



# Short form development for oral health patient-reported outcome evaluation in children and adolescents

Honghu Liu<sup>1,2,3</sup>  · Ron Hays<sup>2</sup> · Yan Wang<sup>3</sup> · Marvin Marcus<sup>1</sup> · Carl Maida<sup>1</sup> · Jie Shen<sup>1</sup> · Di Xiong<sup>3</sup> · Steve Lee<sup>1</sup> · Vladimir Spolsky<sup>1</sup> · Ian Coulter<sup>1</sup> · James Crall<sup>1</sup>

Accepted: 20 February 2018 / Published online: 5 March 2018

© This is a U.S. Government work and not under copyright protection in the US; foreign copyright protection may apply 2018

## Abstract

**Purpose** Children and adolescents are vulnerable to dental problems and oral diseases. This paper presents the development of two multi-item self-report scales for use in assessing oral health status of children and adolescents.

**Methods** Following the Patient-Reported Outcome Measurement Information System framework, survey questions were designed using a newly developed conceptual model. These items were administered to 334 children and adolescents (8–17 years) along with concurrent dental exams. Exploratory and confirmatory factor analyses were conducted and the item response theory graded response model was used to estimate item parameters and oral health status scores and to identify short-form items. The items were selected by high level of information and wide coverage of different domains to assess Child Oral Health Status Index (COHSI) and treatment referral recommendations (RR).

**Results** The long form consists of 28 items. The short-form includes 12 items (8 for COHSI and 7 for RR with 3 common items). The intra-class correlations between long form and short-form were 0.90 for COHSI and 0.87 for RR.

**Conclusion** The short-forms provide a possible solution for the longstanding challenge of oral health evaluation for large populations of children and adolescents. The calibrated long form provides the foundation for computer adaptive test administration. These oral health assessment toolkits can be used for oral health screening, surveillance program, policy planning, and research.

**Keywords** Oral health · Children · Patient-reported outcomes · Short-forms

## Introduction

Oral health is an essential component of daily functioning and well-being. Healthy People 2020 [1] emphasizes the importance of prevention and control of oral and craniofacial diseases, conditions, and injuries, and of enhancing access to preventive services and dental care. Children and adolescents are in the transitional stage for development and often have diets with high levels of sugar in United States and

globally [2–4], making them vulnerable to dental problems and oral diseases [5].

It is challenging to develop oral health measures for cost effectively evaluating large populations for their current oral health status and the treatment needs, especially for children and adolescents. One of the most commonly used oral health surveys is the child version of Oral Impact on Daily Performances (Child-OIDP) [6–9], based on social dental approach for the consequences of oral health diseases [10]. The Child-OIDP measures the impact of oral health on children's life in a Europe population, and it was highly concentrated on the impact of "eating, smiling and speaking" and the impact of the oral health condition on quality of life [10–12]. A number of other questionnaires for measuring oral health-related quality of life (OHRQoL) of children have been developed [13], but none of them provide treatment referral recommendations. For example, the Child Perceptions Questionnaire (CPQ) [14–16] mainly links the frequency of events in the past 3 months to oral health conditions. The Child

✉ Honghu Liu  
hhliu@dentistry.ucla.edu

<sup>1</sup> Public Health & Community Dentistry, School of Dentistry, University of California, Los Angeles (UCLA), Los Angeles, CA, USA

<sup>2</sup> UCLA Department of Medicine, David Geffen School of Medicine, Los Angeles, CA, USA

<sup>3</sup> UCLA Department of Biostatistics, Los Angeles, CA, USA

Oral Health Impact Profile (COHIP) [17, 18] centers on measuring social impact of oral disorders. The Early Child Oral Health Impact Scale (ECOHIS) [19] primarily focuses on the negative influence of dental disease and treatment experience in very young children. Other surveys, such as the Scale of Oral Health Outcomes for 5-year-old children (SOHO-5) [20, 21], the Michigan Oral Health-Related Quality of Life scale (MOHRQoL) [22], the PedsQL Oral Health Scale [23, 24], and the Pediatric Oral Health-Related Quality of Life Measure (POQL) [25], do not identify the levels of treatment need for children. In short, these existing surveys above were created without comprehensive psychometric analysis. Surveys that can measure the current oral health status and estimate treatment needs, and can be used to evaluate the large population efficiently and effectively, are in need for children and adolescents. But these oral health surveys should be developed through rigorous psychometric analysis encompassing global health, physical, mental, and social health components to establish their reliability and validity.

High-quality patient-reported measures are essential for patient-centered care [26–28]. The Patient-Reported Outcome Measurement Information System (PROMIS), a United States National Institutes of Health roadmap initiative, began in 2004 to create self-reported health measures using state-of-the-science qualitative and quantitative methods [29, 30]. The PROMIS framework conceptualizes health as consisting of physical, mental, and social components [29, 31, 32], following the well-recognized World Health Organization model [29, 33]. Each component is then classified into subcomponents, such as function and symptoms in physical health; affect, behavior, and cognition in mental health; and relationships and functions in social health [29, 32, 34]. The subcomponents are further expanded into domains and sub-domains with respect to a particular concept [32]. These domains are intended to be mutually exclusive and collectively exhaustive.

We have developed self-report items assessing oral health by children and adolescents using the PROMIS framework and methodology [35, 36]. Specifically, we employed four operational phases to develop the oral health items [35, 36]: (i) systematic review of the literature to identify instruments and items associated with oral health; (ii) focus groups with racially/ethnically and geographically diverse children, parents, and oral health professionals to help conceptualize oral health [35, 36]; (iii) draft oral health items administered in a series of face-to-face cognitive interviews to ensure that subjects understand the intent of each question; and (iv) field test to evaluate the resulting oral health items.

The aim for this paper is to develop a static fixed-length short-form that measures current oral health status and provides recommendations for different levels of treatment needs, which could eventually serve as a screening tool for

efficiently and effectively assessing oral health of children and adolescents. This paper describes the calibration of this static oral health short-form that can be used to provide a reduced list of survey questions along with a scoring table, and to enable score mapping to clinical categories (e.g., referral treatment categories) guided by our oral health conceptual model (see Fig. 1).

