

# BMJ Open Feasibility and validity of the Health Status Classification System-Preschool (HSCS-PS) in a large community sample: the Generation R study

Xinye Fang,<sup>1,2</sup> Guannan Bai,<sup>3,4</sup> Dafna A Windhorst,<sup>3</sup> David Feeny,<sup>5</sup> Saroj Saigal,<sup>6</sup> Liesbeth Duijts,<sup>7,8</sup> Vincent W V Jaddoe,<sup>4,9,10</sup> Shanlian Hu,<sup>2</sup> Chunlin Jin,<sup>1,2</sup> Hein Raat<sup>3</sup>

**To cite:** Fang X, Bai G, Windhorst DA, *et al.* Feasibility and validity of the Health Status Classification System-Preschool (HSCS-PS) in a large community sample: the Generation R study. *BMJ Open* 2018;**8**:e022449. doi:10.1136/bmjopen-2018-022449

► Prepublication history and additional material for this paper are available online. To view these files, please visit the journal online (<http://dx.doi.org/10.1136/bmjopen-2018-022449>).

Received 18 February 2018  
Revised 27 September 2018  
Accepted 4 October 2018



© Author(s) (or their employer(s)) 2018. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

For numbered affiliations see end of article.

**Correspondence to**  
Professor Hein Raat;  
[h.raat@erasmusmc.nl](mailto:h.raat@erasmusmc.nl)

## ABSTRACT

**Objectives** To evaluate the feasibility, discriminant validity and concurrent validity of the Health Status Classification System-Preschool (HSCS-PS) in children aged 3 years in a large community sample in the Netherlands.

**Design/setting** A prospective population-based cohort in Rotterdam, the Netherlands.

**Participants** A questionnaire was administered to a sample of parents of 4546 children (36.7±1.5 months).

**Outcome measures** Health-related quality of life (HRQOL) of children was measured by HSCS-PS. The HSCS-PS consists of 10 original domains. Two single-item measures of 'General health' and 'Behavior' were added. A disability score was calculated by summing up all 10 original domains to describe the overall health status. Feasibility was assessed by the response rate, percentages of missing answers, score distributions and the presence of floor/ceiling effects. Discriminant validity was analysed between subgroups with predefined conditions: low birth weight, preterm birth, wheezing, Ear-Nose-Throat surgical procedures and behaviour problems. In the absence of another HRQOL measure, this study uses the single-items 'General health' and 'Behavior' as a first step to evaluate concurrent validity of the HSCS-PS.

**Results** Feasibility: response rate was 69%. Ceiling effects were observed in all domains. Discriminant validity: the disability score discriminated clearly between subgroups of children born with a 'very low birth weight', 'very preterm birth', with 'four or more than four times wheezing', 'at least one ear-nose-throat surgical procedures', 'behaviour problems present' and the 'reference' group. Concurrent validity: HSCS-PS domains correlated better with hypothesised parallel additional domains than with other non-hypothesised original domains.

**Conclusions** This study supports the feasibility and validity of the HSCS-PS among preschoolers in community settings. We recommend developing a utility-based scoring algorithm for the HSCS-PS. Further empirical studies and repeated evaluations in varied populations are recommended.

## INTRODUCTION

Patient-reported health status and health-related quality of life (HRQOL) are essential

## Strength and limitations of this study

- This study is the first to apply and to evaluate the Health Status Classification System-Preschool (HSCS-PS) in a very large community sample of preschool children.
- Because our sample consisted of relatively healthy children, we were not able to describe the HSCS-PS in seriously ill populations.
- The population for analysis was relatively well-educated and non-migrant families compared with the non-respondents which may limit the generalisability of the results.
- Some indicators for the reliability of the HSCS-PS, such as inter-rater reliability and test-retest reliability were not able to be evaluated in our study. We recommend to assess this in future studies.
- In the absence of an accepted 'gold standard' in this study, we recommend to assess the association between HSCS-PS with one or more other established measures of health-related quality of life in a future study.

outcome measures in addition to clinical outcomes in both general medicine and paediatrics.<sup>1,2</sup> HRQOL refers to quality of life as modified by the functional states, impairments, perceptions and social opportunities as influenced by chronic conditions, injury, treatment or policy.<sup>3,4</sup> Studies on HRQOL of preschool children are scarce due to the early stage of development and the need for proxy reporting.<sup>5</sup> In addition to 'health profile measures' (eg, the Infant and Toddler Quality of Life Questionnaire; ITQOL) for economic analyses and guiding value-based healthcare, we need preference-based measures where the 'health status description' is 'valued' (ie, 'weighted') by a relevant panel in society.<sup>6-9</sup> The most widely used preference-based measure is the Health Utilities Index (HUI; ie, HUI2 and HUI3) for children/people aged 4 years and above.<sup>1,10</sup> There is a need

for a similar, preference-based instrument for preschool children.

The Health Status Classification System-Preschool (HSCS-PS), developed by Saigal *et al*<sup>11</sup> is a multidimensional system to describe the HRQOL of preschool children aged 2.5–5 years.<sup>11</sup> It is a parental (or clinician) proxy measurement of the health status of the child with a structure similar to the HUI.<sup>12</sup> The instrument includes 10 mutually exclusive domains, that is, 'Vision', 'Hearing', 'Speech', 'Mobility', 'Dexterity', 'Self-care', 'Emotion', 'Learning and remembering', 'Thinking and problem solving', 'Pain and discomfort'. Saigal *et al*<sup>11</sup> proposed two additional parent-reported single-item measures: 'General health' and 'Behavior'.<sup>11</sup>

So far, the HSCS-PS has been validated in clinical cohorts of children with a very low birth weight (VLBW) and children with cerebral palsy.<sup>11</sup> The HSCS-PS was applied in studies regarding the development of health of young children after extremely preterm birth.<sup>13 14</sup> The reliability and validity of the HSCS-PS were supported in previous studies that applied HSCS-PS in patients who were diagnosed with neuroblastoma at 2–5 years of age.<sup>15 16</sup> Little is known about the feasibility and validity among children in generally healthy populations.

