



Reliability and validity of health status measurement by the TAPQOL

E M Bunge, M-L Essink-Bot, M P H M Kobussen, L W A van Suijlekom-Smit, H A Moll and H Raat

Arch. Dis. Child. 2005;90;351-358
doi:10.1136/adc.2003.048645

Updated information and services can be found at:
<http://adc.bmjournals.com/cgi/content/full/90/4/351>

These include:

References

This article cites 25 articles, 8 of which can be accessed free at:
<http://adc.bmjournals.com/cgi/content/full/90/4/351#BIBL>

Rapid responses

You can respond to this article at:
<http://adc.bmjournals.com/cgi/eletter-submit/90/4/351>

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 Public Health](#) (2616 articles)
[Other Pediatrics](#) (1856 articles)

Notes

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

To subscribe to *Archives of Disease in Childhood* go to:
<http://www.bmjournals.com/subscriptions/>

ORIGINAL ARTICLE

Reliability and validity of health status measurement by the TAPQOL

E M Bunge, M-L Essink-Bot, M P H M Kobussen, L W A van Suijlekom-Smit, H A Moll, H Raat

Arch Dis Child 2005;90:351–358. doi: 10.1136/adc.2003.048645

See end of article for authors' affiliations

Correspondence to:
Ms E Bunge, Department of Public Health, Erasmus MC, University Medical Center Rotterdam, PO Box 1738, 3000 DR Rotterdam, Netherlands; e.bunge@erasmusmc.nl

Accepted 11 October 2004

Background: In addition to clinical measures in the evaluation of paediatric interventions, health related quality of life (HRQoL) is an important outcome. The TAPQOL (TNO-AZL Preschool children Quality of Life) was developed to measure HRQoL in preschool children. It is a generic instrument consisting of 12 scales that cover the domains physical, social, cognitive, and emotional functioning.

Aims: To evaluate the feasibility, score distribution, internal consistency, test-retest reliability, and discriminative and concurrent validity of the TAPQOL multi-item scales in preschool children, aged 2–48 months. Also to evaluate the feasibility, reliability, and validity separately for infants (2–12 months old) and toddlers (12–48 months old).

Methods: Parents of a random general population sample of 500 preschool children were sent a questionnaire by mail. A random subgroup of 159 parents who participated received a retest after two weeks.

Results: The response rate was 83% at the test and 75% at the retest. There were few missing answers. Six scales showed ceiling effects. Nine scales had Cronbach's alphas >0.70. In general, score distributions and Cronbach's alphas were comparable for infants and toddlers. Test-retest showed no significant differences in mean scale scores; two scales had intra-class correlations <0.50. Five scales showed significant differences between children with no conditions versus children with two or more parent reported chronic conditions.

Conclusion: Results showed that the TAPQOL is a feasible instrument to measure HRQoL and support the reliability and discriminative validity of the majority of its scales for infants as well as toddlers.

Health status and health related quality of life (HRQoL) measures are used for the evaluation of healthcare intervention in community medicine and clinical practice.^{1–6} Furthermore, HRQoL measures are used for descriptive studies; for example, burden of disease studies in public health^{7–8} and follow up studies of distinct patients groups.⁹ In the future, possibilities may arise for applications in daily medical practice in both community and clinical medicine.¹⁰

Few HRQoL measures are available for preschool children.^{11–17} A reason for this might be that young children show a fast development of cognitive, motor, and behavioural functions, especially during the first years of life.¹⁸ This means that instruments which intend to cover a relatively wide age range (for example, 0–4 years) have to somehow accommodate for this.

The TAPQOL is the first multi-dimensional HRQoL measure that was specifically designed for preschool children aged 1–5 years.^{14–17} As preschool children cannot complete questionnaires by themselves, the TAPQOL uses a proxy, mostly a parent. In this study we evaluated the psychometric properties of the TAPQOL including, for the first time, assessment of the test-retest reliability. Additionally, also for the first time, we applied the TAPQOL to infants (2–12 months old) and specifically evaluated its performance in this subgroup.

The aim of this study was to evaluate the feasibility, score distribution, internal consistency, test-retest reliability, and discriminative and concurrent validity of the TAPQOL multi-item scales in preschool children, aged 2–48 months. In addition, the feasibility, reliability, and validity were evaluated separately for infants (2–12 months old) and toddlers (12–48 months old).

METHODS

The Medical Ethical Committee of the Erasmus MC, University Medical Center Rotterdam, approved this study.

Population and data collection

Parents of a random general population sample of 500 preschool children (2–48 months old) in the eastern part of the Netherlands were sent the TAPQOL questionnaire by mail. The parents themselves decided which parent should participate. In case of non-response each household received maximally two reminder letters; no incentives to participate were given. Two weeks later, a random subgroup of 158 participating parents received the same questionnaire to assess test-retest reliability. The completed TAPQOL questionnaires were returned by mail.

