]

Table 2 Score distributions and internal consistency of CHQ-PF28 scales in two random population samples: regional sample of schoolchildren (4–13 years; n = 1435); nationwide general population sample (4–11 years; n = 2474)

CHQ-PF28 scales	Population	Mean (SD)	Range	% min*	% max*	25th %tile	50th† %tile	75th %tile	Cronbach's α‡
Physical functioning	School sample	96.2 (12.4)	0–100	1	86	100	100	100	0.83
	General popn sample	95.1 (15.4)	0–100	1	86	100	100	100	0.85
Role fct-emo/behav	School sample	95.3 (14.5)	0–100	1	89	100	100	100	na
	General popn sample	96.7 (13.2)	0–100	1	93	100	100	100	na
Role funct-physical	School sample	95.7 (13.9)	0–100	0	90	100	100	100	na
	General popn sample	95.3 (16.6)	0–100	1	91	100	100	100	na
Bodily pain	School sample	81.1 (17.3)	0–100	1	31	80	80	100	na
	General popn sample	85.5 (19.6)	0–100	1	52	80	80	100	na
General behaviour	School sample	70.2 (15.9)	0–100	0	3	59	71	81	0.72
	General popn sample	70.3 (15.9)	0–100	0	3	59	71	84	0.56
Mental health	School sample	80.8 (15.1)	8–100	0	19	75	83	92	0.64
	General popn sample	81.8 (16.2)	8–100	0	27	75	83	100	0.59
Self esteem	School sample	77.0 (14.9)	0–100	0	14	67	75	83	0.75
	General popn sample	79.6 (13.9)	0–100	0	15	75	75	92	0.67
General health	School sample	84.1 (17.0)	15–100	0	21	75	90	96	0.48
	General popn sample	85.2 (17.9)	0–100	0	28	75	90	100	0.53
Parental-emotional	School sample	85.6 (16.7)	0–100	0	40	75	88	100	0.44
	General popn sample	89.9 (15.5)	0–100	0	59	88	100	100	0.34
Parental-time	School sample	94.3 (12.9)	0–100	0	78	100	100	100	0.41
	General popn sample	92.4 (19.8)	0–100	2	81	100	100	100	0.68
Family activities	School sample	88.5 (17.8)	0–100	0	59	75	100	100	0.69
	General popn sample	90.7 (16.3)	0–100	0	67	88	100	100	0.53
Family cohesion	School sample	69.0 (19.2)	0–100	0	11	60	60	85	na
	General popn sample	74.9 (18.1)	0–100	0	20	60	85	85	na
Physical summary§	School sample	55.8 (7.6)	0–70	na	na	55	58	60	0.87
	General popn sample	56.1 (8.7)	0–68	na	na	55	59	61	0.88
Psychos summary§	School sample	51.5 (7.9)	1–66	na	na	48	53	57	0.86
	General popn sample	52.5 (7.2)	13–75	na	na	49	53	57	0.80
Change in health¶	School sample	55.7 (16.4)	0–100	1	9	50	50	50	na

*% Of respondents with the highest, respectively lowest possible CHQ-PF28 scale score (ceiling/floor); †median; ‡average α of the eight multi-item scales 0.62 (school sample), respectively 0.59 (general population sample); §physical and psychosocial CHQ summary measures based on a factor analytical model of a US population sample; a score of 50 represents the mean in the general US population; scores above/below 50 are above/below the average in the US reference population[11]; ¶score of 50 shows child's health rating to be about the same now as one year ago; 0 much worse now than one year ago; 100 much better now than one year ago; na, not applicable (single item scales). Please see the addendum (available on line http//:www.jech.com/supplemental) for a table showing the comparison of mean scores and standard deviations of the CHQ-PF28 in a subgroup of children aged 4–11 years of the regional sample of schoolchildren and the whole nationwide general population sample of Statistics Netherlands.

Reliability

Cronbach's α was used to evaluate the internal consistency of the CHQ scales in both samples; a Cronbach's α of a multi-item scale of 0.70 or higher was considered to indicate sufficient internal consistency.[29] We assessed whether (on average) correlation coefficients (Pearson r correlation coefficients) between the items and their own scale score (without the item under consideration) were higher than the correlation coefficients between these items and any other scale, to evaluate whether the CHQ items represent separate domains.

Test-retest reliability of the CHQ-PF28 scales was, at the individual level, assessed by test-retest intraclass correlation coefficients (ICCs)[30]; ICCs of 0.70 or higher were considered to show excellent test-retest reliability and ICCs of 0.50–0.70 to indicate moderate test-retest reliability.[6] Additionally, at the group level, test-retest reliability was assessed by two sided Wilcoxon's signed ranks tests, and by effect sizes: d = [mean(test)−mean(retest))/SD(test]; $0.20 \leqslant d < 0.50$ indicates a small effect size, $0.50 \leqslant d < 0.80$ a moderate effect size, and $d \geqslant 0.80$ a large effect size.[31]

Concurrent validity

Concurrent validity was evaluated by assessing Spearman rank order correlation coefficients between CHQ scale/summary scores and the VAS rating of the child's current health. Concurrent validity is hypothesised to be expressed in positive correlation coefficients. We also hypothesised that the scale general health perceptions has the highest correlations with the VAS rating by the parents, as the content of the other CHQ scales involves a subset of the overall health state as expressed in the VAS rating.