## Methods

### Conceptual model

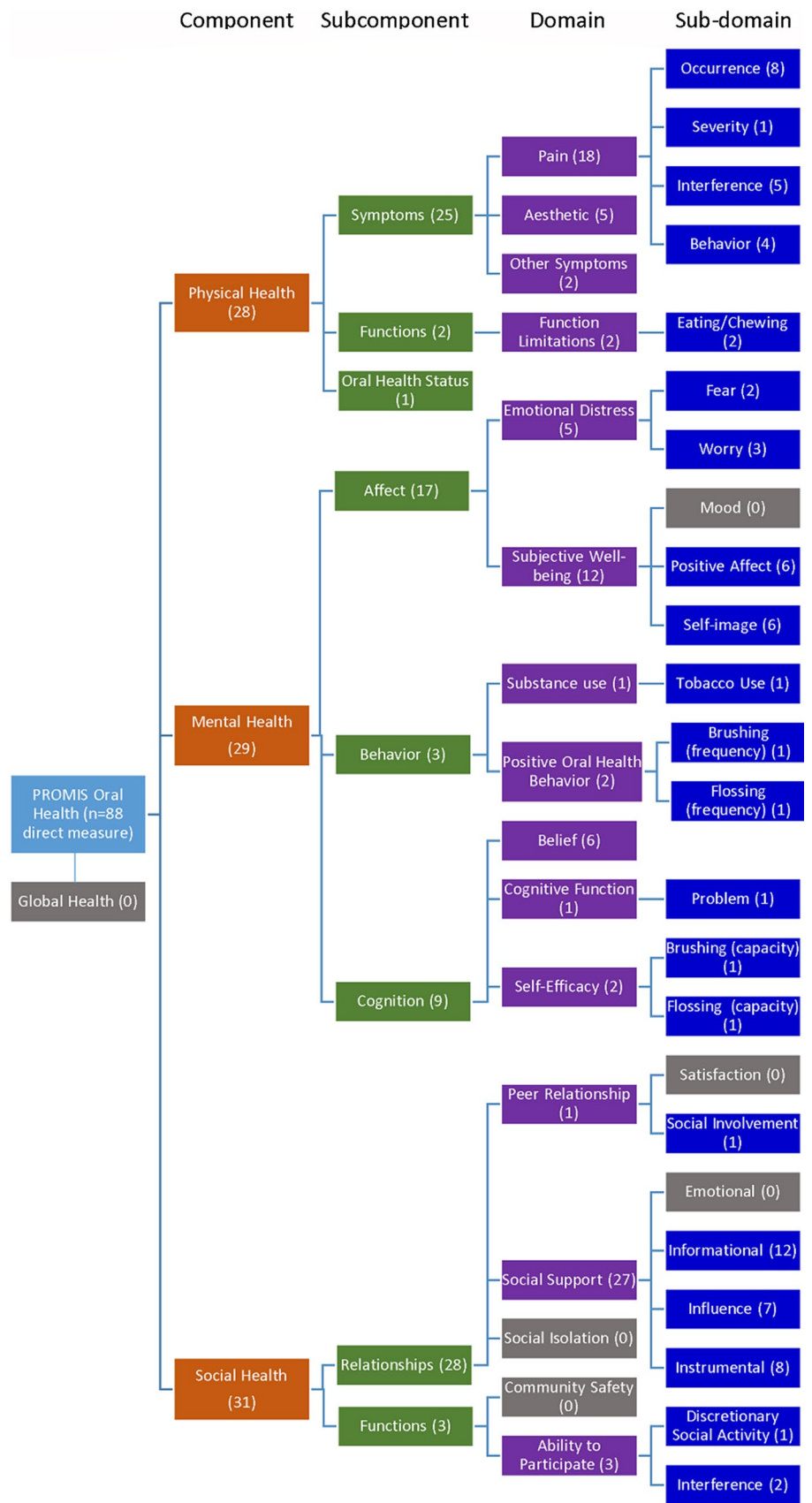
Guided by the PROMIS framework, the oral health conceptual model (Fig. 1) was developed by our expert panel of pediatric dentists, general dentists, social scientists, and PROMIS experts. The conceptual model is divided into three main components: physical, mental, and social health. Each component is further extended to subcomponent, domain, and sub-domain. The colors indicate different levels of the structure, from left to right: component (orange), subcomponent (green), domain (purple), and sub-domain (blue). The gray colored domains and sub-domains were not included in the analysis because they indirectly measure oral health status.

### Dental examination and referral recommendations

Data were obtained from (1) dental exams performed by calibrated dentists [35] (i.e., dentists who use the same standards in the exams) and (2) survey questions answered by children (age 8–17) during field testing in *dental clinics* located in Los Angeles County from August 2015 to October 2016. Details of the study, such as the development of oral health items and the study flowchart, are described elsewhere [35, 36]. Patients (children age 8–17) without orthodontic appliances were recruited with their parents or guardians from participating dental clinics located in Los Angeles County for on-site dental exams and administration of survey questions. Only one child was chosen from each family to form a stratified systematic convenient sample, which aimed at reflecting the race/ethnicity distribution of the general U.S. population [37].

The clinical exams were performed and surveys were administered in the same clinic setting. The data collected at dental clinics in the field test were used to evaluate and calibrate the oral health items by IRT modeling, and to refine the item pool. The oral health items were administered to children and parents by computer using Questionnaire Development System™ (QDS™), developed by Nova Research Company [38]. Children and their parents completed the surveys separately at the clinics. In this paper, we only focused on the oral health items for children.

**Fig. 1** The domain structure of child oral health conceptual model based on PROMIS framework



Using professional dental chairs, the tooth-based dental exam was performed by two experienced dentists who were part of the research team. The dental exam yielded the clinical outcome measures, including Children's Oral Health Status Index (COHSI) and treatment referral recommendations (RRs). COHSI assesses clinical oral health and reflects diseases (missing teeth, decay, and abnormal position) and occlusal status, estimating overall oral health status. The COHSI has a best possible score of 100 and decrements from that are estimated by multiplying previously derived regression coefficients by number of missing teeth ( $-2.27$  for primary anterior teeth,  $-4.55$  for primary posterior teeth and permanent teeth), number of decayed teeth ( $-1.12$  for primary anterior teeth,  $-2.24$  for primary posterior teeth and permanent teeth), occlusion status ( $-4.38$ ), and abnormal positions ( $-1.73$ ) (see [32] for details). The worst possible COHSI score is  $-27.4$  for adolescents with all permanent teeth missing and  $18.16$  for younger children with all primary teeth missing. The regression coefficients, which are derived from a paired preference approach [39], are higher for posterior teeth and permanent teeth that are decayed or missing. The observed range of the COHSI in the field test was  $59.18$ – $100$ . The RR is based on cavities and gingival bleeding and has four categories: see a dentist (i) immediately, (ii) within 2 weeks, (iii) at earliest convenience, or (iv) with regular schedule [40, 41].

RRs employ the 4-level guidelines used in the National Health and Nutrition Examination Survey (NHANES) [42, 43]. Because the NHANES 2009–2014 survey revealed few emergency dental conditions among those who had a dental visit within the past year, we revised the referral conditions to include urgent as well as emergency conditions. In addition to NHANES guidelines, the following criteria were used for RRs: *level 1 emergency dental condition*—if any permanent tooth has decay with cavitation or more than three primary teeth have decay, see dentist immediately; *level 2 urgent condition*—if a child has 1–2 primary teeth decayed, or more than half of teeth (12+) have bleeding on probing, or during mixed dentition, a primary tooth is missing between two abnormally positioned permanent teeth, see a dentist within 2 weeks; *level 3 earliest convenience*—any evidence of malocclusion, space loss, cross bites or crowding; *level 4 continue regular care*—applies when none of the above conditions exist.