Only parents who were considered to be able to adequately read and write Dutch were eligible for analysis. This was operationalised as at least one parent being Dutch or, if both parents were of foreign origin, that they should have an education of higher vocational level or have a university degree.

Questionnaire

The TAPQOL is a 43 item questionnaire consisting of 12 multi-item scales that cover the domains physical, social, cognitive, and emotional functioning (see fig 1). The number of items per scale ranges from three to seven. TAPQOL items generally relate to the past three months, but this may be adjusted for specific research aims. For all scales, the presence of a specific complaint or limitation was scored on a three point scale, namely "never", "occasionally", and "often". For seven TAPQOL scales ("stomach problems", "skin problems", "lung problems", "sleeping", "appetite",

Items on the TAPQOL questionnaire

TAPQOL scales

PHYSICAL FUNCTIONING**Sleeping**

How did your child sleep...

- Items
- (1) did your child sleep restlessly
 - (2) was your child awake at night
 - (3) did your child cry at night
 - (4) did your child have difficulty sleeping through the night

Appetite

How did your child eat and drink...

- (5) was your child's appetite poor
- (6) did your child have difficulty eating enough
- (7) did your child refuse to eat

Lungs

Has your child had.../Has your child been...

- (8) bronchitis
- (9) difficulty breathing or lung problems
- (10) short of breath

Stomach

Has your child had.../Has your child been...

- (11) stomach ache or abdominal pain
- (12) colic
- (13) nauseous

Skin

Has your child had...

- (14) eczema
- (15) itchiness
- (16) dry child

Motor functioning

Did your child have...

- (17) difficulty with walking
- (18) difficulty with running
- (19) difficulty with walking up stairs without help
- (20) difficulty with balance

SOCIAL FUNCTIONING**Social functioning**

How was your child's behaviour with older children

- (21) my child was able to play happily with other children
- (22) my child was at ease with other children
- (23) my child was confident with other children

Problem behaviour

Your child's behaviour...

- (24) my child was short-tempered
- (25) my child was aggressive
- (26) my child was irritable
- (27) my child was angry
- (28) my child was restless or impatient with me
- (29) my child was defiant/awkward with me
- (30) I could not manage my child

COGNITIVE FUNCTIONING**Communication**

Did your child have...

- (31) difficulty in understanding what others said
- (32) difficulty in talking clearly
- (33) difficulty in saying what he/she meant
- (34) difficulty in making it clear what he/she wanted

EMOTIONAL FUNCTIONING**Anxiety**

How was your child...

- (35) frightened
- (36) tense
- (37) anxious

Positive mood

How was your child...

- (38) in good spirits
- (39) cheerful
- (40) happy

Liveliness

How was your child...

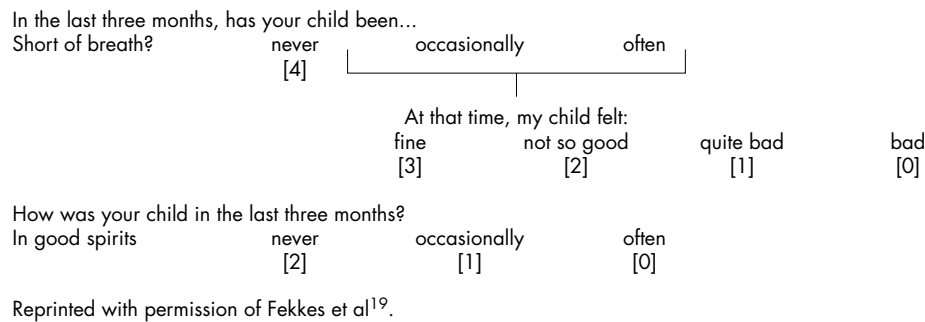
- (41) energetic
- (42) active
- (43) lively

Reprinted with permission of Fekkes et al¹⁹.**Figure 1** Items on the TAPQOL questionnaire.

“motor functioning”, and “communication”), first the presence of a specific complaint or limitation is recorded and, if this is the case, the wellbeing of the child related to that complaint or limitation is measured on a four point scale, namely “fine”, “not so good”, “quite bad”, and “bad”. Scale scores were calculated by adding up item scores within scales, and transforming crude scale scores linearly to a

0–100 scale, with higher scores indicating better quality of life (see fig 2 for an example). The scales “social functioning”, “motor functioning”, and “communication” are only relevant for children aged 1½ years and older.¹⁹

The TAPQOL is available in a Dutch as well as in an English version, translated from Dutch according to international guidelines.²⁰

**Figure 2** Example item scores TAPQOL.

Besides the TAPQOL, demographic variables and the prevalence of chronic conditions and visits to the general practitioner were assessed. Questions about chronic conditions covered: asthma or recurrent problems of the respiratory tract, recurrent otitis or having tympanostomy tubes, defective vision in which glasses are not helpful, regular abdominal pain, allergies, eczema, and other conditions.