The present study uses a large general population sample to describe and evaluate the parent-completed HSCS-PS by using information regarding birth outcomes (eg, low birth weight (LBW) and preterm birth), wheezing and ear-nose-throat (ENT) surgical procedures, and behaviour problem in preschool children identified by the Child Behavior Checklist (CBCL 1.5–5). Previous studies have shown that children with the above-mentioned health condition were reported by their parents or caregivers with relatively low HRQOL.<sup>17–25</sup> For example, the parent-reported HRQOL of preschool children born preterm or born with a VLBW was lower than HRQOL of those who were not born preterm or with an LBW.<sup>17–20</sup>

In the absence of another HRQOL measure, this study uses the above-mentioned parent-reported single-items regarding 'General health'<sup>26</sup> and 'Behavior'<sup>27</sup> as a first step to evaluate concurrent validity of the HSCS-PS.

The aims of this study are to assess: (1) The feasibility of the HSCS-PS considering the response rate, missing answers, score distributions and presence of floor/ceiling effects; (2) The discriminant validity by comparing HSCS-PS scores between subgroups in the general population with presence/absence of LBW, premature delivery, wheezing, ENT surgical procedures and CBCL behaviour problems; and (3) as a first step regarding the concurrent validity by evaluation of the correlations between the original HSCS-PS scores and the 'General health' and 'Behavior' single-item measures.

## METHODS

### Study design

This study was embedded in the Generation R Study, a population-based prospective child cohort study from

fetal life onwards in Rotterdam, the Netherlands.<sup>28</sup> All children were born between April 2002 and January 2006.

A total of 7893 children were included in the postnatal follow-up studies.<sup>28</sup> In the survey after 36 months of birth, parental consent was available for 7294 children. Children whose caregivers did not fill out the questionnaire at age 3 years (n=2280) were excluded. Additionally, we excluded children with missing data on one or more domains of the HSCS-PS (n=468), leaving 4546 children for the analyses (see online supplementary figure S1).

### Public involvement

Generation R discusses the strategy of the cohort study and the outcomes of the studies with the Municipality of Rotterdam on a regular basis; as the Municipality represents parents (and youth) in general. Moreover, participating parents and youth are regularly informed by newsletters regarding general outcomes of Generation R studies, and by a personal 'passport' with findings from the measurements for the participating family. Parents and youth are invited to comment on the outcomes at the website. At the individual level, the 'passport' and the individual results are discussed with the parents (and youth) by a physician, after the measurements on a certain day are finished.

### Health Status Classification System—Preschool

The HSCS-PS is a parent-reported health status questionnaire applicable to children aged 2.5–5 years which consists of 10 mutually exclusive domains, based on the HUI.<sup>11</sup> In addition, Saigal *et al* proposed two additional parent-reported single-item questions regarding 'general health' and 'behaviour', given the relatively high prevalence of general health and behaviour problems among the VLBW infants.<sup>29 30</sup> The HSCS-PS was initially applied to approximately 80 children across Canada by paediatricians and neonatologists regarding the structured and qualitative feedback. After several rounds of refinements, the final version contains 10 domains each with 3–5 levels, and the two additional items. (see online supplementary table S1). The overall health status is described as a 10-element vector consisting of one level for each of the domains. In this study, to facilitate comparisons between groups, a total 'disability score' for the overall health state of a child was calculated as the sum of the level codes for the original domains. Therefore, the range of the disability score varied from 10 (no disability on any domain) to 41 (maximum disability on all 10 domains).<sup>16</sup>

### Birth outcomes

In the present study, birth weight and gestational age at birth was obtained from medical records. LBW was defined as a birth weight less than 2500 g. To construct extreme groups, we further divided LBW into VLBW (<1500 g) and moderate LBW (1500–2500 g).<sup>31</sup> Children were defined as preterm when they were born alive before 37 weeks of gestation. Preterm birth was further

subdivided into very preterm (<32 weeks) and moderate to late preterm (32–37 weeks).<sup>32</sup>

### Wheezing

Parent-reported frequency of wheezing in the past 12 months at age 3 years was assessed using core questions from the International Study of Asthma and Allergies in Childhood and classified as 'no wheezing', '1–3 episodes' and '≥4 episodes'.<sup>33</sup>

### ENT surgical procedures

At age 3 years, the parents were asked whether the child had undergone an ENT surgical procedure (ie, removal of the adenoids, removal of the tonsils and inserting tubes to aerate the middle ear).<sup>34</sup> If at least one of these procedures was reported, the child was classified as 'with ENT surgical procedure(s)'.<sup>35</sup>

### Behaviour problems

The presence of child behaviour problems was assessed at age 3 years by the CBCL parent questionnaire.<sup>35</sup> A borderline cut-off score (83rd percentile of a Dutch norm group) of the CBCL total problem score was used to differentiate between children with and without behaviour problems in the borderline/clinical range.<sup>36</sup>

### Other data

Socio-demographic characteristics were assessed by parent questionnaires, including marital status, educational level and ethnic background of the main caregiver, household income and child's age when the questionnaire was completed. Child's gender was obtained from medical records.

### Statistical analyses

The scores of the HSCS-PS domains and the HSCS-PS disability score were treated as continuous variables. Statistical analyses were conducted in SPSS, V.21.0 for Windows (IBM).

### Feasibility

Feasibility of the HSCS-PS was evaluated by assessing the response rate, percentage of missing answers, score distributions and the presence of floor/ceiling effects (ie, >50% of the respondents in the best/worst option).

### Discriminant validity

We evaluated the ability of the HSCS-PS to discriminate between subgroups with and without LBW, preterm birth, wheezing, ENT surgical procedures and behaviour problems. Additionally, we calculated how many of these five conditions a child had (ie, whether a child had a LBW, was born preterm, is reported to have wheezing, to have had ENT surgical procedures, to have behaviour problems). This cumulative number of conditions was recoded into four categories: no condition, one condition, two conditions and three or more conditions. The ability of the HSCS-PS to discriminate between subgroups differing in the number of conditions was assessed.

Because of the non-normal distribution of HSCS-PS scores, Mann-Whitney non-parametric tests were used to assess differences in HSCS-PS scores between subgroups. Additionally, Cohen's effect sizes ( $d$ ) were calculated by dividing the difference in mean scores between subgroups by the largest SD, and interpreted as:  $0.2 \leq d < 0.5$  small difference,  $0.5 \leq d < 0.8$  moderate difference,  $d \geq 0.8$  large difference.<sup>37</sup> Significant differences were indicated at the level of  $p < 0.05$ . We expected that the disability score would be higher in the subgroups in which the children were reported to have a 'condition' (preterm birth, low birth weight, wheezing, ENT surgical procedure, or behavior problems) compared with the reference group without this condition. Additionally, we hypothesised that the disability score would be higher in the subgroups with a higher number of conditions, compared with the reference subgroup in which the children were reported to have none of the conditions.