Discriminative validity

The ability of the CHQ to discriminate between a subgroup without parent reported conditions and subgroups with asthma, frequent headaches, and problems with hearing was assessed by Mann Whitney U tests, and effect sizes defined as d = [mean(no conditions)−mean(with condition)]/SD in the condition subgroup; as SDs were generally higher in the subgroups with a reported condition than in the subgroup without conditions this choice resulted in comparatively conservative effect size estimates.[31] Based on the content of the scales and summary measures[11] and the nature of the three parent reported chronic conditions, we hypothesised that the effect sizes (d) will be higher for the physical summary measure than for the psychosocial summary measure in the subgroups with asthma and headaches (in comparison with the subgroup with no parent reported chronic conditions) and the other way around for the subgroup with problems with hearing (as such problems are not directly reflected in the CHQ ''physical'' scales); furthermore that the scale general health perceptions will show large effect sizes in each subgroup with a reported condition; and that the scale bodily pain will be the most affected scale in the subgroup of children with parent reported frequent headaches.

All analyses were done in SPSS, version 10.0. The medical ethical review board of Erasmus MC, University Medical Centre Rotterdam approved this study.

RESULTS
Feasibility

With regard to the 2040 schoolchildren in the regional sample, 1435 parents (70%) responded; a retest questionnaire

Raat, Botterweck, Landgraf, et al

Table 3 Correlation matrix with corrected* item-own scale (printed in bold type), item-other scale, and item-summary measure Pearson r correlation coefficients of CHQ-PF28 scale items in the regional sample of schoolchildren (n = 1435)