Analyses

In accordance with the TAPQOL guidelines all items of a three item scale should be completed in order to be eligible for analysis. In scales with four items one missing answer is

allowed; in the seven item scale, "problem behaviour", two missing answers were allowed. In case of non-unique answers (more than one answer per question), one answer was imputed randomly.

Feasibility of the TAPQOL was evaluated by assessing the response rate and by evaluating both the number of missing answers per item and the number of non-unique answers per item. Cronbach's alpha was used to determine the internal consistency of the scales.²¹ Separate analyses were made for the subgroup with two or more parent reported chronic conditions. Average correlation coefficients were calculated between items and their own scale (without the item under

Table 1 Score distribution and internal consistency of TAPQOL scales in 410 children: 92 infants aged 2–12 months and 318 toddlers aged 12–48 months

TAPQOL scales	Population (n=410)	Mean (SD)	Range	% min*	% max†	Median	Cronbach's alpha‡
Sleeping	Total	79.4 (19.0)	13–100	0.0	29.4	81.3	0.89
	Infants	82.1 (18.5)	31–100	0.0	37.4	87.5	0.88
	Toddlers	78.6 (19.1)	13–100	0.0	27.1	81.3	0.89
Appetite	Total	84.1 (13.9)	33–100	0.0	30.4	83.3	0.79
	Infants	91.9 (10.7)	58–100	0.0	55.6	100	0.70
	Toddlers	81.9 (14.0)	33–100	0.0	23.2	83.3	0.78
Lung problems	Total	91.9 (16.5)	0–100	0.3	73.4	100	0.81
	Infants	89.1 (17.2)	42–100	0.0	62.9	100	0.80
	Toddlers	92.7 (16.2)	0–100	0.3	76.4	100	0.82
Stomach problems	Total	91.6 (13.2)	42–100	0.0	63.3	100	0.47
	Infants	90.3 (14.1)	42–100	0.0	58.0	100	0.46
	Toddlers	91.9 (12.9)	50–100	0.0	64.8	100	0.53
Skin problems	Total	91.4 (12.3)	17–100	0.0	48.0	91.7	0.77
	Infants	92.6 (10.3)	50–100	0.0	45.6	91.7	0.73
	Toddlers	91.0 (12.8)	17–100	0.0	48.7	91.7	0.78
Motor functioning	Total	97.8 (8.5)	25–100	0.0	86.7	100	0.89
	Infants	NA	NA	NA	NA	NA	NA
	Toddlers‡	97.8 (8.5)	25–100	0.0	86.7	100	0.89
Problem behaviour	Total	73.0 (16.9)	14–100	0.0	10.8	71.4	0.81
	Infants	87.3 (12.7)	57–100	0.0	34.4	85.7	0.73
	Toddlers	68.9 (15.6)	14–100	0.0	4.1	71.4	0.78
Social functioning	Total	91.8 (15.3)	33–100	0.0	70.8	100	0.69
	Infants	NA	NA	NA	NA	NA	NA
	Toddlers‡	91.8 (15.3)	33–100	0.0	70.8	100	0.69
Communication	Total	89.4 (12.7)	31–100	0.0	41.6	93.8	0.81
	Infants	NA	NA	NA	NA	NA	NA
	Toddlers‡	89.4 (12.7)	31–100	0.0	41.6	93.8	0.81
Positive mood	Total	97.7 (10.4)	0–100	0.2	94.4	100	0.92
	Infants	97.6 (10.4)	50–100	0.0	94.5	100	0.97
	Toddlers	97.7 (10.4)	0–100	0.3	94.3	100	0.91
Anxiety	Total	77.4 (18.7)	0–100	0.2	27.9	100	0.62
	Infants	83.1 (17.8)	33–100	0.0	41.1	83.3	0.56
	Toddlers	75.8 (18.6)	0–100	0.3	24.2	83.3	0.63
Liveliness	Total	96.0 (12.3)	0–100	0.2	88.3	100	0.76
	Infants	96.0 (10.9)	50–100	0.0	85.7	100	0.58
	Toddlers	96.1 (12.7)	0–100	0.3	89.0	100	0.80

*Respondents with the worst possible score.

†Respondents with the best possible score.

‡Average alpha (total group) of the 12 scales is 0.77.

§Only for children aged 18 months or older (n=264).