## Statistical analyses

We constructed 88 questions to evaluate oral health based on existing ( $n = 187$ ) and newly written ( $n = 9$ ) items from focus group [36]. If a response option was selected by fewer than three field test participants, we collapsed the category with the adjacent lower level option for scoring. Item skewness was examined as one indicator of poor fit of the item

to the sample. Highly skewed items (skewness  $< -7.25$ ) were excluded [44]. We examined product–moment correlations of the remaining 88 items with the COHSI and RR and found 25 items that were significantly positively correlated or had high correlation ( $r > 0.20$ ) with one or both of the clinical outcomes. To ensure full coverage of the oral health conceptual model, we selected ten other items (by criterion of relative high correlation within a given domain) for domains that had no items strongly correlating with the clinical outcomes, resulting in 35 items: 11 items correlated with the COHSI only, 10 items with the RR only, and 14 items with both.

Exploratory factor analysis (EFA) was performed on the 35 items, and 6 items with standardized loadings  $< 0.30$  were then excluded, resulting in 29 items. Then, two confirmatory factor analysis (CFA) models were fitted to evaluate the unidimensionality of the (a) 19 items that correlated substantially with the COHSI; and (b) the 22 items that correlated with the RR. Nine items were included in both factors. While these two sets of items were hypothesized to assess a common higher-order factor, it is possible for the “COHSI” score to be low due to occlusal conditions, such as overbite, crowded, and spacing, but no need to see a dentist urgently—that is, the child could wait until the occlusal issues are ready for future orthodontic intervention. The practical fit of the CFA models was evaluated using the comparative fit index (CFI) ( $> 0.90$ ), the root mean square error of approximation (RMSEA) ( $< 0.06$ ), and standardized root mean square residual (SRMR) ( $< 0.08$ ) [45–48]. Local independence was evaluated by looking for non-trivial residual correlations (0.20 or above).

Samejima's graded response model was used to estimate item discrimination and difficulty parameters for the 19 items in the “COHSI” scale and the 22 items in the “RR” scale [49–52]. The monotonicity assumption of item response theory (IRT) was evaluated by checking item characteristics curves (ICCs) to see if the probability of choosing response options representing more positive oral health increases with better underlying oral health. Finally, we evaluate differential item functioning (DIF) [53, 54] regarding age (8–12 vs. 13–17) and gender and their impact, if any, on IRT modeling using ordinal logistic regression (OLR) and nested OLR with pseudo  $R^2$  [55].

We identified short-form versions of the 19-item “COHSI” and 22-item “RR” scales based on item parameters and inputs from content experts. We gave preference to items with higher discrimination and items with a range of difficulties. We also wanted the subset of items selected to represent a wide range of domains in the conceptual model. The extent to which the short-forms capture the information in the full-length scales was assessed by estimating intra-class correlations between estimated scores and comparing reliability (information) of measurement [50].

The readability of long- and short-form items was evaluated by the Flesch–Kincaid (F–K) grade level and flesch reading ease (FRE) scores using Microsoft Word Software [56, 57]. The F–K and FRE are based on average number of syllables per word and number of words per sentence [58].

All items were transformed linearly to a 0–100 possible range with a higher score representing better oral health [50, 59]. The COHSI and RR scales were scored on a *t* score metric with a mean of 50 and standard deviation of 10 in the field test sample.

SAS, STATA, and Mplus software [60–62] were used to calibrate the items and generate comparison plots.

## Results

The sample included 334 individuals recruited from August 2015 to October 2016 at 12 dental clinics and practices in Greater Los Angeles, including general clinics and pediatric clinics. Participant characteristics are shown in Table 1. The overall mean COHSI was 88.97 with a standard deviation 8.8 (COHSI ranged from 59.18 to 100). There were 24 (7%) children who were classified as needing immediate care, while 52% children were classified to continue their regular routine care. The age groups and gender groups were

**Table 1** Characteristics of the children and adolescents in the field test

Variables	Mean (SD) or <i>N</i> (%)
Children’s Oral Health Status Index (COHSI)	89.0 (8.8)
Clinical recommendation	
Continue your regular routine care	174 (52.1%)
See a dentist at your earliest convenience	53 (15.9%)
See a dentist within the next 2 weeks	83 (24.9%)
See a dentist immediately	24 (7.2%)
Mean age	12.08 (2.81%)
8–12	193 (57.8%)
13–17	141 (42.2%)
Gender	
Male	173 (51.9%)
Female	160 (47.8%)
Female to male transgender	1 (0.3%)
Race/ethnicity	
Alaska Native/American Indian	1 (0.3%)
Asian	43 (12.9%)
Black/African American	25 (7.5%)
Hispanic/Latino	140 (41.9%)
Pacific Islander	2 (0.6%)
Caucasian/White	71 (21.3%)
Multiracial	26 (7.8%)
Other	26 (7.8%)

balanced with 58% age 8–12 and 52% males. The sample had 42% Hispanic and 8% multiracial children.

The CFI, SRMR, and RMSEA were 0.94, 0.04, and 0.06, respectively, for the 15-item COHSI scale. These fit indices were 0.90, 0.04, and 0.08, respectively, for the 22-item RR scale. The residual correlations range from  $-0.55$  to  $0.37$  and  $-0.69$  to  $0.52$  for COHSI and RR, respectively. There are 26 pairs with absolute residual correlation larger than 0.2 for COHSI and 33 pairs for RR.

Item parameters from the GRM are shown in Tables 2 and 3. Discrimination (slope) parameters ranged from 1.47 to 2.62 for the COHSI scale and from 0.00 to 3.18 for the RR scale. Item characteristic curves provided support for the monotonicity assumption.

Ordinal logistic regression (OLR) on the IRT-calibrated latent trait, oral health items selected in long form, with and without age group (8–12 vs. 13–17), or gender (male vs. female) were performed to check the no DIF assumption [55]. Items with DIF between age group 8–12 and 13–17 are presented in Tables 2 and 3.

The eight items for COHSI and seven items for RR selected for the short-forms are bolded in Tables 2 and 3 and shown in Table 4. In Table 4, the items are divided into two parts showing the measures of COHSI and RRs separately or together. Using Tables 5 and 6, we can convert the summation of the short-form score to corresponding clinical outcomes—COHSI and RRs. The first column of the two tables is the raw summation of short-form score. The second column and third column are the corresponding latent variable *t* score and standard error to estimate COHSI and RRs. We also report the sample size (% of the sample) in both of the two tables. As noted above, these items were selected based on the slope (higher slope), threshold parameters (wider range), domain representation, and expert panel opinions. The overall F–K and FRE scores are 5.0 and 85.7 for long form, and 3.3 and 95.3 for short-form, which correspond to 3rd or 4th grade reading level (children 8–9 years old).

Information curves for the full-length and short-form scales are provided in Fig. 2. Intra-class correlations between the estimated scores for the corresponding full-length and short-form scales were 0.86 for COHSI and 0.87 for RR. The estimated correlation between the COHSI and the RR scale was 0.90 (full-length) and 0.74 (short-form).