	CHQ-PF28 scales and summary measures														
CHQ-PF28 items‡:	PF	REB	RP	BP	BE	MH	SE	GH	PE	PT	FA	FC	PHS†	PSS†	CH
Physical functioning (PF)															
PFALOT	**0.73¶**	0.20¶	0.57¶	0.37¶	0.11¶	0.19¶	0.13¶	0.39¶	0.30¶	0.32¶	0.25¶	0.05	0.76¶	0.11¶	−0.06
PFSOME	**0.80¶**	0.17¶	0.55¶	0.31¶	0.10¶	0.15¶	0.10¶	0.32¶	0.29¶	0.34¶	0.25¶	0.02	0.73¶	0.09¶	−0.05
PFBEND	**0.58¶**	0.18¶	0.36¶	0.18¶	0.10¶	0.16¶	0.09¶	0.21¶	0.20¶	0.28¶	0.25¶	0.06§	0.51¶	0.11¶	−0.02
Role funct-emotional/beha (REB)															
REBAMNT	0.21¶	**na***	0.17¶	0.18¶	0.35¶	0.40¶	0.31¶	0.19¶	0.49¶	0.45¶	0.41¶	0.14¶	0.17¶	0.63¶	−0.04
Role funct-physical (RP)															
RPKIND	0.58¶	0.17¶	**na***	0.42¶	0.08¶	0.19¶	0.14¶	0.29¶	0.36¶	0.33¶	0.27¶	0.01	0.77¶	0.12¶	−0.01
Bodily pain (BP)															
BPFREQ	0.34¶	0.18¶	0.42¶	**na***	0.15¶	0.28¶	0.19¶	0.32¶	0.44¶	0.26¶	0.22¶	0.08¶	0.65¶	0.22¶	0.04
General behaviour (BE)															
BEARGUE	0.06§	0.25¶	0.03	0.14¶	**0.55¶**	0.36¶	0.28¶	0.14¶	0.36¶	0.26¶	0.34¶	0.35¶	−0.03	0.58¶	−0.02
BEATTEN	0.14¶	0.31¶	0.08¶	0.11¶	**0.54¶**	0.38¶	0.35¶	0.17¶	0.35¶	0.28¶	0.36¶	0.22¶	0.00	0.62¶	−0.04
BELIES	0.03	0.17¶	0.02	0.07§	**0.44¶**	0.28¶	0.21¶	0.09¶	0.24¶	0.16¶	0.28¶	0.24¶	−0.07§	0.46¶	0.00
BEGLOBAL	0.12¶	0.30¶	0.12¶	0.13¶	**0.53¶**	0.38¶	0.45¶	0.21¶	0.38¶	0.26¶	0.37¶	0.39¶	0.03	0.62¶	−0.01
Mental health (MH)															
MHLONELY	0.19¶	0.36¶	0.17¶	0.17¶	0.31¶	**0.43¶**	0.38¶	0.19¶	0.40¶	0.30¶	0.29¶	0.20¶	0.11¶	0.60¶	−0.01
MHNERVES	0.11¶	0.24¶	0.10¶	0.22¶	0.35¶	**0.44¶**	0.31¶	0.21¶	0.35¶	0.24¶	0.23¶	0.20¶	0.07¶	0.57¶	−0.06§
MHUPSET	0.14¶	0.33¶	0.16¶	0.25¶	0.42¶	**0.48¶**	0.32¶	0.22¶	0.47¶	0.32¶	0.34¶	0.23¶	0.12¶	0.64¶	−0.02
Self esteem (SE)															
SESCHOOL	0.11¶	0.22¶	0.11¶	0.15¶	0.33¶	0.29¶	**0.54¶**	0.14¶	0.32¶	0.17¶	0.16¶	0.16¶	0.03	0.52¶	0.05
SEFRIEND	0.10¶	0.25¶	0.10¶	0.13¶	0.34¶	0.38¶	**0.58¶**	0.14¶	0.33¶	0.20¶	0.19¶	0.23¶	0.01	0.56¶	−0.01
SEOVERAL	0.10¶	0.28¶	0.14¶	0.19¶	0.42¶	0.43¶	**0.63¶**	0.21¶	0.42¶	0.25¶	0.27¶	0.31¶	0.05§	0.63¶	0.00
General health (GH)															
GHISHLTH	0.37¶	0.17¶	0.31¶	0.35¶	0.17¶	0.22¶	0.02§	**0.45¶**	0.35¶	0.31¶	0.28¶	0.08¶	0.55¶	0.22¶	−0.09¶
GHNVRILL	0.16¶	0.12¶	0.08¶	0.10¶	0.07¶	0.11¶	0.11¶	**0.19¶**	0.12¶	0.11¶	0.11¶	0.07¶	0.34¶	0.05	−0.18¶
GHWORRY	0.22¶	0.22¶	0.17¶	0.16¶	0.09¶	0.16¶	0.29¶	**0.29¶**	0.28¶	0.23¶	0.22¶	0.04§	0.41¶	0.14¶	−0.13¶
GHGLOBAL	0.28¶	0.22¶	0.29¶	0.36¶	0.29¶	0.30¶	0.29¶	**0.33¶**	0.40¶	0.24¶	0.28¶	0.24¶	0.43¶	0.33¶	−0.05§
Parental-emotional (PE)															
PEPHYSI	0.42¶	0.22¶	0.48¶	0.50¶	0.14¶	0.25¶	0.19¶	0.44¶	**0.29¶**	0.36¶	0.31¶	0.06§	0.62¶	0.30¶	−0.02
PEEMBEH	0.12¶	0.52¶	0.14¶	0.25¶	0.53¶	0.56¶	0.47¶	0.22¶	**0.29¶**	0.48¶	0.44¶	0.26¶	0.13¶	0.78¶	−0.06§
Parental-time (PT)															
PTPHYSI	0.46¶	0.17¶	0.49¶	0.28¶	0.08¶	0.13¶	0.06§	0.30¶	0.31¶	**0.27¶**	0.38¶	0.03	0.56¶	0.17¶	−0.05
PTEMBEH	0.16¶	0.50¶	0.09¶	0.16¶	0.40¶	0.42¶	0.30¶	0.21¶	0.51¶	**0.27¶**	0.50¶	0.20¶	0.16¶	0.63¶	−0.02
Family activities (FA)															
FATYPE	0.36¶	0.37¶	0.36¶	0.28¶	0.33¶	0.32¶	0.20¶	0.35¶	0.47¶	0.58¶	**0.54¶**	0.17¶	0.42¶	0.42¶	−0.03
FAINTERU	0.17¶	0.35¶	0.14¶	0.13¶	0.46¶	0.33¶	0.24¶	0.21¶	0.38¶	0.42¶	**0.54¶**	0.24¶	0.14¶	0.48¶	−0.05§
Family cohesion (FC)															
FCCOHSN	0.05	0.14¶	0.01	0.08¶	0.39¶	0.28¶	0.28¶	0.14¶	0.22¶	0.16¶	0.24¶	**na***	−0.01	0.37¶	0.01
Change in health (CH)															
CHANHLTH	−0.05§	−0.04	−0.01	0.04	−0.02	−0.04	0.02	−0.20¶	−0.05§	−0.04	−0.05	0.01	−0.08¶	−0.02	**na***

*Each item was correlated with the applicable scale excluding the item under consideration from the scale score; with regard to two item scales this equals the inter-item correlation; corrected item-own scale correlations cannot be applied to single item scales (na, not applicable); †physical (PHS) and psychosocial (PSS) CHQ summary measures based on a factor analytical model of a US population sample[11]; ‡each item was indicated by an acronym as applied in the CHQ user's manual[11], §p<0.05; ¶p<0.01.