NA, not applicable.

consideration) and between items and every other scale, to determine whether the items were well chosen and if the scales represent different domains. The average corrected item-own scale correlation coefficients are expected to be higher than the average item-other scale correlation coefficients. At the group level, test-retest reliability was assessed by the Wilcoxon signed ranks test. We used non-parametric tests, because data were skewed and TAPQOL scales are not continuous but have a lowest and highest possible value that will show ceiling effects. Cohen's effect size,²² which relates the difference in mean scores between test and retest to the dispersion of the scores of the test, were calculated: $d = [\text{mean}(a) - \text{mean}(b)]/\text{SD}$ at the test.²² At the individual level, the intra-class correlation (ICC) was applied to assess test-retest reliability.²³ Discriminant validity was evaluated by comparing TAPQOL scale scores of a subgroup of children with no parent reported chronic conditions with those of a subgroup of children with two or more parent reported chronic conditions. The Mann-Whitney U test was used to determine differences in mean scale scores between the two groups. Cohen's effect sizes were calculated: $d = [\text{mean}(a) - \text{mean}(b)]/\text{SD}$ of the subgroup with parent reported conditions. Comparisons were also made of TAPQOL scale scores between a subgroup of children with zero or one visit to the general practitioner and a subgroup of children with four or more visits to the general practitioner in the last year. Spearman's rank order correlation coefficients were applied to evaluate concurrent validity of the TAPQOL with a single item general health rating: "In general, would you

say your child's health is: excellent, very good, good, fair, or poor".

SPSS 10.0 was used for the analysis.

RESULTS

Response, feasibility, and sample characteristics

Response rate was 83.0%; five (1.2%) questionnaires were not eligible for analysis (see methods). Response rate at the retest was 75.3%; one questionnaire was not eligible for analysis; 115 retest questionnaires could be matched to a test questionnaire (same child, same respondent). Mean age of the parent respondents was 33.1 (SD 7.1) years; 97% of the respondents were mothers. Most lived together with their partner (98%); 50% of the respondents had a part-time job and 36% were homemakers; 33% of the respondents had an education at intermediate vocational level, 39% had a lower, and 27% a higher educational level than intermediate vocational education.

Fifty per cent of the children eligible for analysis were girls and 22% of the children were infants (between 2 and 12 months old).

There were few missing answers on the TAPQOL (circa 1% per item) and very few non-unique answers (less than 1% per item).

Score distribution and internal consistency

Six scales had ceiling effects (that is, >50% of the respondents had the maximum score). When the total group

Table 2 Average inter-item, corrected item-own scale, and item-other scale correlations* of the TAPQOL scales in 410 children: 92 infants aged 2–12 months and 318 toddlers aged 12–48 months

TAPQOL scales	Population	Average inter-item correlation	Average item-own scale correlation*	Average item-other scale correlation
Sleeping	Total	0.67	0.76	-0.16
	Infants	0.66	0.75	-0.22
	Toddlers	0.67	0.76	-0.15
Appetite	Total	0.55	0.63	-0.10
	Infants	0.44	0.52	-0.17
	Toddlers	0.54	0.62	-0.09
Lung problems	Total	0.58	0.67	-0.10
	Infants	0.56	0.64	-0.15
	Toddlers	0.59	0.68	-0.10
Stomach problems	Total	0.22	0.30	-0.09
	Infants	0.22	0.29	-0.16
	Toddlers	0.25	0.36	-0.09
Skin problems	Total	0.53	0.61	-0.05
	Infants	0.48	0.55	-0.08
	Toddlers	0.55	0.62	-0.05
Motor functioning	Total	0.67	0.76	-0.10
	Infants	NA	NA	NA
	Toddlers†	0.67	0.76	-0.10
Problem behaviour	Total	0.38	0.55	-0.11
	Infants	0.27	0.45	-0.15
	Toddlers	0.33	0.50	-0.10
Social functioning	Total	0.44	0.52	0.07
	Infants	NA	NA	NA
	Toddlers†	0.44	0.52	0.07
Communication	Total	0.51	0.63	-0.09
	Infants	NA	NA	NA
	Toddlers†	0.51	0.63	-0.09
Positive mood	Total	0.79	0.84	0.15
	Infants	0.93	0.94	0.23
	Toddlers	0.76	0.81	0.14
Anxiety	Total	0.35	0.43	-0.10
	Infants	0.29	0.37	-0.11
	Toddlers	0.36	0.44	-0.09
Liveliness	Total	0.57	0.63	0.05
	Infants	0.33	0.41	0.05
	Toddlers	0.64	0.70	0.05

*Each item was correlated with the applicable scale excluding the item under consideration from the scale score.

†Only for children aged 18 months or older (n = 264).

NA, not applicable.

Table 3 Test-retest reliability of the TAPQOL in a subgroup of 115 preschool children: 28 infants aged 2–12 months and 87 toddlers aged 12–48 months