### Concurrent validity

In the absence of a 'gold standard' measure of HRQOL, as a first step to evaluate the concurrent validity of the 10-domains HSCS-PS, it was assessed whether specific HSCS-PS domains correlated better with their assumed 'parallel' single-item measures of 'General health' and/or 'Behavior' than with a 'non-parallel' measure. Considering the non-normal distribution of the data, Spearman rank correlation was applied. We calculated bootstrapped 95% CIs for Spearman correlation coefficients. When (a) the 95% CI is not 'across 0'; and (b) the  $p$  value  $< 0.05$ , the correlation coefficient was regarded as statistically significant. We hypothesised relatively high correlation coefficients between the following 'parallel' pairs of a HSCS-PS domain/single-item parent-rated measure (in italics): 'Pain and discomfort'/'General health'; 'Self-care'/'Behavior'; 'Emotion'/'Behavior'; 'Learning and remembering'/'Behavior'; 'Thinking and problem solving'/'Behavior'; and we hypothesised the correlation coefficients for all other pairs to be lower.

### Non-response analysis

Children with missing data on the HSCS-PS at age 3 years, including children whose parents did not complete the entire questionnaire ( $n=2748$ ) were compared with children who did not have missing data on any HSCS-PS domain and thus were included in the analyses ( $n=4546$ ).

## RESULTS

Of the respondents, 94.3% were mothers. Children's mean age at the HSCS-PS questionnaire was 36.7 months (SD 1.5); 49.6% were boys; 5.2% of the children had a LBW (<2500 g); 6% of the children were born preterm (gestational age at birth <37 weeks); 12.7% had wheezing in the previous year; 11.5% had any previous ENT surgical procedure and 5.7% had parent-reported behaviour problems (see online supplementary table S2).

**Table 1** Score distributions of the HSCS-PS domains in the population sample (n=4546)

HSCS-PS domains	Mean (SD)	Range	% of min*	% of max†	25th %tile	50th %tile‡	75th %tile
<b>Original domains</b>							
Vision	1.02 (0.18)	1–4	98.9	0	1	1	1
Hearing	1.02 (0.14)	1–3	98.3	0	1	1	1
Speech	1.35 (0.51)	1–4	66.8	0.2	1	1	2
Mobility	1.02 (0.17)	1–4	98.1	0.1	1	1	1
Dexterity	1.01 (0.11)	1–3	99.5	0	1	1	1
Self-care	1.12 (0.36)	1–4	89.4	0.2	1	1	1
Emotion	1.01 (0.08)	1–3	99.5	0§	1	1	1
Learning and remembering	1.02 (0.14)	1–3	98.5	0	1	1	1
Thinking and problem solving	1.02 (0.19)	1–4	98.0	0.1	1	1	1
Pain and discomfort	1.12 (0.33)	1–3	88.5	0	1	1	1
Disability score¶	10.69 (1.11)	10–26	54.7	0	10	10	11
<b>Additional domains</b>							
General health	1.05 (0.24)	1–4	95.8	0.1	1	1	1
Behaviour	1.05 (0.23)	1–4	95.1	0§	1	1	1

\*Percentage of respondents with the best possible score (ceiling).

†Percentage of respondents with the worst possible score (floor).

‡Median.

§<0.1% were observed at the maximum (floor).

¶Sum of the 10 original domains.

max, maximum; min, minimum; HSCS-PS, Health Status Classification System-Preschool.

### Non-response analysis

Significant differences were present in all characteristics, except for children's age, gender and presence of ENT surgical procedures and wheezing. Excluded children relatively more often had single parents ( $p<0.001$ ); parents with a low educational level ( $p<0.001$ ); and more often a non-Dutch parent ( $p<0.001$ ) (see online supplementary table S3).

### Feasibility

The response rate of the questionnaire at 36 month after birth was 69%.<sup>28</sup> Considering all questionnaires that were received at age 3 years (n=5014), there were on average 1.7% missing answers regarding the HSCS-PS items in the questionnaire; this was highest for 'Vision' (4.19%) and 'Hearing' (3.19%). Score distributions of the HSCS-PS domains and the total 'disability score' are presented in table 1. Floor effects were absent. Near to perfect scores (level 1=normal health/no impairment) were reported by >90.0% in 7 out of 10 domains; exceptions were 'Speech' (66.8%), 'Self-care' (89.4%) and 'Pain and discomfort' (88.5%). All HSCS-PS domains and the total 'disability score' showed a ceiling effect.

### Discriminant validity

Table 2 shows the ability of the total 'disability score' to discriminate between the subgroup of children born with a 'VLBW' (effect size 0.39;  $p<0.05$ ) and 'very preterm birth' (0.42;  $p<0.01$ ), and the 'reference' subgroup; the differences between the subgroups with 'moderate LBW' (effect size 0.17;  $p\geq 0.05$ ) and 'moderate to late

preterm' (0.20;  $p<0.001$ ) compared with the 'reference' subgroup were lower, as hypothesised. The domains 'Learning and remembering', 'Self-care' showed the largest discriminant validity regarding 'VLBW' (effect size 0.44;  $p<0.001$ ) and 'very preterm birth' (0.42;  $p<0.001$ ).

Table 3 shows the ability of the total 'disability score' to discriminate between the subgroup of children with '≥4 times wheezing in the previous year' (effect size 0.27;  $p<0.01$ ) and 'at least 1 ENT surgical procedure' (0.33;  $p<0.001$ ) and CBCL 'behaviour problems present' (0.52;  $p<0.001$ ), and the 'reference' subgroup, as hypothesised. In these three comparisons (table 4) the single 'original domains' that showed the largest discriminant validity were, respectively, 'Pain and discomfort' (effect size 0.31;  $p<0.001$ ), 'Pain and discomfort' (0.35;  $p<0.001$ ), 'Self-care' (0.40;  $p<0.001$ ).

The total 'disability score' discriminated clearly between the subgroup with a total of '≥3 conditions present' (0.47;  $p<0.001$ ), and the 'reference' subgroup, as hypothesised. The domain 'Pain and discomfort' (0.56;  $p<0.001$ ) showed the largest discriminant validity (online supplementary table S4).