Table 4 Test-retest reliability of the CHQ-PF28 in a random subgroup of the regional sample of schoolchildren (4–12 years old; n = 234)

	Test n = 234 mean (SD)	Retest n = 234 mean (SD)	p Value (Wilcoxon)*	Effect size (d)†	Intraclass correlation coefficients test-retest‡
CHQ-PF28 scales (range 0–100):					
Physical functioning	97 (9)	97 (9)	0.60	−0.04	0.36
Role funct-emotional/behavioural	96 (14)	98 (11)	0.03	0.14	0.50
Role funct-physical	97 (12)	98 (9)	0.49	0.06	0.14
Bodily pain	83 (15)	85 (15)	0.00	0.18	0.42
General behaviour	68 (16)	70 (16)	0.02	0.11	0.75
Mental health	80 (15)	82 (15)	0.04	0.11	0.60
Self esteem	78 (14)	80 (14)	0.03	0.11	0.60
General health	84 (16)	85 (15)	0.03	0.11	0.61
Parental-emotional	86 (16)	88 (14)	0.00	0.15	0.61
Parental-time	94 (13)	96 (12)	0.07	0.11	0.45
Family activities	88 (19)	90 (17)	0.12	0.09	0.65
Family cohesion	70 (18)	71 (17)	0.30	0.05	0.59
Physical summary§	57 (6)	57 (6)	0.06	0.08	0.44
Psychosocial summary§	51 (8)	52 (8)	0.00	0.15	0.78
Change in health¶	56 (16)	56 (14)	0.45	−0.05	0.44

*Non-parametric test for differences between scale scores at the test and at the retest: two sided Wilcoxon's signed ranks test; †difference of the means divided by SD at the first measurement; all effects can be classified as minor (<0.20)[31]; ‡all correlation coefficients are significant (p<0.01); §physical and psychosocial CHQ summary measures based on a factor analytical model of a US population sample; a score of 50 represents the mean in the general US population; scores above/below 50 are above/below the average in the US reference population[11]; ¶score of 50 shows child's health rating to be about the same now as one year ago; 0 much worse now than one year ago; 100 much better now than one year ago.

was send to parents of 329 children (a random subgroup) that resulted in a response by 234 parents (71%). Mean respondent age was 37.7 years (range 23–60; SD 5.2); 86% were mothers; 8% were of non-Dutch ethnic origin; 18% had completed higher vocational education/university, 4% had only elementary or no education; most were employed (54%) or homemakers (35%). The schoolchildren ranged from 4–13 years of age (mean 8.1; SD 2.4); 51% were girls; 7% belonged to a single parent family. Twenty four of 28 CHQ items had less than 1.5% missing answers; the highest percentage found was 1.7% (behaviour item "lied or cheated?"). Twenty seven of 28 CHQ items had less than 0.5% non-unique answers; the single item scale family cohesion had the highest percentage (0.8%).

Score distributions
In both our regional school sample and the nationwide general population sample, seven CHQ scales showed ceiling effects (>25% of the respondents had the maximum score); five scales even showed a profound ceiling effect in both samples (>50% at the extreme) (table 2). (Please see the addendum (available on line http//:www.jech.com/supplemental) for a table showing the comparison of mean scores and standard deviations of the CHQ-PF28 in a subgroup of children aged 4–11 years of the regional sample of schoolchildren and the whole nationwide general population sample of Statistics Netherlands.) In three subgroups of the sample of schoolchildren with a specific condition (asthma, frequent headaches, and problems with hearing; see table 6) fewer ceiling effects were present: bodily pain did not show a ceiling effect in any subgroup; parental-emotional did not show a ceiling effect in the subgroup with problems with hearing; family activities showed less ceiling effect (<50% at the extreme) in subgroups with headaches and problems with hearing.

Internal consistencies
Only one multi-item scale, that of physical functioning, and both CHQ summary measures showed adequate internal consistency in both samples (Cronbach's α>0.70). All multi-

item scales, except for parental-emotional and parental-time, showed higher (corrected) item-own scale correlation coefficients than (on average) item-other scale correlation coefficients (see table 3). Items of scales with a "physical" content showed comparatively high correlation coefficients with the physical summary score, and items of scales with a "psychosocial" content with the psychosocial summary score (table 3).