TAPQOL scales	Population	Test mean (SD)	Retest mean (SD)	p value (WSR)†	Effect size‡	Intra-class correlation Test-retest
Sleeping	Total	79 (19)	79 (18)	0.96	−0.02	0.65*
	Infants	88 (14)	86 (17)	0.49	−0.11	0.63*
	Toddlers	76 (19)	76 (18)	0.70	0.0	0.63*
Appetite	Total	84 (14)	85 (14)	0.33	0.05	0.69*
	Infants	96 (8)	92 (13)	0.20	−0.50	0.19
	Toddlers	81 (14)	83 (14)	0.05	0.17	0.76*
Lung problems	Total	92 (17)	93 (15)	0.18	0.08	0.74*
	Infants	95 (11)	99 (5)	0.02	0.33	0.63*
	Toddlers	91 (18)	91 (16)	0.67	0.03	0.74*
Stomach problems	Total	91 (13)	90 (14)	0.94	−0.03	0.35*
	Infants	92 (14)	96 (7)	0.20	0.29	0.02
	Toddlers	90 (13)	88 (15)	0.43	−0.14	0.41*
Skin problems	Total	92 (11)	93 (11)	0.14	0.09	0.84*
	Infants	94 (7)	95 (8)	0.78	0.04	0.70*
	Toddlers	91 (12)	92 (12)	0.13	0.10	0.86*
Motor functioning	Total	98 (9)	97 (12)	0.16	−0.14	0.88*
	Infants	NA	NA	NA	NA	NA
	Toddlers§	98 (9)	97 (12)	0.16	−0.14	0.88*
Problem behaviour	Total	72 (19)	73 (16)	0.28	0.05	0.81*
	Infants	91 (11)	89 (8)	0.30	−0.18	0.57*
	Toddlers	65 (16)	67 (15)	0.09	0.13	0.74*
Social functioning	Total	92 (15)	95 (11)	0.05	0.21	0.51*
	Infants	NA	NA	NA	NA	NA
	Toddlers§	92 (15)	95 (11)	0.05	0.21	0.51*
Communication	Total	88 (14)	90 (11)	0.13	−0.13	0.68*
	Infants	NA	NA	NA	NA	NA
	Toddlers§	88 (14)	90 (11)	0.13	−0.13	0.68*
Positive mood	Total	96 (14)	96 (14)	0.77	0.03	0.48*
	Infants	98 (10)	98 (9)	0.32	0.06	0.95*
	Toddlers	95 (16)	96 (15)	0.89	0.02	0.42*
Anxiety	Total	76 (21)	78 (20)	0.14	0.13	0.50*
	Infants	84 (20)	94 (12)	0.02	0.51	0.25
	Toddlers	73 (20)	73 (20)	0.77	0.0	0.50*
Liveliness	Total	97 (10)	96 (13)	0.15	−0.14	0.62*
	Infants	96 (11)	98 (10)	0.41	0.10	0.74*
	Toddlers	98 (9)	95 (14)	0.06	−0.23	0.59*

*p<0.01.

†Non-parametric test for differences between the average scale scores at the test and at the retest: two-sided Wilcoxon signed ranks test.

‡Differences between the means divided by SD at the first measurement.

§Only for children aged 18 months or older (n=76).

NA, not applicable.

showed a ceiling effect, then both the infants and toddlers as subgroups did so. On only one item (“appetite”) did the subgroup infants show a ceiling effect whereas the subgroup toddlers and total group did not. In the total group, nine scales had Cronbach’s alpha >0.70. The subgroup infants (except for “liveliness”) and the subgroup toddlers showed sufficient internal consistency for the same scales as the total group, but in general, the subgroup infants had somewhat lower Cronbach’s alphas than the subgroup toddlers (table 1). For the subgroup with two or more parent reported chronic conditions, five scales (same scales as for the total group except for “lung problems”) showed ceiling effects. Overall, the percentages of respondents with a maximum score were lower in this subgroup than in the total group. In this subgroup, eight scales had Cronbach’s alphas >0.70. These were the same scales as in the total group; only liveliness showed in this subgroup a Cronbach’s alpha below 0.70.

There were no differences with regard to scale means between boys and girls, except that girls had a higher mean score than boys on the scale “communication” (p < 0.01).

All scales had higher average corrected item-own scale correlation coefficients than the corresponding average item-other scale correlation coefficients in the total group, as well as in the subgroups infants and toddlers (table 2).

Test-retest reliability

In the total group there were no significant differences in mean scale scores between test and retest. The subgroup

infants showed significant differences between mean scores for “lung problems” and “anxiety”; toddlers did not show significant differences. Two scales in the total group, three scales in the subgroup infants, and two scales in the subgroup toddlers had an ICC <0.50. For most scales ICCs were lower in the subgroup of infants than in the subgroup of toddlers, except for “liveliness and “positive mood” (table 3).

Discriminant validity

The most prevalent parent reported chronic conditions were asthma (20%), eczema (14%), and regular otitis or having tympanostomy tubes (11%); the remainder of the conditions were prevalent in less than 6% of the children. For the total group, five scales (“sleeping”, “appetite”, “lung problems”, “skin problems”, and “problem behaviour”) showed significantly different mean scores between the subgroup of children with zero parent reported chronic conditions versus the subgroup of children with two or more conditions. Cohen’s effect sizes were large for the scales “lung problems” and “skin problems”. In general, the subgroup infants showed the same effect sizes as the subgroup toddlers, except for “sleeping” and “lung problems”, where the effect sizes for the subgroup infants were much larger than for the subgroup toddlers (table 4).