### Concurrent validity

All five hypothesised correlation coefficients between the 'parallel' HSCS-PS-domains and 'General health'/'Behavior' were positive (0.21, 0.17, 0.16, 0.16, 0.18, respectively;  $p<0.01$ ). All 15 'non-hypothesised correlations' were lower than the hypothesised correlations (see table 4).

**Table 2** Discriminative ability of the HSCS-PS between subgroups differing in birth outcomes: birth weight (n=4541) and gestational age at birth (n=4526)†

HSCS-PS domains	Birth weight		Gestational age									
	≥2500 g (n=4307)		1500–2500 g (n=206)		<1500 g (n=28)		≥37 weeks (n=4256)		32–37 weeks (n=239)		<32 weeks (n=31)	
	Mean (SD)‡	Effect size§	Mean (SD)‡	Effect size§	Mean (SD)‡	Effect size§	Mean (SD)‡	Effect size‡‡	Mean (SD)‡	Effect size‡‡	Mean (SD)‡	Effect size‡‡
Original domains												
Vision	1.01 (0.16)	0.14	1.06 (0.37)**	0.16	1.07 (0.38)	0.16	1.01 (0.16)	0.14	1.06 (0.35)**	0.14	1.06 (0.36)	0.14
Hearing	1.02 (0.14)	0.00	1.02 (0.15)	0.11	1.04 (0.19)	0.11	1.02 (0.14)	0.00	1.02 (0.14)	0.00	1.03 (0.18)	0.06
Speech	1.34 (0.51)	0.02	1.35 (0.54)	0.16	1.43 (0.57)	0.16	1.34 (0.50)	0.14	1.42 (0.58)	0.14	1.48 (0.57)	0.25
Mobility	1.02 (0.17)	0.09	1.04 (0.22)	0.21	1.11 (0.42)*	0.21	1.02 (0.15)	0.12	1.06 (0.34)**	0.12	1.10 (0.40)*	0.20
Dexterity	1.01 (0.10)	0.09	1.03 (0.23)***	0.10	1.00 (0.00)	0.10	1.01 (0.09)	0.12	1.04 (0.25)***	0.12	1.00 (0.00)	0.11
Self-care	1.11 (0.34)	0.16	1.19 (0.50)*	0.42	1.46 (0.84)**	0.42	1.11 (0.34)	0.14	1.18 (0.50)*	0.14	1.45 (0.81)***	0.42
Emotion	1.01 (0.08)	0.13	1.00 (0.07)	0.13	1.00 (0.00)	0.13	1.01 (0.08)	0.00	1.01 (0.09)	0.00	1.00 (0.00)	0.13
Learning and remembering	1.01 (0.13)	0.14	1.05 (0.28)**	0.44	1.18 (0.39)***	0.44	1.01 (0.13)	0.15	1.05 (0.27)**	0.15	1.16 (0.37)***	0.41
Thinking and problem solving	1.02 (0.17)	0.16	1.07 (0.32)**	0.20	1.14 (0.59)*	0.20	1.02 (0.17)	0.12	1.06 (0.34)*	0.12	1.16 (0.58)**	0.24
Pain and discomfort	1.12 (0.33)	0.09	1.15 (0.35)	0.06	1.14 (0.36)	0.06	1.12 (0.33)	0.08	1.15 (0.36)	0.08	1.13 (0.34)	0.03
Disability score§§	10.67 (1.05)	0.17	10.98 (1.78)	0.39	11.57 (2.28)*	0.39	10.66 (1.03)	0.20	11.05 (1.92)***	0.20	11.58 (2.17)**	0.42
Additional domains												
General health	1.05 (0.23)	0.07	1.07 (0.29)	0.33	1.32 (0.82)***	0.33	1.04 (0.23)	0.17	1.10 (0.35)**	0.17	1.29 (0.78)***	0.32
Behaviour	1.05 (0.23)	0.08	1.07 (0.25)	0.19	1.11 (0.32)	0.19	1.05 (0.23)	0.04	1.06 (0.24)	0.04	1.10 (0.30)	0.17

Cohen's effect size (d): 0.2≤d<0.5 indicates a small difference, 0.5≤d<0.8 indicates a moderate difference, d≥0.8 indicates a large difference.

†Data are presented as mean or SD.

‡Two-sided Mann-Whitney U-test given a non-normal distribution of the data.

§Difference of the means divided by the largest SD between children born with a moderately low birth weight and children born with a normal birth weight.

||Difference of the means divided by the largest SD between children born with a very low birth weight and children born with a normal birth weight.

††Difference of the means divided by the largest SD between moderate to late preterm children and term children.

‡‡Difference of the means divided by the largest SD between very preterm children and term children.

§§ Sum of the 10 original domains.

\*P<0.05;

\*\*P<0.01;

\*\*\*P<0.001.

HSCS-PS, Health Status Classification System-Preschool.

**Table 3** Discriminative ability of the HSCS-PS between subgroups differing in: wheezing chest in the previous year (n=4407), ear-nose-throat (ENT) surgical procedures ever (n=4346), Child Behavior Checklist behaviour problem present in the last 2 months (n=4490)†