## Discussion

Guided by the PROMIS framework, the initial items for measuring children’s oral health were developed from a literature review of existing survey questions that measure oral health status of children and adolescents [35]. Newly designed questions from focus group interview [36] were also included in the initial item pool. The revised items serve

**Table 2** Item calibration statistics for Children’s Oral Health Status Index

Subcomponent	Items <sup>b</sup>	Outcome	Skewness	DIF	Slope Threshold				
					1	2	3	4	5
PHY/symptoms	<b>It hurts my teeth to chew</b>	Both	-2.34		2.62	-6.90	-4.70	-2.55	
PHY/symptoms	I had a tooth that hurts	Both	-1.56	Gender	1.68	-5.25	-3.14	-1.12	
PHY/symptoms	My mouth hurts	Both	-1.63	Age <sup>a</sup>	1.47	-4.21	-2.39	-1.14	
PHY/symptoms	In the last 4 weeks, how much of the time did you have pain or discomfort?	Both	-2.07		1.46	-5.18	-4.67	-3.71	-2.26 -0.79
PHY/symptoms	During the past twelve months, how much pain or discomfort from dental problems did you have?	Both	-1.19		1.31	-5.19	-3.13	-0.44	
PHY/functions	<b>In the last 4 weeks, how much of the time did you limit the kinds or amounts of foods because of problems with your mouth, tongue, teeth, jaws, or gums?</b>	Both	-1.76		1.25	-5.15	-4.41	-3.47	-1.84 -0.63
MEN/affect	When I look at my teeth	Index	-0.45		1.22	-3.77	0.04		
MEN/affect	Compared to your classmates and friends how do you think your teeth look?	Index	0.06		1.07	-3.22	-0.26	1.67	
MEN/affect	<b>In the last 4 weeks, how much of the time were you pleased or happy with the look of your mouth, teeth, jaws, or gums?</b>	Both	-0.92	Age <sup>a</sup>	1.03	-3.80	-2.84	-1.48	-0.77 0.80
SOC/functions	<b>Have you ever avoided laughing or smiling because of the way your teeth look?</b>	Index	-1.61		0.93	-1.72			
MEN/affect	Compared to others my age	Index	-0.04		0.88	-3.39	0.66		
PHY/functions	<b>In general, would you say your overall oral health is</b>	Both	0.04		0.87	-3.11	-0.29	1.90	
PHY/symptoms	<b>My teeth are straight</b>	Both	-0.75	Age <sup>a</sup>	0.87	-0.39			
PHY/symptoms	My teeth have problems (space, crooked and crowded)	Both	-0.34		0.85	-0.85			
PHY/functions	In the last 4 weeks, how much of the time were you able to swallow comfortably?	Index	-2.17	Age <sup>a</sup>	0.70	-3.06	-2.72	-2.19	-1.89 -1.34
SOC/functions	In the last 4 weeks, how much of the time did your oral health interfere with your social activities?	Index	-1.74	Age <sup>a</sup>	0.68	-3.52	-2.88	-2.14	-1.48 -0.63
MEN/affect	<b>How much are you afraid to go to a dentist?</b>	Both	-1.52	Age <sup>a</sup>	0.67	-3.48	-2.36	-0.49	
MEN/behavior	<b>How often do you brush your teeth?</b>	Index	-2.27	Age <sup>a</sup> , gender	0.48	-3.48	-3.03	-2.44	-0.64
MEN/affect	Did any of the following reasons ever keep you from visiting a dentist?: I was afraid the treatment might be painful or the dentist might hurt me	Both	-2.17		0.46	-1.96			

<sup>a</sup>No DIF with age if within age group

<sup>b</sup>Items in bold indicate being selected in the Static short-form

**Table 3** Item calibration statistics for referral

Subcomponent	Items <sup>b</sup>	Outcome	Skewness	DIF	Slope	Threshold				
						1	2	3	4	5
PHY/symptoms	<b>It was hard for me to eat because of the pain in my mouth</b>	Referral	-2.53		3.18	-7.56	-6.50	-4.78	-2.63	
MEN/cognition	<b>It was hard for me to pay attention because of the pain in my mouth</b>	Referral	-3.22		3.17	-7.58	-5.74	-4.11		
PHY/symptoms	I felt stressed because of the pain in my mouth	Referral	-3.23		3.03	-7.27	-5.38	-4.14		
PHY/symptoms	It hurts my teeth to chew	Both	-2.34		2.92	-7.34	-5.07	-2.82		
PHY/symptoms	My mouth hurts	Both	-1.63	Age <sup>a</sup>	2.35	-6.24	-3.85	-1.43		
PHY/symptoms	It was hard for me to sleep because of the pain in my mouth	Referral	-3.83		2.05	-6.19	-4.84	-3.45		
PHY/symptoms	I had a tooth that hurts	Both	-1.56	Gender	1.97	-4.73	-2.76	-1.35		
PHY/symptoms	In the last 4 weeks, how much of the time did you have pain or discomfort?	Both	-2.07	Age <sup>a</sup>	1.79	-5.60	-5.05	-4.07	-2.52	-0.91
PHY/symptoms	During the past twelve months, how much pain or discomfort from dental problems did you have?	Both	-1.19	Age	1.44	-5.35	-3.24	-0.48		
PHY/functions	<b>In the last 4 weeks, how much of the time did you limit the kinds or amounts of foods because of problems with your mouth, tongue, teeth, jaws, or gums?</b>	Both	-1.76		1.33	-5.23	-4.48	-3.53	-1.91	-0.67
SOC/relationships	<b>Do other students make jokes about the way your teeth look?</b>	Referral	-5.49	Gender	1.21	-4.12				
MEN/affect	<b>In the last 4 weeks, how much of the time were you pleased or happy with the look of your mouth, teeth, jaws, or gums?</b>	Both	-0.92		0.68	-3.53	-2.61	-1.36	-0.71	0.72
MEN/cognition	Did any of the following reasons ever keep you from visiting a dentist? I thought the dental trouble I had would go away	Referral	-3.48		0.58	-2.80				
MEN/affect	How much are you afraid to go to a dentist?	Both	-1.52	Age <sup>a</sup>	0.52	-3.41	-2.31	-0.48		
PHY/functions	<b>In general, would you say your overall oral health is</b>	Both	0.04		0.52	-2.92	-0.28	1.75		
PHY/symptoms	My teeth have some problems (space, crooked, or crowded)	Both	-0.34		0.46	-0.36				
PHY/symptoms	My teeth are straight	Both	-0.75	Age	0.45	-0.78				
MEN/affect	Did any of the following reasons ever keep you from visiting a dentist? I was afraid the treatment might be painful or the dentist might hurt me	Both	-2.17		0.42	-1.95				
PHY/symptoms	How often do you have bad breath?	Referral	-0.19	Gender	0.41	-3.66	-2.34	0.27	2.21	
MEN/cognition	<b>Flossing my teeth, I can</b>	Referral	-3.56		0.27	-4.03	-2.34			
MEN/behavior	How often do you use dental floss on your teeth?	Referral	-0.17		0.16	-1.53	-0.68	-0.15	0.68	2.06
MEN/cognition	Brushing my teeth, I can	Referral	-2.95		0.00	-4.41	-2.05			