Table 5 Concurrent validity of the CHQ-PF28 assessed by Spearman rank order correlation coefficients between CHQ-PF28 scale/summary scores and a 0–100 VAS rating of the child's health in the regional sample of schoolchildren (4–13 years; n = 1435)

	0–100 VAS rating of the child's health
	Spearman rank order correlation coefficients*
Single CHQ item general health:	
Single item general health†	0.50
CHQ-PF28 scales:	
Physical functioning	0.27
Role funct-emotional/behavioural	0.17
Role functioning-physical	0.23
Bodily pain	0.35
General behaviour	0.21
Mental health	0.27
Self esteem	0.23
General health	0.39
Parental-emotional	0.34
Parental-time	0.22
Family activities	0.23
Family cohesion	0.15
Physical summary‡	0.32
Psychosocial summary‡	0.27

*All correlation coefficients are significant (p<0.01); †the first item of the CHQ-PF28 scale general health perceptions is: "In general, would you say your child's health is: (1) excellent; (2) very good; (3) good; (4) fair; or (5) poor"; ‡physical and psychosocial CHQ summary measures based on a factor analytical model of a US population sample.[11]

Table 6 Discriminative ability of the CHQ-PF28 between a subgroup without parent reported chronic conditions of the child (n = 954), compared with subgroups with parent reported asthma (n = 134), frequent headaches (n = 42), and problems with hearing (n = 38) in the regional sample of schoolchildren

	Presence of parent reported chronic conditions in subgroups of the regional sample						
	No conditions	Asthma	Asthma versus no conditions	Headaches	Headaches versus no conditions	Problems with hearing	Problems with hearing versus no conditions
	n = 954 mean (SD)	n = 134 mean (SD)	Effect size d*	n = 42 mean (SD)	Effect size d*	n = 38 mean (SD)	Effect size d*
CHQ-PF28 scales (range 0–100):							
Physical functioning	98 (10)	90 (15)‡	0.55**	86 (29)‡	0.40¶	92 (19)‡	0.33¶
Role funct-emo/behave	97 (12)	94 (18)†	0.18	89 (23)‡	0.34¶	89 (22)‡	0.36¶
Role function-physical	97 (11)	91 (19)‡	0.34¶	90 (21)‡	0.36¶	88 (22)‡	0.43¶
Bodily pain	84 (15)	73 (22)‡	0.49¶	64 (21)‡	0.92††	69 (23)‡	0.65**
General behaviour	71 (15)	68 (18)†	0.20¶	62 (17)‡	0.57**	57 (17)‡	0.87††
Mental health	82 (15)	78 (15)‡	0.23¶	76 (17)†	0.34¶	74 (15)‡	0.51**
Self esteem	78 (15)	75 (16)	0.16	72 (14)†	0.41¶	70 (16)‡	0.50**
General health	88 (14)	69 (22)‡	0.85††	76 (20)‡	0.61**	71 (21)‡	0.80††
Parental-emotional	88 (15)	80 (19)‡	0.44¶	74 (23)‡	0.58**	69 (24)‡	0.79**
Parental-time	95 (12)	91 (16)‡	0.25¶	88 (21)‡	0.36¶	90 (15)†	0.32¶
Family activities	90 (16)	86 (19)‡	0.24¶	81 (23)‡	0.41¶	78 (22)‡	0.56**
Family cohesion	70 (19)	67 (20)	0.12	63 (20)†	0.33¶	63 (19)†	0.36¶
Physical summary§	57 (6)	50 (10)‡	0.72**	49 (13)‡	0.58**	51 (11)‡	0.57**
Psychosocial summary§	52 (8)	51 (9)	0.16	48 (9)‡	0.48¶	46 (9)‡	0.71**

*Difference of the means divided by SD in the subgroup with a condition[31]; †p<0.05 (two sided Mann-Whitney U test of differences between the subgroup with a given condition compared with the subgroup with no conditions); ‡p<0.01 (two sided Mann-Whitney U test of differences between the subgroup with a given condition compared with the subgroup with no conditions); §physical and psychosocial CHQ summary measures based on a factor analytical model of a US population sample; a score of 50 represents the mean in the general US population; scores above/below 50 are above/below the average in the US reference population[11]; ¶indicates a small effect (0.20 ⩽ d<0.50)[31]; **indicates a medium effect (0.50 ⩽ d<0.80)[31]; ††indicates a large effect (d⩾0.80).[31]

Test-retest reliability

Seven CHQ scales and the psychosocial summary measure had significantly higher mean retest scores (p<0.05), although with only minor effect sizes (<0.20). Test/retest ICCs ranged from 0.14 to 0.78 (average 0.50; p<0.01) (table 4); general behaviour and the psychosocial summary measure showed excellent test-retest reliability (ICC 0.70 or higher); seven other CHQ scales showed moderate test-retest reliability (ICC 0.50–0.70).[6]

Concurrent validity

All correlation coefficients between CHQ scales/summary measures and the VAS rating of the child's health were positive and significant (p<0.01) as was hypothesised. General health perceptions (specifically the item GHGLOBAL—see table 3—of this scale), correlated best with the VAS rating, as hypothesised (Spearman correlation coefficients 0.39/0.50) (table 5).