HSCS-PS domains	Wheezing in the previous year				ENT surgical procedures				Behaviour problems present in the last 2 months			
	No wheezing (n=3849)		1-3 times (n=454)		No ENT surgical procedures (n=3956)		At least 1 ENT surgical procedures (n=390)		No behaviour problems present (n=4235)		Behaviour problems present (n=255)	
	Mean (SD)‡	Effect size§	Mean (SD)‡	Effect size‡	Mean (SD)‡	Effect size‡	Mean (SD)‡	Effect size‡	Mean (SD)‡	Effect size‡	Mean (SD)‡	Effect size‡
Original domains												
Vision	1.02 (0.18)	0.00	1.02 (0.16)	0.09	1.02 (0.18)	0.06	1.01 (0.10)	0.06	1.02 (0.17)	0.07	1.04 (0.27)	0.07
Hearing	1.02 (0.14)	0.06	1.03 (0.17)	0.14	1.01 (0.12)	0.21	1.07 (0.28)***	0.21	1.02 (0.13)	0.12	1.05 (0.25)**	0.12
Speech	1.34 (0.50)	0.06	1.37 (0.54)	0.05	1.33 (0.49)	0.25	1.48 (0.60)***	0.25	1.33 (0.50)	0.38	1.57 (0.63)***	0.38
Mobility	1.02 (0.15)	0.11	1.05 (0.27)***	0.09	1.02 (0.14)	0.17	1.08 (0.36)***	0.17	1.02 (0.15)	0.19	1.09 (0.36)***	0.19
Dexterity	1.00 (0.08)	0.11	1.02 (0.19)*	0.10	1.01 (0.09)	0.06	1.02 (0.18)	0.06	1.00 (0.08)	0.19	1.06 (0.31)***	0.19
Self-care	1.11 (0.34)	0.13	1.17 (0.47)*	0.18	1.11 (0.34)	0.11	1.16 (0.46)	0.11	1.10 (0.33)	0.40	1.35 (0.62)***	0.40
Emotion	1.00 (0.07)	0.08	1.01 (0.12)	0.00	1.00 (0.07)	0.17	1.02 (0.12)**	0.17	1.00 (0.05)	0.22	1.05 (0.23)***	0.22
Learning and remembering	1.01 (0.13)	0.05	1.02 (0.19)	0.19	1.01 (0.12)	0.21	1.07 (0.28)***	0.21	1.01 (0.11)	0.31	1.13 (0.39)***	0.31
Thinking and problem solving	1.02 (0.17)	0.07	1.04 (0.27)	0.06	1.02 (0.17)	0.11	1.05 (0.27)**	0.11	1.02 (0.14)	0.29	1.15 (0.45)***	0.29
Pain and discomfort	1.11 (0.31)	0.17	1.18 (0.41)***	0.31	1.10 (0.31)	0.35	1.27 (0.48)***	0.35	1.11 (0.32)	0.31	1.25 (0.45)***	0.31
Disability score§§	10.65 (1.01)	0.15	10.90 (1.64)***	0.27	10.62 (0.98)	0.33	11.22 (1.84)***	0.33	10.62 (0.97)	0.52	11.75 (2.17)***	0.52
Additional domains												
General health	1.03 (0.18)	0.25	1.13 (0.40)***	0.65	1.03 (0.19)	0.34	1.20 (0.50)***	0.34	1.04 (0.22)	0.25	1.15 (0.44)***	0.25
Behaviour	1.05 (0.23)	0.04	1.06 (0.23)	0.20	1.05 (0.22)	0.10	1.08 (0.30)*	0.10	1.03 (0.18)	0.56	1.35 (0.57)***	0.56

Cohen's effect size (d): 0.2≤d<0.5 indicates a small difference, 0.5≤d<0.8 indicates a moderate difference, d≥0.8 indicates a large difference.

†Data are presented as mean or SD.

‡Two-sided Mann-Whitney U-test given a non-normal distribution of the data.

§Difference of the means divided by the largest SD between children with 1-3 times of wheezing in the past 1 year and children with no wheezing.

¶Difference of the means divided by the largest SD between children with at least 4 times of wheezing in the past 1 year and children with no wheezing.

‡‡Difference of the means divided by the largest SD between children had at least one ENT surgical procedures and children had no ENT surgical procedures.

‡‡‡Difference of the means divided by the largest SD between children had behaviour problems and children had no behaviour problems.

§§Sum of the 10 original domains.

\*P<0.05; \*\*P<0.01; \*\*\*P<0.001.

HSCS-PS, Health Status Classification System-Preschool.

**Table 4** Concurrent validity of the HSCS-PS assessed by Spearman correlations between original HSCS-PS domains and two additional domains (n=4546)†

HSCS-PS domains	General health	Behaviour
Vision	0.04 (–0.003 to 0.094)	0.02 (–0.011 to 0.060)
Hearing	0.09 * (0.039 to 0.143)	0.04 (–0.002 to 0.085)
Speech	0.08* (0.047 to 0.111)	0.09* (0.059 to 0.126)
Mobility	0.13* (0.066 to 0.187)	0.06* (0.015 to 0.106)
Dexterity	0.11* (0.038 to 0.178)	0.07* (0.013 to 0.129)
Self-care	0.09* (0.051 to 0.129)	0.17* (0.129 to 0.218)
Emotion	0.00 (–0.016 to 0.033)	0.16* (0.088 to 0.221)
Learning and remembering	0.11* (0.051 to 0.178)	0.16* (0.091 to 0.228)
Thinking and problem solving	0.11* (0.056 to 0.176)	0.18* (0.123 to 0.245)
Pain and discomfort	0.21* (0.161 to 0.250)	0.08* (0.040 to 0.109)

Values presented in this table are values of Spearman correlation coefficient (CC) and 95% CI of Spearman's CC.

† Correlations with predefined related general health/behaviour are in italics; other (spurious) are in standard font.

\* When (a) 95% CI is not 'across 0'; and (b) p value <0.05, the correlation coefficient was regarded as statistically significant.

HSCS-PS, Health Status Classification System-Preschool.

## DISCUSSION

The present study evaluated the HSCS-PS among children at 3 years of age in a community setting using a large general population sample. The results support the feasibility, discriminant validity and concurrent validity of the HSCS-PS.

### Feasibility

The HSCS-PS was well accepted by parents, as shown by the high response and relatively few missing answers. All levels of the potential answer categories were observed for 5 of the 10 domains in this community sample: 'Speech'; 'Mobility'; 'Self-care'; 'Emotion'; 'Thinking and problem solving'. Yet, the study showed considerable ceiling effects, specifically regarding 'Dexterity' and 'Emotion'. Such ceiling effects are a common phenomenon in community samples with a generally healthy population; they were also observed in studies with other HRQOL measures.<sup>6,10</sup> However, they may limit the use of the instrument to detect changes and to describe health beyond the average. The domain 'Speech' showed the highest variation in the obtained scores, which may be related to the individual differences in the development of children's speech skills.<sup>38</sup>

### Discriminant validity

The results support the ability of HSCS-PS to discriminate across subgroups characterised by absence/presence of adverse perinatal conditions or distinct chronic or medical conditions. The total 'disability score' showed consistent differences between the subgroups with no adverse condition/situation (the reference subgroup), and subgroups with a 'mild condition/situation' (if present), and subgroups with a 'severe condition/situation', concerning six outcomes (birth weight; gestational age at birth; wheezing in the previous year; ENT surgical procedures; behaviour problem present in the last

2 months; total number of chronic/medical conditions), as hypothesised.