<sup>a</sup>No DIF with age if within age group

<sup>b</sup>Items in bold indicate being selected in the Static short-form

**Table 4** PROMIS child oral health short-form items

Subcomponent	Item	Response	Outcome
PHY/symptoms	It hurts my teeth to chew	Always/almost always/often(0); sometimes (2); almost never (4); never (5)	Index
SOC/functions	Have you ever avoided laughing or smiling because of the way your teeth look?	Yes (0); no (5)	Index
PHY/symptoms	My teeth are straight	Yes (0); no (5)	Index
MEN/affect	How much are you afraid to go to a dentist?	A great deal (0); somewhat (2); a little bit (4); not at all (5)	Index
MEN/behavior	How often do you brush your teeth?	Never/a few (2–3) times a month (0); once a week (2); a few (2–3) times a week (3); once a day (4); two or more times a day (5)	Index
PHY/functions	In general, would you say your overall oral health is	Fair/poor (0); good (2); very good (4); excellent (5)	Both index and referral
PHY/functions	In the last 4 weeks, how much of the time did you limit the kinds or amounts of foods because of problems with your mouth, tongue, teeth, jaws, or gums?	Always (0); almost always (1); often (2); sometimes (3); almost never (4); never (5)	Both index and referral
MEN/affect	In the last 4 weeks, how much of the time were you pleased or happy with the look of your mouth, teeth, jaws, or gums?	Never (0); almost never (1); sometimes (2); often (3); almost always (4); always (5)	Both index and referral
PHY/symptoms	It was hard for me to eat because of the pain in my mouth	Always/almost always(0); often(2); sometimes (3); almost never (4); never (5)	Referral
MEN/cognition	It was hard for me to pay attention because of the pain in my mouth	Always/almost always/often(0); sometimes (2); almost never (4); never (5)	Referral
SOC/relationships	Do other students make jokes about the way your teeth look?	Yes (0); no (5)	Referral
MEN/cognition	Flossing my teeth, I can	It will make no difference to my oral health(0); keep my oral health the same(3); make my oral health better (5)	Referral



**Table 5** *T* score conversion table for Children's Oral Health Status Index

Raw score	<i>t</i> score	SE	<i>N</i> (%)
11	28.50	6.18	1 (0.3)
12	30.52	6.86	1 (0.3)
14	28.80	5.42	3 (0.9)
15	30.54	5.90	1 (0.3)
16	30.29	5.54	2 (0.6)
18	33.49	5.20	1 (0.3)
19	36.40	5.71	3 (0.9)
20	36.52	5.61	6 (1.8)
21	37.50	5.70	3 (0.9)
22	38.09	5.53	5 (1.5)
23	38.65	5.54	5 (1.5)
24	39.83	5.66	9 (2.7)
25	43.32	5.87	8 (2.4)
26	43.31	5.81	3 (0.9)
27	42.04	5.59	9 (2.7)
28	44.04	5.81	17 (5.1)
29	44.37	5.74	16 (4.8)
30	47.76	6.02	22 (6.6)
31	48.70	6.00	26 (7.8)
32	49.44	6.00	17 (5.1)
33	50.48	6.06	24 (7.2)
34	51.92	6.18	19 (5.7)
35	52.80	6.18	22 (6.6)
36	53.36	6.19	16 (4.8)
37	55.29	6.31	27 (8.1)
38	57.16	6.37	33 (9.9)
39	61.12	6.75	24 (7.2)
40	65.30	7.24	11 (3.3)

SE standard error on *t* score metric

as the foundation for developing static short-forms, and later computer adaptive testing (CAT) with different purposes for oral health evaluation, screening, and surveillance monitoring. A sample size of 250 or more was enough to estimate the parameters for GRM [63], but a sample size 500 and over was recommended for estimating latent variable accurately and calibrating stable parameters of the items. Although this is ideal, a smaller sample can still provide useful information. In the sample, we did not include those who have orthodontic appliances and who could not speak English.

The on-site clinical examinations were performed independently by two dentists on the research team who were calibrated with regard to the examination instrument and standards prior to the examination. The exam checked the entire mouth for occlusion, oral mucosal tissue, overbite, overjet, cross bite, space loss, and crowding issue for primary and permanent teeth. Examination was done on a tooth-by-tooth basis, including missing, decayed, filled,

sound but fractured, abnormal positioned, intrinsic/extrinsic enamel opacity, and bleeding. The RRs are given immediately at the site to the participants and their parent or guardian. The two outcomes, COHSI score and RRs, represent two perspectives of the oral health status, namely the overall evaluation of the child's current oral health status and the examiner's determination as to whether a dental appointment is necessary beyond regular dental visit. This dentist-based clinical examination model is, however, not feasible for evaluating oral health of large populations of children and it has been very challenging to find a feasible and cost effective alternative approach to monitor oral health of large populations of school children.

The aim for this paper is to develop a static short-form that measures current oral health status and provides recommendations for different levels of treatment needs, which could eventually serve as a screening tool for efficiently assessing oral health of children and adolescents. The domain hierarchy for oral health developed in this paper is designed to comprehensively cover different aspects of oral health from physical, mental, and social perspectives. The long form includes all domains in the conceptual model that directly measure oral health conceptually from physical, mental, and social aspects of health. The final combined short-forms of COHSI and RRs include 12 items, which minimize the number of questions and significantly reduce the burden for subjects, while at the same time maintain as much information in the long form as possible. The short-form can be effectively applied to large populations of children to evaluate their oral health status and provide a treatment recommendation for dental care. However, one does not have to administer both COHSI and RR together. If needed, one can only administer items for COHSI or RR to obtain its corresponding *t* score. The summation score for the short-form responses, used together with Table 5 for COHSI and Table 6 for RRs, can estimate the current overall oral health status and treatment needs, respectively.

This study has limitations. We recruited 334 children/adolescents with age 8–17, and all of them had access to dental care, with fewer oral health problems than the general population of children. Very few subjects needed a recommendation to see a dentist immediately or within 2 weeks. Given the complexity of residential mobility in Los Angeles, we did not take into account the variation among different areas, either demographic differences or differential oral health status in the analysis. We pooled the data collected from different clinics together to conduct the analysis and develop long and short-forms.