Discriminative validity

The prevalences of parent reported chronic conditions were: asthma (9%); frequent headaches (3%); and problems with hearing (3%); 66% had no parent reported condition at all. All CHQ scale and summary scores were lower in the

Key points

- The child health questionnaire parent form 28 items (CHQ-PF28), a short generic measure of health status and health related quality of life with eight multi-item and five single item scales is a feasible measurement instrument for large scale applications in community medicine by paper and pencil (mailed) questionnaires as well as oral interviews.
- This study provides reference (norm) scores for the CHQ-PF28 for age groups 4–11/4–13 years, derived from large random school based and general population samples.
- Score distributions and concurrent and discriminative validity of the "short" CHQ-PF28 items are comparable to the results of its longer counterpart, the CHQ-PF50 items.
- The physical and psychosocial summary measures of the "short" CHQ-PF28 show adequate internal consistency and its psychosocial summary measure has excellent test-retest reliability, but individual CHQ-PF28 scales may have a low internal consistency and test-retest reliability.

Policy implications

- Feasible, reliable, and validated measures are needed to describe health and health related quality of life of children in evaluation studies of community health and clinical interventions, burden of disease studies, or in community health and clinical practice; these measures may be applied in addition to condition specific health related quality of life measures and in addition to various clinical measures.
- The child health questionnaire parent forms (CHQ-PF) provide a comparatively well evaluated option for measurement of generic health status and health related quality of life in child populations.
- The "short" CHQ-PF28 items offers an acceptable alternative to the longer CHQ-PF50 items if the physical and psychosocial summary measures are of primary interest and reliable estimates of each separate CHQ scale are not required for the purpose of health measurement.

subgroups with a reported condition compared with the subgroup with no reported condition; all differences but three (regarding the subgroup with asthma compared with no conditions) were significant (p<0.05) (table 6). In the subgroups with reported asthma and frequent headaches, the impact on the physical summary measure as expressed by the effect sizes was higher than on the psychosocial summary measure, while the reverse was true in the subgroup with problems with hearing, as hypothesised (table 6). The scale general health perceptions was significantly affected in each of the three subgroups with a condition, as hypothesised, but only resulted in large effect sizes in the subgroups with asthma and problems with hearing; in the subgroup with frequent headaches the scale bodily pain, as hypothesised, showed the largest effect size (table 6).

DISCUSSION

This first evaluation of the 28 item short form of the CHQ for parents in an independent large random sample of school children re-established the feasibility of the CHQ-PF as a paper and pencil health status questionnaire[12] [14] [19]; large scale oral CHQ-PF28 interviewing by trained interviewers in a nationwide random general population sample also proved to be feasible.[24] This study supports the concurrent and discriminative validity of the CHQ-PF28 and provides reference (norm) scores derived from school based and general population samples, but gives rise to some concerns about ceiling effects, the internal consistency of (short) multi-item scales, and test-retest reliability, requiring further investigation.

Limitations

Limitations of our study include the choice of the sample(s) and study design issues. Our primary study group consisted of a random sample of predominantly healthy school children (66% had no parent reported chronic conditions). So, the results of this study are of primary interest for community health applications such as burden of disease studies,[5] evaluations of preventive interventions in the general population, or future applications in the daily practice of community medicine, for example, applications by school based nurses.[7] [8] We recommend, however, evaluations in other populations as well.

Evaluation of test-retest reliability in our study did not include an assessment of health transition that may occur between test and retest, which we recommend to include in future studies. The responsiveness of the CHQ to changes in medical/social conditions was not evaluated and therefore remains to be studied. Comparisons between our sample and the national sample of Statistics Netherlands provide only preliminary insights as in the Statistics Netherlands study a different mode of data gathering and questionnaire administration had been applied.

In this study, only the CHQ-PF28 items were administered, as mingling the extra items of the CHQ-PF50 with the regular CHQ-PF28 items may influence the results. However, additional analyses may be recommended in other, existing, datasets that do include the CHQ-PF50 concerning results regarding the subset of CHQ-PF28 items in comparison with the results regarding all CHQ-PF50 items.

Score distributions

In both samples in this study, seven CHQ-PF28 scales showed a percentage of respondents higher than 25% that have the maximum score of the scale. This finding (that is, ceiling effect) is common in paediatric health measurement and health related quality of life studies; it is equally apparent in studies with other CHQ versions and with other measurement instruments.[11] [12] [14] [15] [17–20] [32–34] However, it limits the use

of these instruments to detect changes in a generally healthy population, or to describe excellent health beyond the average in comparatively healthy populations. In specific populations with chronic conditions ceiling effects may be less pronounced as was shown in this study.[14]

Internal consistency

Levels of reliability of health status measures in children may be low relative to instruments designed for adults,[1–3] especially in the case of shortened scales. With regard to the CHQ-PF28 we recommend to restrict the evaluations to the CHQ summary measures, which showed to have adequate internal consistency.

Evaluation of CHQ-PF28 (corrected) item-own scale correlations and item-other scale correlation coefficients showed that all CHQ scales concerning the health status of the children themselves represented separate entities.