For the number of visits to the general practitioner, six scales (“sleeping”, “appetite”, “lung problems”, “stomach problems”, “skin problems”, and “problem behaviour”) showed significant mean scale score differences between

Table 4 Mean (SD) scores of the TAPQOL scales, separately for all ages, infants, and toddlers, for the subgroup without parent reported conditions (n = 240; of which 57 infants and 183 toddlers), the subgroup with one condition (n = 113; of which 24 infants and 89 toddlers), and the subgroup with two or more conditions (n = 57; of which 11 infants and 46 toddlers)

TAPQOL scales	Population	Number of parent reported chronic conditions per child				
		0 conditions n = 240 Mean (SD)	1 condition n = 113 Mean (SD)	1 v 0 conditions Effect size Size†	≥2 conditions n = 57 Mean (SD)	≥2 v 0 conditions Effect size
Sleeping	Total	83 (18)	77 (19)	0.34**	69 (21)	0.68**
	Infants	86 (16) ¹	81 (20) ²	0.27	62 (16) ³	1.50**
	Toddlers	82 (18) ⁴	76 (18) ⁵	0.35**	71 (21) ⁶	0.53**
Appetite	Total	85 (14)	85 (13)	0.04	79 (17)	0.34*
	Infants	93 (10)	90 (11)	0.27	88 (14)	0.39*
	Toddlers	83 (14)	83 (13)	-0.03	77 (17)	0.31
Lung problems	Total	97 (10)	89 (19)	0.40**	75 (22)	1.02**
	Infants	96 (10)	85 (17)	0.64**	60 (17)	2.18**
	Toddlers	97 (9)	90 (20)	0.34**	78 (21)	0.88**
Stomach problems	Total	93 (12)	91 (14)	0.12	88 (16)	0.29
	Infants	92 (13)	89 (15)	0.17	86 (19)	0.30
	Toddlers	93 (12)	92 (13)	0.10	88 (16)	0.29
Skin problems	Total	96 (6)	88 (13)	0.60**	80 (19)	0.80**
	Infants	96 (6)	91 (12)	0.40	81 (17)	0.87**
	Toddlers	96 (6)	87 (13)	0.64**	80 (20)	0.78**
Motor functioning	Total	99 (7)	97 (11)	0.18	96 (9)	0.27
	Infants	NA	NA	NA	NA	NA
	Toddlers‡	99 (7)	97 (11)	0.18	96 (9)	0.27
Problem behaviour	Total	73 (17)	75 (16)	-0.10	67 (17)	0.38*
	Infants	88 (11)	90 (13)	-0.14	80 (17)	0.45
	Toddlers	69 (16)	71 (15)	-0.14	64 (16)	0.34
Social functioning	Total	93 (15)	90 (18)	0.14	92 (12)	0.04
	Infants	NA	NA	NA	NA	NA
	Toddlers‡	93 (15)	90 (18)	0.14	92 (12)	0.04
Communication	Total	91 (11)	88 (15)	0.19	87 (16)	0.25
	Infants	NA	NA	NA	NA	NA
	Toddlers‡	91 (11)	88 (15)	0.19	87 (16)	0.25
Positive mood	Total	98 (8)	97 (12)	0.08	95 (15)	0.23
	Infants	98 (10)	100 (0)	NA	91 (20)	0.35
	Toddlers	99 (7)	97 (14)	0.13	96 (13)	0.20
Anxiety	Total	78 (19)	79 (17)	-0.06	74 (19)	0.18
	Infants	84 (18)	84 (13)	0.01	76 (23)	0.37
	Toddlers	76 (19)	77 (18)	-0.09	74 (18)	0.09
Liveliness	Total	96 (12)	97 (13)	-0.05	95 (12)	0.05
	Infants	96 (12)	98 (6)	-0.42	94 (15)	0.10
	Toddlers	96 (13)	96 (14)	-0.01	96 (11)	0.04

*Significant at the 0.05 level.

**Significant at the 0.01 level.

†Differences of the means divided by SD in the subgroup with condition(s).

‡Only for children aged 18 months old or older.

NA, not applicable.

Differences in mean scores between the subgroups without conditions and with conditions were evaluated by Cohen's effect size and by Mann Whitney U-test.

the subgroup of children with zero or one visit and the subgroup of children with four or more visits in the last year for the total group. "Appetite" and "stomach problems" showed no significant differences in the subgroup infants. The scales "sleeping" and "lung problems" had large Cohen's effect sizes, especially in the subgroup infants.

Concurrent validity

There were low but significant Spearman's correlation coefficients in the expected direction between nine TAPQOL scales and a single item general health rating. Six of the nine scales suitable for the subgroup infants showed larger correlation coefficients between TAPQOL scales and a single item general health rating than in the subgroup toddlers (table 5).

DISCUSSION

This study, with a very high response rate,²⁴ established the feasibility of the parent completed TAPQOL questionnaire for preschool children in a large general population sample; psychometric properties were generally adequate in the total group as well as in the subgroups infants and toddlers.