The long and the short-forms created in this study provide a possible alternative solution to the long-lasting challenge of assessing oral health in large populations of children in care, who have already visited a dentist and has a dental clinic (dental home) to visit again when in need. For more

**Table 6** *T* score conversion table for referral recommendations

Raw score	<i>t</i> score	SE	Category	<i>N</i> (%)
8	21.13	4.84	Level 1—see dentist immediately	1 (0.3)
10	21.13	4.84		1 (0.3)
13	24.87	4.30		1 (0.3)
15	29.46	3.65		1 (0.3)
16	29.46	3.65		1 (0.3)
17	29.46	3.65		1 (0.3)
18	31.73	4.59		3 (0.9)
20	35.65	3.62		Level 2—see dentist within the next 2 weeks
21	36.47	4.21	5 (1.5)	
22	36.47	4.21	8 (2.4)	
23	39.60	4.48	7 (2.1)	
24	40.78	4.62	16 (4.8)	
25	45.10	5.24	Level 3—see dentist at your earliest convenience	13 (3.9)
26	45.10	5.24		20 (6.0)
27	47.93	5.76	Level 4—see dentist at your regular routine	31 (9.3)
28	50.37	6.16		25 (7.5)
29	51.24	6.30		42 (12.6)
30	51.74	6.42		42 (12.6)
31	53.29	6.62		29 (8.7)
32	54.69	6.88		31 (9.3)
33	55.16	6.95		20 (6.0)
34	58.88	7.63		24 (7.2)
35	61.26	8.04		10 (3.0)

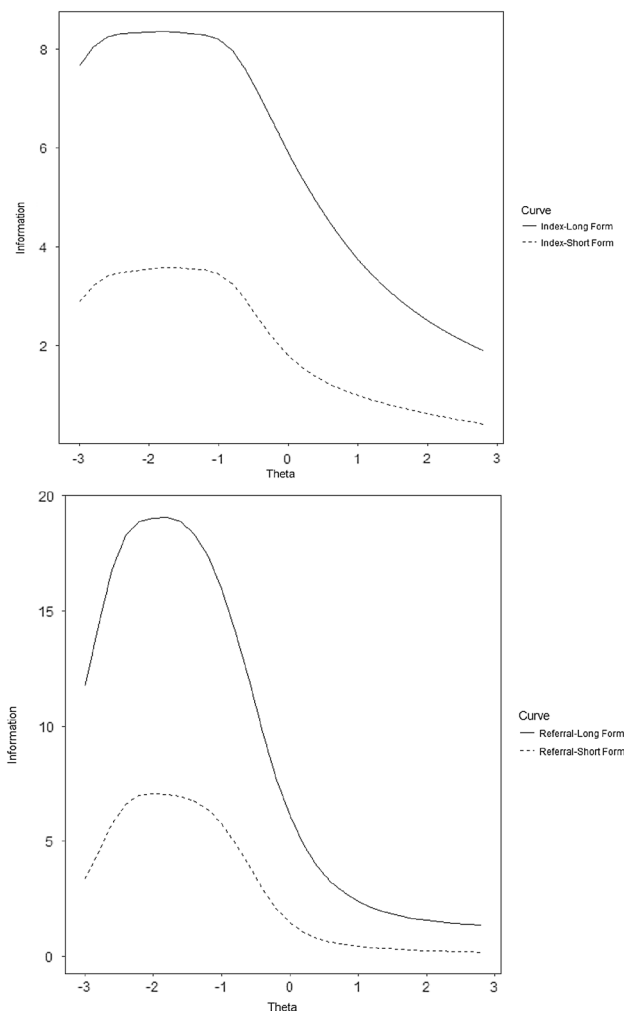
*SE* standard error on *t* score metric

general populations, we will need to incorporate additional questions. For example, we will need to consider the unique situations of those who are not in care yet. This may include items related to frequency of dental visits, and some clinical questions, such as dry mouth, gingival bleeding when brushing or flossing, etc.

We recommend that people use the short-forms as they are. However, it is also possible for future users to add other items and evaluate whether they contribute additional value (information) beyond the existing items. Furthermore, future evaluation of the existing items in other samples is needed. In addition, those items that are excluded for score estimation due to skewness may need to be included for other purposes, e.g., smoking or tobacco use as an important indicator for poor oral health. The short-form was developed based on the sample recruited from dental clinics located within Los Angeles County, with higher percentages of Latinos and Asians than the national proportion. This, together with the higher COHSI score and fewer subjects with an urgent need to see a dentist, made it more difficult to directly generalize the current short-form (as well as the current long form) to a broader general population without modifications. Specific adjustments that consider the variation of a particular subgroup should be considered when using the short-form. This

version of the short-form not only provides recommendations for needed care but also estimates current overall oral health status. While the screening by short-forms can never totally replicate the examination by dental professionals, it can provide a cost-effective way to make oral health screening feasible for large populations of children and adolescents. It can be used in practices at local, state, and even national level for triaging the large children and adolescent populations with oral health needs and setting the priorities within populations with varying urgency for dental and oral health care. This is especially true with school-aged children and adolescents who may need immediate care, and therefore contribute to gaining more timely access to needed care.

The initial short-forms presented in this paper are not disease-targeted, although it could be designed to focus on dental caries, missing teeth, periodontal diseases, etc. The generic version of the short-forms for oral health can be used to compare the relative burden of oral-related diseases, and evaluate overall oral health status and need for care of large populations. In the process, we flagged items with DIF instead of directly excluding them from the design, as these items provide the insights of fully measuring oral health, for example, the aesthetic domain might not be the same for males and females.



**Fig. 2** Information curve for both COHSI and referral (long form vs. short-form). Conversion of raw scores to  $t$  scores are given in Table 5 (COHSI) and Table 6 (RR)

In conclusion, this is the initial paper for developing the generic version of oral health short-forms that can cover the different domains in the oral health conceptual structure including components of global health, physical, mental, and social health. The overall 12-item short-form can be effectively used in large populations of children to evaluate their current overall oral health status and treatment needs at different levels. Using a standardized  $t$  score metric, it is easy to compare oral health status across individuals and populations. With further tuning, the short-form can potentially have better accuracy and higher sensitivity and it can be eventually be implemented through computer, internet, and smart phone apps at subject's preferred place and time. With the created long form shown in Tables 2 and 3, CAT system, a dynamic and efficient survey tool, could be developed in the next step, in order to be more effective and accurate in survey process, notably for ease of scoring and

immediate plotting results in real-time [31]. The oral health item bank system created in this paper provides the foundation for other purposes such as creating specific targeted short-form(s) for program evaluation and/or oral health policy planning and others.

**Acknowledgements** We would like to acknowledge the following individuals for their assistance in both collecting the samples and providing the space for the field test: Drs. Suzanne Berger, Christie Burnett, Adrienne Fang, Roger Fieldman, Ms. Dale Gorman, Karen Ho, Oariona Lowe, Mr. Pastor Miranda, Evangelos Rossopoulos, Megha Sata, Justin Shuffer and Ms. Audrey Simons, and Mr. Dennis Young.

**Funding** This research was supported by a National Institute of Dental and Craniofacial Research/NIH grant to the University of California, Los Angeles [U01DE022648].