Test-retest reliability

One CHQ-PF28 scale and a summary measure showed excellent test-retest reliability and seven scales moderate test-retest reliability. The CHQ single item scales showed, overall, lower test-retest reliability than multi-item scales/measures, which illustrates that measurement of concept via multiple items may increase reliability.[6] [29] Seven scales and the psychosocial summary measure had statistically significantly higher retest scores, although effect sizes were minor. This might reflect a somewhat lower prevalence of, for example, viral infections at the retest (later in the spring season).

Validity

The statistically significant, positive correlation coefficients between CHQ scales/summary measures and the VAS rating of the child's health by the parent supported the concurrent validity of the CHQ-PF28. However, the results illustrate that CHQ scales other than general health (for example, those related to role functioning, psychosocial health, and family cohesion) measure concepts that extend beyond the mere measurement of health in general such as was done by the VAS rating by the parents. This study showed the ability of the CHQ-PF to discriminate between absence/presence of three parent reported conditions, with scoring patterns that generally confirm hypotheses that were based on the nature of the three conditions. This is in accordance with earlier reports on the CHQ-PF50 and CHQ-CF87.[11–20] We recommend further assessments of the validity of the CHQ-PF28 by comparing scores between clinical groups with reported medical conditions, in addition to this study that included parent reports.[1–3]

Conclusions

There is a clear need for feasible, reliable, and valid measures to describe generic health status and health related quality of life in child populations; this is equally true for community health and for clinical applications, and in the future possibly for applications in daily medical (preventive) practice.[1–8] The CHQ is such a measure, and short forms like the CHQ-PF28 are especially welcome given the overload of items in most questionnaires. This study showed the score distributions and concurrent and discriminative validity of the CHQ-PF28 to be comparable to its longer counterpart, the CHQ-PF50. However, the internal consistency and test-retest reliability of many individual CHQ-PF28 scales were comparatively low. The two CHQ summary score measures however did show adequate internal consistency, while the psychosocial summary measure also showed excellent test-retest reliability. In community health applications, therefore, the CHQ-PF28 offers an acceptable alternative to the longer CHQ-PF50 if

evaluation of the summary measures suffices and reliable estimates of each separate CHQ scale are not required for the purpose of health measurement. In addition to our study, we recommend that further assessment of the CHQ-PF28 be made in varied clinical samples, as well as a close evaluation of both the responsiveness to change and test-retest characteristics.

ACKNOWLEDGEMENTS

The addendum is available on line (http://www.jech.com/supplemental).

. . . .

Authors' affiliations
H Raat, M-L Essink-Bot, Department of Public Health, Erasmus MC, University Medical Centre Rotterdam, Netherlands
A M Botterweck, Statistics Netherlands, Voorburg/Heerlen, Netherlands
J M Landgraf, HealthAct, Boston, USA
W C Hoogeveen, Department of Youth, GGD, Municipal Health Service Rotterdam, Netherlands

Conflicts of interest: none declared.

The interpretations reported in this article are the author's (AMB) and do not necessarily correspond with the policy of Statistics Netherlands.