Regarding 'VLBW', the highest effect sizes were found in 'Self-care' and 'Learning and remembering'. A similar pattern was observed regarding 'very preterm birth'. Previous studies showed that perinatal adversity may be associated with neurodevelopmental disabilities that may cause cognitive impairment and attention problems during child development.<sup>39</sup> The results of our study are also supported by Msall and Tremont who measured functional outcomes in self-care in infants with VLBW: of children with and without neurodevelopmental impairment, 41% and 13%, respectively, had self-care limitations.<sup>40</sup>

Regarding wheezing frequency, the highest effect sizes were found regarding 'Pain and discomfort' between children with at least four times of wheezing in the past year and children with no wheezing. This is in accordance with earlier reports that wheezing is associated with low HRQOL, especially in the domain of bodily pain.<sup>41</sup> Similarly, regarding ENT surgical procedures, the highest effect size was found regarding 'Pain and discomfort', which is consistent with studies on pain after ENT procedures.<sup>42</sup>

Regarding behaviour problems, the highest effect sizes were found regarding 'Self-care'. This is in line with the results of previous studies that documented the impact of behaviour problems on self-care.<sup>43</sup>

### Concurrent validity

The finding that the five hypothesised correlations between the HSCS-PS domains and respective parallel single-item parent reports on 'General health'/'Behavior' were higher than the 15 'non-hypothesised correlations' supports the concurrent validity of the HSCS-PS. The strength of the five hypothesised correlations varied. It should be noted that 11 of the 15 'non-hypothesised

correlations' were statistically significant, although they were relatively small. The correlation between 'Pain and discomfort' and 'General health' was the highest which is consistent with previous studies; pain may play a major role in the rating of quality of life.<sup>44</sup> The relatively strong correlation between HSCS-PS 'Thinking and problem solving' and 'Behaviour' is consistent with previous reports regarding associations between cognition deficits in children and behaviour problems.<sup>45</sup> The observed associations to assess concurrent validity in this community sample are slightly lower than those that were found by Saigal *et al* who used selected clinical cohorts. This may be explained by the relatively low prevalence of serious impairments in a general population sample such as in our study.<sup>11</sup>

It should be noted that relatively little is known about the acceptance and validity of parent-report single-items to describe 'General health' and 'Behavior/Mental health' of children compared with the body of knowledge regarding the validity of such measures in adult populations.<sup>26 27</sup> Therefore, in the future, we recommend the concurrent validity of the HSCS-PS should be evaluated by comparing it with an accepted 'gold standard' HRQOL measure such as the ITQOL.<sup>6</sup> The evaluation of the concurrent validity of the 10-domains HSCS-PS in this study is a first step and results should be interpreted with caution.

### Methodological considerations

First, in this study, measurements were primarily done using parent questionnaires, including accepted validated instruments such as the CBCL parent questionnaire.<sup>35</sup> Only the birth outcomes were obtained from medical files. 'Reporting tendency' by, for example, 'optimistic' or 'pessimistic' parents may have applied to all measures in the questionnaires and may have induced relatively high statistical associations in this study. For future validation studies we recommend to use as many as possible 'objective' external measures to validate the 10-domains HSCS-PS.

Second, no formal power calculations were made with regard to the validation study, given multiple comparisons and studies of associations. However, the size of the population for analysis (n=4526) is relatively large for a validation study; therefore many associations, even with a small effect size, were statistically significant. The smallest subgroups regarding the evaluation of discriminative validity (birth weight <1500 g, n=28; gestational age <32 weeks, n=31) resulted in almost half of the comparisons being statistically significant. All other subgroups regarding the evaluation of discriminative validity ranged from n=104 up to n=4307.

Third, in our study, the non-participants were children from vulnerable families, who more often had single parent, and whose parents more often had lower educational level or had an immigrant background. These children may have more health conditions/problems than their counterparts from non-vulnerable families. This

issue may impose an impact on results. For instance, the high ceiling effect may be caused by the relatively better health status of the participants. In addition, the generalisability of results in the present study may be limited due to this issue.

Fourth, while a utility-based scoring algorithm for HSCS-PS has not yet been developed, a total 'disability score' summing up the scores regarding each of the 10 original domains was applied in this study.<sup>16</sup> Two previous studies supported the feasibility and validity of the HSCS-PS total 'disability score' in absence of a utility-based scoring algorithm, which we recommend to be developed in future studies.<sup>15 16 46</sup> Given the relative paucity of experience with the HSCS-PS system, no specific guidelines for clinically important differences are available; we recommend such guidelines to be developed. Regarding the HUI for patients aged 4 years and above, it was proposed that a difference of one level within any domain may be interpreted as a clinically important difference.<sup>12</sup> In our case, for example, the subgroup with CBCL 'behaviour problems present' and the subgroup with '≥3 chronic/medical conditions' have both a mean total 'disability score' that is more than one point (one level) higher compared with the reference group, which may be interpreted as a clinically important difference. From a statistical point of view, we propose to apply Cohen's effect size (d), and to interpret 0.50 (half an SD) as a meaningful difference. Effect sizes were relatively small in this study, which reflects that the general population in a society with modern and accessible healthcare is relatively healthy.<sup>37 47</sup>

Fifth, we would like to note that regarding the procedure of developing the HSCS-PS, items were mainly derived from the HUI system and additionally two new items were based on experts' opinion. Qualitative studies, such as using focus group interviews have not been mentioned in this procedure; we recommend that qualitative research may be applied in the future, for example, to reduce the number of items, or to evaluate the content of the items.

Finally, in the present study, indicators of the reliability of the HSCS-PS, such as test-retest reliability were not evaluated. We recommend assessing this in future studies in the large varied community population.

### CONCLUSION

This study is the first to apply and to evaluate the HSCS-PS in a large community sample of preschool children. This is a relevant addition to previous studies among VLBW children and children with cerebral palsy. For the assessment of the validity, we applied objectively measured conditions (birth weight, gestational age at birth) in addition to validated parent-reported outcome measures (CBCL). This study supports the feasibility and validity of the HSCS-PS among preschool children in community settings. We recommend developing utility-based scoring algorithms for the HSCS-PS, and conducting empirical studies of what changes are meaningful, as



well as repeated studies of reliability and validity in large varied populations with objectively measured, external benchmarks. In the meantime, the HSCS-PS may be used by clinicians and researchers as parent-reported health outcome in addition to clinical outcomes for economic evaluations, and may be used to support the development of value-based healthcare regarding interventions for preschool children.