## Compliance with ethical standards

**Conflict of interest** The authors declare that they have no conflict of interest.

**Ethical approval** All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

**Informed consent** Informed consent was obtained from all individual participants included in the study.

## References

1. U.S. Department of Health and Human Services (USDHHS), Office of Disease Prevention and Health Promotion. (2000). Healthy People 2010. Washington, DC: U.S. Government Printing Office.
2. Munoz, K. A., Krebs-Smith, S. M., Ballard-Barbash, R., & Cleveland, L. E. (1997). Food intakes of US children and adolescents compared with recommendations. *Pediatrics*, *100*(3), 323–329.
3. Frary, C. D., Johnson, R. K., & Wang, M. Q. (2004). Children and adolescents' choices of foods and beverages high in added sugars are associated with intakes of key nutrients and food groups. *Journal of Adolescent Health*, *34*(1), 56–63.
4. Petersen, P. E., Bourgeois, D., Ogawa, H., Estupinan-Day, S., & Ndiaye, C. (2005). The global burden of oral diseases and risks to oral health. *Bulletin of the World Health Organization*, *83*(9), 661–669.
5. Kalsbeek, H., & Verrips, G. (1994). Consumption of sweet snacks and caries experience of primary school children. *Caries Research*, *28*(6), 477–483.
6. Ilma de Souza Cortes, M., Marcenes, W., & Sheiham, A. (2002). Impact of traumatic injuries to the permanent teeth on the oral health-related quality of life in 12–14-year-old children. *Community Dentistry and Oral Epidemiology*, *30*(3), 193–198.
7. Erić, J., Stančić, I., Šojić, L. T., Popovac, A. J., & Tsakos, G. (2012). Validity and reliability of the Oral Impacts on Daily Performance (OIDP) scale in the elderly population of Bosnia and Herzegovina. *Gerodontology*, *29*(2), e902–e908.
8. Åström, A. N., & Okullo, I. (2003). Validity and reliability of the Oral Impacts on Daily Performance (OIDP) frequency scale: A

- cross-sectional study of adolescents in Uganda. *BMC Oral Health*, 3(1), 5.
9. Ilha, L., Martins, A., & Abegg, C. (2016). Oral impact on daily performance: Need and use of dental prostheses among Brazilian adults. *Journal of Oral Rehabilitation*, 43(2), 119–126.
  10. Agrawal, N., Pushpanjali, K., Gupta, N., & Garg, A. K. (2014). Child-Oral impacts on daily performances: A socio dental approach to assess prevalence and severity of oral impacts on daily performances in South Indian school children of Bangalore city: A cross-sectional survey. *Journal of Indian Association of Public Health Dentistry*, 12(2), 88.
  11. Yusuf, H., Gherunpong, S., Sheiham, A., & Tsakos, G. (2006). Validation of an English version of the Child-OIDP index, an oral health-related quality of life measure for children. *Health and Quality of Life Outcomes*, 4(1), 38.
  12. Bernabé, E., Sheiham, A., & Tsakos, G. (2008). A comprehensive evaluation of the validity of Child-OIDP: Further evidence from Peru. *Community Dentistry and Oral Epidemiology*, 36(4), 317–325.
  13. Ferreira, M. C., Goursand, D., Bendo, C. B., Ramos-Jorge, M. L., Pordeus, I. A., & Paiva, S. M. (2012). Agreement between adolescents' and their mothers' reports of oral health-related quality of life. *Brazilian Oral Research*, 26(2), 112–118.
  14. Jokovic, A., Locker, D., Tompson, B., & Guyatt, G. (2004). Questionnaire for measuring oral health-related quality of life in eight- to ten-year-old children. *Pediatric Dentistry*, 26(6), 512–518.
  15. Jokovic, A., Locker, D., Stephens, M., Kenny, D., Tompson, B., & Guyatt, G. (2002). Validity and reliability of a questionnaire for measuring child oral-health-related quality of life. *Journal of Dental Research*, 81(7), 459–463.
  16. Jokovic, A., Locker, D., & Guyatt, G. (2006). Short forms of the Child Perceptions Questionnaire for 11–14-year-old children (CPQ 11–14): Development and initial evaluation. *Health and Quality of Life Outcomes*, 4(1), 4.
  17. Broder, H. L., & Wilson-Genderson, M. (2007). Reliability and convergent and discriminant validity of the Child Oral Health Impact Profile (COHIP Child's version). *Community Dentistry and Oral Epidemiology*, 35(s1), 20–31.
  18. Slade, G. D., & Spencer, A. J. (1994). Development and evaluation of the Oral Health Impact Profile. *Community Dental Health*, 11(1), 3–11.
  19. Pahl, B. T., Rozier, R. G., & Slade, G. D. (2007). Parental perceptions of children's oral health: The Early Childhood Oral Health Impact Scale (ECHOIS). *Health and Quality of Life Outcomes*, 5(1), 6.
  20. Tsakos, G., Allen, P. F., Steele, J. G., & Locker, D. (2012). Interpreting oral health-related quality of life data. *Community Dentistry and Oral Epidemiology*, 40(3), 193–200.
  21. Tsakos, G., Blair, Y. I., Yusuf, H., Wright, W., Watt, R. G., & Macpherson, L. M. (2012). Developing a new self-reported scale of oral health outcomes for 5-year-old children (SOHO-5). *Health and Quality of Life Outcomes*, 10(1), 62.
  22. Filstrup, S. L., Briskie, D., Da Fonseca, M., Lawrence, L., Wandera, A., & Inglehart, M. R. (2003). Early childhood caries and quality of life: Child and parent perspectives. *Pediatric Dentistry*, 25(5), 431–440.
  23. Pakpour, A. H., Yekaninejad, M. S., Zarei, F., Hashemi, F., Steele, M. M., & Varni, J. W. (2011). The PedsQL™ Oral Health Scale in Iranian children: Reliability and validity. *International Journal of Paediatric Dentistry*, 21(5), 342–352.
  24. Bendo, C. B., Paiva, S. M., Viegas, C. M., Vale, M. P., & Varni, J. W. (2012). The PedsQL™ Oral Health Scale: Feasibility, reliability and validity of the Brazilian Portuguese version. *Health and Quality of Life Outcomes*, 10(1), 42.
  25. Huntington, N. L., Spetter, D., Jones, J. A., Rich, S. E., Garcia, R. I., & Spiro, I. I., A (2011). Development and validation of a measure of pediatric oral health-related quality of life: The POQL. *Journal of Public Health Dentistry*, 71(3), 185–193.
  26. Perlin, J. B., Kolodner, R. M., & Roswell, R. H. (2004). The Veterans Health Administration: Quality, value, accountability, and information as transforming strategies for patient-centered care. *The American Journal of Managed Care*, 10(11 Pt 2), 828–836.
  27. Snyder, C. F., Jensen, R. E., Segal, J. B., & Wu, A. W. (2013). Patient-reported outcomes (PROs): Putting the patient perspective in patient-centered outcomes research. *Medical Care*, 51(8 0 3), S73.
  28. Forrest, C. B., Bevans, K. B., Pratiwadi, R., Moon, J., Teneralli, R. E., Minton, J. M., et al. (2014). Development of the PROMIS (R) pediatric global health (PGH-7) measure. *Quality of Life Research*, 23(4), 1221–1231. <https://doi.org/10.1007/s11136-013-0581-8>.
  29. Cella, D., Riley, W., Stone, A., Rothrock, N., Reeve, B., Yount, S., et al. (2010). The Patient-Reported Outcomes Measurement Information System (PROMIS) developed and tested its first wave of adult self-reported health outcome item banks: 2005–2008. *Journal of Clinical Epidemiology*, 63(11), 1179–1194.
  30. Cella, D., Yount, S., Rothrock, N., Gershon, R., Cook, K., Reeve, B., et al. (2007). The Patient-Reported Outcomes Measurement Information System (PROMIS): Progress of an NIH Roadmap cooperative group during its first two years. *Medical Care*, 45(5 Suppl 1), S3.
  31. Cella, D., Gershon, R., Lai, J. S., & Choi, S. (2007). The future of outcomes measurement: Item banking, tailored short-forms, and computerized adaptive assessment. *Quality of Life Research*, 16(Suppl 1), 133–141. <https://doi.org/10.1007/s11136-007-9204-6>.
  32. Fries, J., Bruce, B., & Cella, D. (2005). The promise of PROMIS: Using item response theory to improve assessment of patient-reported outcomes. *Clinical and Experimental Rheumatology*, 23(5), S53.
  33. WHO (2005). WHO STEPS surveillance manual: The WHO STEPwise approach to chronic disease risk factor surveillance.
  34. Ader, D. N. (2007). Developing the patient-reported outcomes measurement information system (PROMIS). Philadelphia: Lippincott Williams & Wilkins.
  35. Liu, H., Hays, R. D., Marcus, M., Coulter, I., Maida, C., Ramos-Gomez, F., et al. (2016). Patient-Reported oral health outcome measurement for children and adolescents. *BMC Oral Health*, 16(1), 95.
  36. Maida, C. A., Marcus, M., Hays, R. D., Coulter, I. D., Ramos-Gomez, F., Lee, S. Y., et al. (2015). Child and adolescent perceptions of oral health over the life course. *Quality of Life Research*, 24(11), 2739–2751.
  37. Myers, D. (2013). California's diminishing resource: Children. Palo Alto: Lucile Packard Foundation for Children's Health.
  38. NOVA Research Company. (2006). Questionnaire development system. Bethesda, MA: Nova Research Company.
  39. Koch, A. L., Gershen, J. A., & Marcus, M. (1985). A children's oral health status index based on dentists' judgment. *The Journal of the American Dental Association*, 110(1), 36–42.
  40. Gershen, J., Marcus, M., & Koch, A. (1979). Using dentists' judgments to identify the components of children's oral health. *ASDC Journal of Dentistry for Children*, 47(6), 419–424.
  41. Hagan, P., Levy, S., & Machen, J. (1986). Validation of the children's Oral Health Status Index (COHSI). *ASDC Journal of Dentistry for Children*, 54(2), 110–113.
  42. CDC (2015–2016). National health and nutrition examination survey (nhanes): Oral health examiners manual.
  43. NCHS (2013). National health and nutrition examination survey questionnaire (examination protocol). [https://wwwn.cdc.gov/nchs/data/nhanes/2013-2014/manuals/Oral\\_Health\\_Examiners.pdf](https://wwwn.cdc.gov/nchs/data/nhanes/2013-2014/manuals/Oral_Health_Examiners.pdf).