Because our study was limited to a random general population sample, we could not evaluate the applicability

of the TAPQOL in clinical populations. We had only one cross-sectional assessment and a retest; therefore, we could not evaluate responsiveness to change in health status over time. Another limitation is that we are unaware of the adequacy of proxy rating (by parents) which are indispensable for this age group. Proxy rating may be confounded by many factors.^{13 25 26}

Our results can be compared only with those from the study by Fekkes and colleagues,¹⁴ and our data confirm their results concerning ceiling effects. The phenomenon "ceiling effects" may limit the use of the TAPQOL to detect changes and to describe health beyond the average in relatively healthy populations. In general, our Cronbach's alphas were somewhat higher than in the study of Fekkes *et al*, but especially for "skin problems", "motor functioning", "communication", and "positive mood" our Cronbach's alphas were much higher than reported by Fekkes and colleagues.¹⁴ For discriminant validity, both Fekkes *et al* and our study found significant differences in mean scale scores in the physical functioning domain for children with and without parent reported chronic conditions.

Test-retest reliability was low for some scales; a phenomenon that has also been reported in evaluations of other instruments.^{27 28} For the scales "stomach problems" and

Table 5 Concurrent validity: Spearman's correlation coefficients and Pearson correlations between TAPQOL scales and the CHQ-IT general health question in 410 children: 92 infants aged 2–12 months and 318 toddlers aged 12–48 months

TAPQOL scales	Population	Single item general health rating	
		Spearman rho correlation coefficient	Pearson r correlation coefficient
Sleeping	Total	0.36**	0.35**
	Infants	0.42**	0.39**
	Toddlers	0.34**	0.33**
Appetite	Total	0.24**	0.25**
	Infants	0.28**	0.31**
	Toddlers	0.22**	0.24**
Lung problems	Total	0.34**	0.40**
	Infants	0.55**	0.59**
	Toddlers	0.28**	0.34**
Stomach problems	Total	0.16**	0.21**
	Infants	0.26*	0.34**
	Toddlers	0.14*	0.17**
Skin problems	Total	0.13**	0.16**
	Infants	0.08	0.16**
	Toddlers	0.15**	0.16**
Motor functioning	Total	0.13*	0.19**
	Infants	NA	NA
	Toddlers‡	0.13*	0.19**
Problem behaviour	Total	0.09	0.09
	Infants	0.08	0.11
	Toddlers	0.05	0.07
Social functioning	Total	0.10	0.11
	Infants	NA	NA
	Toddlers‡	0.10	0.11
Communication	Total	0.18**	0.21**
	Infants	NA	NA
	Toddlers‡	0.18**	0.21**
Positive mood	Total	0.12*	0.17**
	Infants	0.28**	0.33**
	Toddlers	0.08	0.12*
Anxiety	Total	0.15**	0.14**
	Infants	0.37**	0.28**
	Toddlers	0.08	0.10
Liveliness	Total	0.03	0.04
	Infants	−0.04	0.08
	Toddlers	0.05	0.03

*Significant at the 0.05 level.

**Significant at the 0.01 level.

NA, not applicable.

‡Only for children aged 18 months or older (n = 264)

“anxiety” the Cronbach's alpha is also low. We suggest further research on this topic, as test-retest reliability should be adequately shown, especially in studies with repeated measurements.

Ceiling effects were present in the total group as well as in the subgroup with parent reported chronic conditions, although in the subgroup to a lesser degree (five scales instead of six scales with ceiling effects and fewer respondents with maximum scores). This can be interpreted as follows. The chronic conditions mentioned by the parent mostly affected physical functioning. The scales belonging to this domain did show differences between children with and without parent reported chronic conditions. The other domains seemed not to be affected in these conditions. We suggest further evaluation of the TAPQOL in patient groups with distinct conditions that affect the emotional, social, and cognitive TAPQOL domains, such as children with attention deficit hyperactivity disorder (ADHD) or mental retardation.

In conclusion, our study that was conducted in the setting of community medicine showed that the TAPQOL is a feasible and reliable instrument to measure health status and health related quality of life. Our results suggest that the TAPQOL will also be applicable in the clinical setting with conditions that affect physical functioning, since it clearly discriminated between children with and without parent

reported chronic conditions with a physical nature. Although the TAPQOL was not originally designed for infants, our study supports the reliability and discriminative validity of the majority of its scales, not only for toddlers but also for infants. We propose further research, including cross-cultural validation,²⁹ evaluations in clinical samples, and evaluations of responsiveness to community or clinical interventions.

ACKNOWLEDGEMENTS

Community Care Salland was responsible for the data collection. We like to thank the physicians, nurses, physician's assistants, and managers of the Home Care for facilitating this project. We are also very grateful to the parents who participated in this study.