#### Author affiliations

<sup>1</sup>Shanghai Medical Information Center, Shanghai, China

<sup>2</sup>Shanghai Health Development Research Center, Shanghai, China

<sup>3</sup>Department of Public Health, Erasmus MC - University Medical Center Rotterdam, Rotterdam, The Netherlands

<sup>4</sup>The Generation R Group, Erasmus MC - University Medical Center Rotterdam, Rotterdam, The Netherlands

<sup>5</sup>Department of Economics and Centre for Health Economics and Policy Analysis, McMaster University, Hamilton, Ontario, Canada

<sup>6</sup>Department of Pediatrics, McMaster University, Hamilton, Ontario, Canada

<sup>7</sup>Division of Respiratory Medicine and Allergology, Department of Pediatrics, Erasmus MC - University Medical Center Rotterdam, Rotterdam, The Netherlands

<sup>8</sup>Division of Neonatology, Department of Pediatrics, Erasmus MC - University Medical Center Rotterdam, Rotterdam, The Netherlands

<sup>9</sup>Department of Pediatrics, Erasmus MC - University Medical Center Rotterdam, Rotterdam, The Netherlands

<sup>10</sup>Department of Epidemiology, Erasmus MC - University Medical Center Rotterdam, Rotterdam, The Netherlands

**Acknowledgements** The Generation R Study was conducted by the Erasmus Medical Center, Rotterdam, the Netherlands, in close collaboration with the School of Law and Faculty of Social Sciences of the Erasmus University, Rotterdam; the Municipal Health Service, Rotterdam area; the Rotterdam Homecare Foundation; and the Stichting Trombosedienst & Artsenlaboratorium Rijnmond (STAR), Rotterdam. We gratefully acknowledge the contribution of general practitioners, hospitals, midwives and pharmacies in Rotterdam, and all of the women participating in the present study. We also acknowledge Henning W. Tiemeier who collected data of CBCL in the Generation R Study.

**Contributors** XF conceptualised and designed the study, carried out the analyses, interpreted the data, drafted the initial manuscript, critically reviewed and revised the manuscript. GB conceptualised and designed the study, carried out the analyses, interpreted the data, critically reviewed and revised the manuscript. DAW, DF and SS conceptualised and designed the study, interpreted the data, critically reviewed and revised the manuscript. SH and CJ interpreted the data, critically reviewed and revised the manuscript. LD and VVVJ aided in the collection of data, interpreted the data, critically reviewed and revised the manuscript. HR initiated the study, conceptualised and designed the study, collected the data, interpreted the data, and critically reviewed and revised the manuscript. All authors approved the final manuscript as submitted and agree to be accountable for all aspects of the work.

**Funding** XF is supported by a fellowship from the China Medical Board-Collaborating Program in Evidence-based Health Policy-making (grant number: CMB-CP14-190, <https://www.chinamedicalboard.org/>). GB is supported by a China Scholarship Council (CSC) PhD Fellowship for her PhD study in Erasmus MC, Rotterdam, the Netherlands. The scholarship file number is 201406010271, CSC URL: [<http://www.csc.edu.cn/>].

**Competing interests** None declared.

**Patient consent** Not required.

**Ethics approval** Medical Ethics Committee of the Erasmus University Medical Center, Rotterdam (MEC-2007-413).

**Provenance and peer review** Not commissioned; externally peer reviewed.

**Data sharing statement** The data (actual SPSS data) on which our study is based are derived from a large birth cohort in Rotterdam (also known as the Generation R Study). Individual researchers do not have the right to distribute data from the Generation R Study. For the purpose of verification/validation/replication/meta-analyses, the external researcher can contact our data managers (Claudia J. Kruitthof, [c.kruitthof@erasmusmc.nl](mailto:c.kruitthof@erasmusmc.nl) or [datamanagementgen@erasmusmc.nl](mailto:datamanagementgen@erasmusmc.nl)) and

contact the Director Generation R, Vincent Jaddoe ([v.jaddoe@erasmusmc.nl](mailto:v.jaddoe@erasmusmc.nl)). Postal address: Generation R, Postbox 2040, 3000 CA Rotterdam, the Netherlands. Data will be made available via these contact persons after a written agreement about the use of the data has been made via the Technology Transfer Office of Erasmus MC.

**Open access** This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

#### REFERENCES

- Raat H, Mohangoo AD, Grootenhuys MA. Pediatric health-related quality of life questionnaires in clinical trials. *Curr Opin Allergy Clin Immunol* 2006;6:180–5.
- Black N. Patient reported outcome measures could help transform healthcare. *BMJ* 2013;346:f167.
- Kaplan RM. *Quality of life measures: measurement strategies in health psychology*. New York: John Wiley, 1985.
- Patrick DL, Erickson P. Health status and health policy: quality of life in health care evaluation and resource allocation.. 1993.
- Theunissen NC, Vogels TG, Koopman HM, *et al*. The proxy problem: child report versus parent report in health-related quality of life research. *Qual Life Res* 1998;7:387–97.
- Raat H, Landgraf JM, Oostenbrink R, *et al*. Reliability and validity of the Infant and Toddler Quality of Life Questionnaire (ITQOL) in a general population and respiratory disease sample. *Qual Life Res* 2007;16:445–60.
- Porter ME. What is value in health care? *N Engl J Med* 2010;363:2477–81.
- Krabbe P. *The measurement of health and health status: concepts, methods and applications from a multidisciplinary perspective*: Academic Press, 2016.
- Krabbe PF. A generalized measurement model to quantify health: the multi-attribute preference response model. *PLoS One* 2013;8:e79494.
- 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.
- Saigal S, Rosenbaum P, Stoskopf B, *et al*. Development, reliability and validity of a new measure of overall health for pre-school children. *Qual Life Res* 2005;14:243–52.
- Furlong WJ, Feeny DH, Torrance GW, *et al*. The Health Utilities Index (HUI) system for assessing health-related quality of life in clinical studies. *Ann Med* 2001;33:375–84.
- Msall ME. Neurodevelopmental surveillance in the first 2 years after extremely preterm birth: evidence, challenges, and guidelines. *Early Hum Dev* 2006;82:157–66.
- Klassen AF, Lee SK, Raina P, *et al*. Health status and health-related quality of life in a population-based sample of neonatal intensive care unit graduates. *Pediatrics* 2004;113:594–600.
- Nathan PC, Furlong W, Horsman J, *et al*. Inter-observer agreement of a comprehensive health status classification system for pre-school children among patients with Wilms' tumor or advanced neuroblastoma. *Qual Life Res* 2004;13:1707–14.
- Nathan PC, Furlong W, De Pauw S, *et al*. Health status of young children during therapy for advanced neuroblastoma. *Pediatr Blood Cancer* 2004;43:659–67.
- Vederhus BJ, Eide GE, Natvig GK, *et al*. Health-related quality of life and emotional and behavioral difficulties after extreme preterm birth: developmental trajectories. *PeerJ* 2015;3:e738.
- Vederhus BJ, Markestad T, Eide GE, *et al*. Health related quality of life after extremely preterm birth: a matched controlled cohort study. *Health Qual Life Outcomes* 2010;8:53.
- Theunissen NC, Veen S, Fekkes M, *et al*. Quality of life in preschool children born preterm. *Dev Med Child Neurol* 2001;43:460–5.
- Chien LY, Chou YH, Ko YL, Yi K, *et al*. Health-related quality of life among 3–4-year-old children born with very low birthweight. *J Adv Nurs* 2006;56:9–16.
- Schiariti V, Houbè JS, Lisonkova S, *et al*. Caregiver-reported health outcomes of preschool children born at 28 to 32 weeks' gestation. *J Dev Behav Pediatr* 2007;28:9–15.
- Hafkamp-de Groen E, Mohangoo AD, Landgraf JM, *et al*. The impact of preschool wheezing patterns on health-related quality of life at age 4 years. *Eur Respir J* 2013;41:952–9.