44. Rose, M., Bjorner, J. B., Gandek, B., Bruce, B., Fries, J. F., & Ware, J. E. (2014). The PROMIS physical function item bank was calibrated to a standardized metric and shown to improve measurement efficiency. *Journal of Clinical Epidemiology*, *67*(5), 516–526.
45. Amtmann, D., Cook, K. F., Jensen, M. P., Chen, W.-H., Choi, S., Revicki, D., et al. (2010). Development of a PROMIS item bank to measure pain interference. *Pain*, *150*(1), 173–182.
46. Reise, S. P., & Haviland, M. G. (2005). Item response theory and the measurement of clinical change. *Journal of Personality Assessment*, *84*(3), 228–238.
47. Suhr, D. (2003). Reliability, exploratory & confirmatory factor analysis for the scale of athletic priorities. *Diambil Pada Tanggal*, *2*, 274–28.
48. Hu, L. t., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*(1), 1–55.
49. Buysse, D. J., Yu, L., Moul, D. E., Germain, A., Stover, A., Dodds, N. E., et al. (2010). Development and validation of patient-reported outcome measures for sleep disturbance and sleep-related impairments. *Sleep*, *33*(6), 781–792.
50. Watt, T., Bjorner, J. B., Groenvold, M., Cramon, P., Winther, K. H., Hegedüs, L., et al. (2015). Development of a short version of the thyroid-related patient-reported outcome ThyPRO. *Thyroid*, *25*(10), 1069–1079.
51. Yu, L., Buysse, D. J., Germain, A., Moul, D. E., Stover, A., Dodds, N. E., et al. (2012). Development of short forms from the PROMIS™ sleep disturbance and sleep-related impairment item banks. *Behavioral Sleep Medicine*, *10*(1), 6–24.
52. Hays, R. D., Morales, L. S., & Reise, S. P. (2000). Item response theory and health outcomes measurement in the 21st century. *Medical Care*, *38*(9 Suppl), I128.
53. Reeve, B. B., Hays, R. D., Bjorner, J. B., Cook, K. F., Crane, P. K., Teresi, J. A., et al. (2007). Psychometric evaluation and calibration of health-related quality of life item banks: Plans for the Patient-Reported Outcomes Measurement Information System (PROMIS). *Medical Care*, *45*(5), S22–S31.
54. Teresi, J. A., Ocepek-Welickson, K., Kleinman, M., Eimicke, J. P., Crane, P. K., Jones, R. N., et al. (2009). Analysis of differential item functioning in the depression item bank from the Patient Reported Outcome Measurement Information System (PROMIS): An item response theory approach. *Psychology Science Quarterly*, *51*(2), 148.
55. Paz, S. H., Spritzer, K. L., Morales, L. S., & Hays, R. D. (2013). Evaluation of the patient-reported outcomes information system (PROMIS®) Spanish-language physical functioning items. *Quality of Life Research*, *22*(7), 1819–1830.
56. Calderón, J. L., Morales, L. S., Liu, H., & Hays, R. D. (2006). Variation in the readability of items within surveys. *American Journal of Medical Quality*, *21*(1), 49–56.
57. Bevans, K. B., Gardner, W., Pajer, K., Riley, A. W., & Forrest, C. B. (2012). Qualitative development of the PROMIS® pediatric stress response item banks. *Journal of Pediatric Psychology*, *38*(2), 173–191.
58. Klare, G. R. (1976). A second look at the validity of readability formulas. *Journal of Reading Behavior*, *8*(2), 129–152.
59. Stewart, A. L., Hays, R. D., & Ware, J. E. (1988). The MOS short-form general health survey: Reliability and validity in a patient population. *Medical Care*, *26*(7), 724–735.
60. Pope, R. (2015). Item response theory models in Stata. In *2015 Stata Conference, 2015* (Vol. 23). Stata Users Group.
61. STATA (2015). Stata Statistical Software Release. College Station: Stata Press.
62. SAS Institute (2015). Base SAS 9.4 Procedures Guide, SAS Institute.
63. Reeve, B. B., & Fayers, P. (2005). Applying item response theory modeling for evaluating questionnaire item and scale properties. *Assessing Quality of Life in Clinical Trials: Methods of Practice*, *2*, 55–73.