Authors' affiliations

E M Bunge, M-L Essink-Bot, H Raat, Department of Public Health, Erasmus MC, University Medical Center Rotterdam, Netherlands
M P H M Kobussen, Community Care Salland, Ommen, Netherlands
L W A van Suijlekom-Smit, H A Moll, Department of Paediatrics, Erasmus MC–University Medical Center Rotterdam, Netherlands
H Raat, GGD–Municipal Health Service, Rotterdam, Netherlands

Funding: This study was funded by the Netherlands Organisation for Health Research and Development (ZonMw) NWO-Health Care Efficiency Research Program Grant # 2200.0128

Competing interests: none declared

REFERENCES

- 1 Eiser C. Children's quality of life measures. *Arch Dis Child* 1997;**77**:350-4.
- 2 Connolly MA, Johnson JA. Measuring quality of life in paediatric patients. *Pharmacoeconomics* 1999;**16**:605-25.
- 3 Jenney ME, Campbell S. Measuring quality of life. *Arch Dis Child* 1997;**77**:347-50.
- 4 Lohr KN. Health outcomes methodology symposium: summary and recommendations. *Med Care* 2000;**38**(9 suppl):1194-208.
- 5 Brazier J, Deverill M, Green C, et al. A review of the use of health status measures in economic evaluation. *Health Technol Assess* 1993;**3**(9):i-iv, 1-164.
- 6 Gold MR, Siegel JE, Russel LB, et al. *Cost-effectiveness in health and medicine*. New York/Oxford: Oxford University Press, 1996.
- 7 Manuel DG, Schultz SE, Kopec JA. Measuring the health burden of chronic disease and injury using health adjusted life expectancy and the Health Utilities Index. *J Epidemiol Community Health* 2002;**56**:843-50.
- 8 Gold MR, Muennig P. Measure-dependent variation in burden of disease estimates: implications for policy. *Med Care* 2002;**40**:260-6.
- 9 Taylor A, Butt W, Ciardulli M. The functional outcome and quality of life of children after admission to an intensive care unit. *Intensive Care Med* 2003;**29**:795-800.
- 10 Higginson IJ, Carr AJ. Measuring quality of life: using quality of life measures in the clinical setting. *BMJ* 2001;**322**:1297-300.
- 11 Stein RE, Jessop DJ. Functional status II(R). A measure of child health status. *Med Care* 1990;**28**:1041-55.
- 12 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.
- 13 Hack M. Consideration of the use of health status, functional outcome, and quality-of-life to monitor neonatal intensive care practice. *Pediatrics* 1999;**103**(1 suppl E):319-28.
- 14 Fekkes M, Theunissen NC, Brugman E, et al. Development and psychometric evaluation of the TAPQOL: a health-related quality of life instrument for 1-5-year-old children. *Qual Life Res* 2000;**9**:961-72.
- 15 Sturms LM, van der Sluis CK, Groothoff JW, et al. The health-related quality of life of pediatric traffic victims. *J Trauma* 2002;**52**:88-94.
- 16 Veen S, Fekkes M, Koopman HM, et al. Quality of life in preschool children born preterm. *Dev Med Child Neurol* 2001;**43**:460-5.
- 17 Rovers MM, Krabbe PF, Straatman H, et al. Randomised controlled trial of the effect of ventilation tubes (grommets) on quality of life at age 1-2 years. *Arch Dis Child* 2001;**84**:45-9.
- 18 Behrman RE, Kliegman RM, Arvin AM, Nelson WE, eds. *Nelson textbook of pediatrics*, 15th edn. Philadelphia: WB Saunders Company, 1996.
- 19 TNO. *TNO-AZL pre-school children quality of life users manual*. TNO PG, Leiden, Netherlands, 2004.
- 20 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.
- 21 Bland JM, Altman DG. Cronbach's alpha. *BMJ* 1997;**314**:572.
- 22 Cohen J. *Statistical power analysis for the behavioral sciences*. New York: Academic Press, 1977.
- 23 Deyo RA, Diehr P, Patrick DL. Reproducibility and responsiveness of health status measures. Statistics and strategies for evaluation. *Control Clin Trials* 1991;**12**(4 suppl):142S-158S.
- 24 Asch DA, Jedrzejewski MK, Christakis NA. Response rates to mail surveys published in medical journals. *J Clin Epidemiol* 1997;**50**:1129-36.
- 25 Addington-Hall J, Kalra L. Who should measure quality of life? *BMJ* 2001;**322**:1417-20.
- 26 Waters E, Doyle J, Wolfe R, et al. Influence of parental gender and self-reported health and illness on parent-reported child health. *Pediatrics* 2000;**106**:1422-8.
- 27 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.
- 28 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.
- 29 Anderson RT, Aaronson NK, Bullinger M, et al. A review of the progress towards developing health-related quality-of-life instruments for international clinical studies and outcomes research. *Pharmacoeconomics* 1996;**10**:336-55.

Register now!

10th European Forum on Quality Improvement in Health Care
 13-15 April 2005, ExCel Conference Centre, London
 For further information on how to register please go to:
<http://www.quality.bmj.com>