23. Stewart MG, Friedman EM, Sulek M, *et al.* Quality of life and health status in pediatric tonsil and adenoid disease. *Arch Otolaryngol Head Neck Surg* 2000;126:45–8.
24. Richards M, Giannoni C. Quality-of-life outcomes after surgical intervention for otitis media. *Arch Otolaryngol Head Neck Surg* 2002;128:776–82.
25. Charach A, McLennan JD, Bélanger SA, *et al.* Screening for disruptive behaviour problems in preschool children in primary health care settings. *J Can Acad Child Adolesc Psychiatry* 2017;26:172–8.
26. Macias C, Gold PB, Öngür D, *et al.* Are single-item global ratings useful for assessing health status? *J Clin Psychol Med Settings* 2015;25:1–64.
27. Ahmad F, Jhaji AK, Stewart DE, *et al.* Single item measures of self-rated mental health: a scoping review. *BMC Health Serv Res* 2014;14:398.
28. Jaddoe VW, van Duijn CM, Franco OH, *et al.* The Generation R Study: design and cohort update 2012. *Eur J Epidemiol* 2012;27:739–56.
29. Vohr BR, Wright LL, Dusick AM, *et al.* Neurodevelopmental and functional outcomes of extremely low birth weight infants in the National Institute of Child Health and Human Development Neonatal Research Network, 1993–1994. *Pediatrics* 2000;105:1216–26.
30. Bhutta AT, Cleves MA, Casey PH, *et al.* Cognitive and behavioral outcomes of school-aged children who were born preterm: a meta-analysis. *JAMA* 2002;288:728–37.
31. Wardlaw TM. *Low birthweight: country, regional and global estimates*: UNICEF, 2004.
32. Timmermans S, Jaddoe VW, Hofman A, *et al.* Periconception folic acid supplementation, fetal growth and the risks of low birth weight and preterm birth: the Generation R Study. *Br J Nutr* 2009;102:777–85.
33. Solé D, Vanna AT, Yamada E, *et al.* International Study of Asthma and Allergies in Childhood (ISAAC) written questionnaire: validation of the asthma component among Brazilian children. *J Investig Allergol Clin Immunol* 1998;8:376–82.
34. Mamie C, Habre W, Delhumeau C, *et al.* Incidence and risk factors of perioperative respiratory adverse events in children undergoing elective surgery. *Paediatr Anaesth* 2004;14:218–24.
35. Achenbach TM, Rescorla LA. *Manual for the ASEBA preschool forms and profiles*. Burlington, 2010.
36. Tick NT, van der Ende J, Koot HM, Verhulst FC, *et al.* 14-year changes in emotional and behavioral problems of very young Dutch children. *J Am Acad Child Adolesc Psychiatry* 2007;46:1333–40.
37. Cohen J. Statistical power analysis. Current directions in psychological science. 1992;1:98–101.
38. Moeller MP. Early intervention and language development in children who are deaf and hard of hearing. *Pediatrics* 2000;106:e43.
39. Butler AS, Behrman RE. *Preterm birth: causes, consequences, and prevention*: National Academies Press, 2007.
40. Msall ME, Tremont MR. Functional outcomes in self-care, mobility, communication, and learning in extremely low-birth weight infants. *Clin Perinatol* 2000;27:381–401.
41. Mohangoo AD, de Koning HJ, Mangunkusumo RT, *et al.* Health-related quality of life in adolescents with wheezing attacks. *J Adolesc Health* 2007;41:464–71.
42. Turan A, Emet S, Karamanlioğlu B, *et al.* Analgesic effects of rofecoxib in ear-nose-throat surgery. *Anesth Analg* 2002;95:1308–11.
43. Adams H, de Bleeck EA, Mink JW, *et al.* Standardized assessment of behavior and adaptive living skills in juvenile neuronal ceroid lipofuscinosis. *Dev Med Child Neurol* 2006;48:259–64.
44. Arnstein P. The mediation of disability by self efficacy in different samples of chronic pain patients. *Disabil Rehabil* 2000;22:794–801.
45. Dodge KA. Social cognition and children's aggressive behavior. *Child Dev* 1980;51:162–70.
46. Raat H, Bonsel GJ, Hoogveen WC, *et al.* Feasibility and reliability of a mailed questionnaire to obtain visual analogue scale valuations for health states defined by the Health Utilities Index Mark 3. *Med Care* 2004;42:13–18.
47. Norman GR, Sloan JA, Wyrwich KW. Interpretation of changes in health-related quality of life: the remarkable universality of half a standard deviation. *Med Care* 2003;41:582–92.