REFERENCES

1 **Pal DK**. Quality of life assessment in children: a review of conceptual and methodological issues in multidimensional health status measures. *J Epidemiol Community Health* 1996;**50**:391–6.
2 **Jenney ME**, Campbell S. Measuring quality of life. *Arch Dis Child* 1997;**77**:347–50.
3 **Connolly MA**, Johnson JA. Measuring quality of life in paediatric patients. *Pharmacoeconomics* 1999;**16**:605–25.
4 **Lohr KN**. Health outcomes methodology symposium: summary and recommendations. *Med Care* 2000;**38**(suppl):II194–208.
5 **Morrow RH**, Hyder AA, Murray CJ, et al. Measuring the burden of disease. *Lancet* 1998;**352**:1859–61.
6 **Andresen EM**, Catlin TK, Wyrwich KW, et al. Retest reliability of surveillance questions on health related quality of life. *J Epidemiol Community Health* 2003;**57**:339–43.
7 **Higginson IJ**, Carr AJ. Measuring quality of life: Using quality of life measures in the clinical setting. *BMJ* 2001;**322**:1297–300.
8 **Mangunkusumo RT**, Moorman PW, Van den Berg-de Ruiter AE, et al. Internet-administered adolescent health questionnaires compared to a paper version in a randomized study. *J Adolesc Health* (in press).
9 **Coste J**, Guillemin F, Pouchot J, et al. Methodological approaches to shortening composite measurement scales. *J Clin Epidemiol* 1997;**50**:247–52.
10 **Kurtin PS**, Landgraf JM, Abetz L. Patient-based health status measurements in pediatric dialysis: expanding the assessment of outcome. *Am J Kidney Dis* 1994;**24**:376–82.
11 **Landgraf JM**, Abetz L, Ware JE. *The CHQ user's manual.* Boston: The Health Institute, New England Medical Center, 1996.
12 **Landgraf JM**, Maunsell E, Speechley KN, et al. Canadian-French, German and UK versions of the child health questionnaire: methodology and preliminary item scaling results. *Qual Life Res* 1998;**7**:433–45.
13 **Nixon Speechley K**, Maunsell E, Desmeules M, et al. Mutual concurrent validity of the child health questionnaire and the health utilities index: an exploratory analysis using survivors of childhood cancer. *Int J Cancer* 1999;**12**(suppl):95–105.
14 **Raat H**, Bonsel GJ, Essink-Bot ML, et al. Reliability and validity of comprehensive health status measures in children. The child health questionnaire in relation to the health utilities index. *J Clin Epidemiol* 2002;**55**:67–76.
15 **Raat H**, Landgraf JM, Bonsel GJ, et al. Reliability and validity of the child health questionnaire-child form (CHQ-CF87) in a Dutch adolescent population. *Qual Life Res* 2002;**11**:575–81.
16 **Ruperto N**, Ravelli A, Pistorio A, et al. Cross-cultural adaptation and psychometric evaluation of the childhood health assessment questionnaire (CHAQ) and the child health questionnaire in 32 countries. Review of the general methodology. *Clin Exp Rheumatol* 2001;**19**(suppl 23):S1–9.
17 **Wulffraat N**, van der Net JJ, Ruperto N, et al. The Dutch version of the childhood health assessment questionnaire (CHAQ) and the child health questionnaire (CHQ). *Clin Exp Rheumatol* 2001;**19**(suppl 23):S111–15.
18 **Joos R**, Ruperto N, Wouters C, et al. The Belgian-Flemish version of the childhood health assessment questionnaire (CHAQ) and the child health questionnaire (CHQ). *Clin Exp Rheumatol* 2001;**19**(suppl 23):S20–4.
19 **Waters E**, Salmon L, Wake M. The parent-form child health questionnaire in Australia: comparison of reliability, validity, structure, and norms. *J Pediatr Psychol* 2000;**25**:381–91.
20 **Waters EB**, Salmon LA, Wake M, et al. The health and well-being of adolescents: a school-based population study of the self-report child health questionnaire. *J Adolesc Health* 2001;**29**:140–9.
21 **Hemingway H**, Stafford M, Stansfeld S, et al. Is the SF-36 a valid measure of change in population health? Results from the Whitehall II study. *BMJ* 1997;**315**:1273–9.
22 **Landgraf JM**, Abetz LN. Functional status and well-being of children representing three cultural groups: initial self-reports using the CHQ-CF87. *Psychology and Health* 1997;**12**:839–54.
23 **Landgraf JM**. Measuring pediatric outcomes in applied clinical settings: an update about the child health questionnaire (CHQ). *Quality of Life Newsletter MAPI Research Institute* 1999;**23**:5–6.
24 **Botterweck A**, Frenken F, Janssen S, et al. Feasibility, reliability and validity of the new measurements in the Dutch Health Interview Survey 2001 [Plausibiliteit nieuwe metingen algemene gezondheid en leefstijlen 2001]. Voorburg/Heerlen, Netherlands: Statistics Netherlands, 2003.
25 **Guillemin F**, Bombardier C, Beaton D. Cross-cultural adaptation of health-related quality of life measures: literature review and proposed guidelines. *J Clin Epidemiol* 1993;**46**:1417–32.
26 **Ware JE Jr**, Gandek B, Kosinski M, et al. The equivalence of SF-36 summary health scores estimated using standard and country-specific algorithms in 10 countries: results from the IQOLA Project. International quality of life assessment. *J Clin Epidemiol* 1998;**51**:1167–70.
27 **Brooks R**. EuroQol: the current state of play. *Health Policy* 1996;**37**:53–72.
28 **Krabbe PF**, Essink-Bot ML, Bonsel GJ. The comparability and reliability of five health-state valuation methods. *Soc Sci Med* 1997;**45**:1641–52.
29 **Bland JM**, Altman DG. Cronbach's alpha. *BMJ* 1997;**314**:572.
30 **Deyo RA**, Diehr P, Patrick DL. Reproducibility and responsiveness of health status measures. Statistics and strategies for evaluation. *Control Clin Trials* 1991;**12**:142–58S.
31 **Cohen J**. *Statistical power analysis for the behavioral sciences.* New York: Academic Press, 1977.
32 **Gemke RJ**, Bonsel GJ. Reliability and validity of a comprehensive health status measure in a heterogeneous population of children admitted to intensive care. *J Clin Epidemiol* 1996;**49**:327–33.
33 **Ravens-Sieberer U**, Bullinger M. Assessing health-related quality of life in chronically ill children with the German KINDL: first psychometric and content analytical results. *Qual Life Res* 1998;**7**:399–407.
34 **Vogels T**, Verrips GH, Verloove-Vanhorick SP, et al. Measuring health-related quality of life in children: the development of the TACQOL parent form. *Qual Life Res* 1998;**7**:457–65.