

Mapping of EQ-5D-Y-5L responses from PedsQL: A feasibility study using data collected from children and adolescents in Hong Kong

Authors: Ivan C. H. AU ^{1#} BSc, Nan LUO ² PhD, Zhihao YANG ^{3,4} PhD, Carlos K. H. WONG ^{5,6,7,8} PhD

Affiliations:

¹ School of Public Health, LKS Faculty of Medicine, The University of Hong Kong, Hong Kong SAR, China

² Saw Swee Hock School of Public Health, National University of Singapore, Singapore, Singapore

³ Department of Health Services Management, Guizhou Medical University, Guiyang, China

⁴ Center of Medicine Economics and Management Research, Guizhou Medical University, Guiyang, China

⁵ Department of Pharmacology and Pharmacy, LKS Faculty of Medicine, The University of Hong Kong, Hong Kong SAR, China

⁶ Laboratory of Data Discovery for Health (D²4H), Hong Kong SAR, China

⁷ Department of Family Medicine and Primary Care, School of Clinical Medicine, LKS Faculty of Medicine, The University of Hong Kong, Hong Kong SAR, China

⁸ Department of Infectious Disease Epidemiology & Dynamics, Faculty of Epidemiology and Population Health, London School of Hygiene and Tropical Medicine, London, United Kingdom

first author

Corresponding author:

Dr Carlos King Ho Wong, Department of Family Medicine and Primary Care, School of Clinical Medicine, Li Ka Shing Faculty of Medicine, The University of Hong Kong, Hong Kong SAR, China. Phone: (+852) 2831-5055 Fax: (+852) 2814-7475 Email: carlosho@hku.hk

Abstract

Objectives: Currently, there is no EQ-5D-Y-5L (Y-5L) value set for direct mapping of utility scores from Pediatric Quality of Life Inventory (PedsQL) Generic Core scale, and response mapping algorithms of Y-5L from PedsQL have not yet been developed. This study assessed associations between Y-5L responses, and PedsQL responses and summary scores among children and adolescents, and explored the feasibility of mapping Y-5L responses from PedsQL scores.

Methods: A large sample of children and adolescents from 12 primary and 7 secondary schools in Hong Kong, China, were invited to complete the Y-5L and PedsQL questionnaires between May and July 2018. Mean PedsQL summary scores by Y-5L levels in each dimension were estimated, and the associations between Y-5L responses and PedsQL scores and items were measured by Spearman rank correlation. Mapping of Y-5L responses from PedsQL summary scores was estimated using ordered logistic regression.

Results: A total of 7,089 children and adolescents (male: 44.4%; age 6-8: 24.1%; age 9-11: 33.2%; age 12-17: 42.6%; mean age: 11.7, SD: 3.2) completed the Y-5L and PedsQL questionnaires independently without aid from teachers. Proportions of those with no problems were high in “Mobility” (MO) (93.4%), “Looking after myself” (SC) (95.7%) and “Doing usual activities” (UA) (92.7%), while less participants reported no problems in “Having pain or discomfort” (PD) (71.9%) and “Feeling worried, sad, or unhappy” (AD) (57.5%). The mean total PedsQL score decreased as Y-5L level increased in MO (L1: 81.8, L2: 69.7, L3: 66.7, L4: 54.7, L5: 53.5), SC (L1: 81.3, L2: 73.2, L3: 70.6, L4: 58.2, L5: 60.2), UA (L1: 81.9, L2: 69.4, L3: 63.9, L4: 49.5, L5: 53.3), PD (L1: 83.9, L2: 74.6, L3: 66.3, L4: 60.4, L5: 60.0), and AD (L1: 85.9, L2: 76.8, L3: 68.4, L4: 62.7, L5: 58.7). PD and AD had moderate associations with the PedsQL summary scores. Mapping of Y-5L dimensions from physical and psychosocial health summary scores had more accurate predictions for L1 in all five dimensions (MO: 99.8%, SC: 100.0%, UA: 99.7%, PD: 95.2%, AD: 83.1%) and less accurate predictions for L2-L5 (MO: 0.6%, SC: 0.0%, UA: 0.8%, PD: 13.5%, AD: 33.3%).

Conclusion: In this healthy cohort sampled from the school-attending children and adolescents in Hong Kong, our mapping algorithms demonstrated inadequate predictive performance in differentiating those with and without health problems, primarily due to insufficient prevalence of poor health states. Future development of Y-5L value set will facilitate direct mapping of Y-5L utility score from the PedsQL. Study findings call for collaborative efforts in collecting

both PedsQL and Y-5L responses from clinical samples with impaired health profiles for future mapping studies. Nevertheless, researchers are recommended to collect data using preference-based instruments even if a mapping algorithm exists. Mapping can only serve as an alternative when data collection using preference-based instruments is not feasible.

Main Text

Introduction

Health-related quality of life (HRQoL) in children and adolescents is an important outcome when evaluating the effectiveness and cost-effectiveness of emerging technologies for the youth population. A previous systematic review study identified more than 35 published Patient-Reported Outcome Measures (PROMs) for the pediatric demographic covering multiple age groups¹. Among these, the EQ-5D-Y and the Pediatric Quality of Life Inventory™ (PedsQL) are two widely utilized generic instruments in measuring HRQoL for children and adolescents, as well as pediatric patients with acute or chronic health conditions^{2,3}.

In studies where direct EQ-5D-Y measurement is unavailable, mapping algorithms serve as a pragmatic alternative, offering indirect evidence albeit at the cost of increased uncertainty and partial information loss. A systematic review study identified 30 studies mapping non-preference-based measures to generic preference-based measures, and highlighted varied performance among these mapping models and challenges in generalizing them across different instruments⁴. The development of mapping algorithms is an active area of methodological and applied research in health economics, particularly for estimating utility using other data in cost-utility analyses. According to the Health Economics Research Centre (HERC) Database of Mapping Studies, there are currently over 240 studies mapping to EQ-5D or EQ-5D-5L⁵, yet no studies mapped either to or from EQ-5D-Y were reported. A study in 2014 mapped PedsQL onto EQ-5D health utility scores among school-attending children aged 11-15 in the UK general population, yet the EQ-5D adult value set was used in the estimations⁶.

EQ-5D-Y has two versions. The EQ-5D-Y-3L (Y-3L), adapted from the standard EQ-5D-3L instrument for adults, is the youth version used to measure health in the youth populations⁷. The EQ-5D-Y-5L (Y-5L), a five-level version of the three-level Y-3L, was designed to increase sensitivity and reduce ceiling effects⁸. Currently, there is no Y-5L value set for direct mapping of utility scores from the PedsQL Generic Core scale, and response mapping algorithms of Y-5L from PedsQL are not yet available. A prior study showed high correlations between similar items on Y-5L and PedsQL, and moderate associations across PedsQL total score and Y-5L level sum and EQ VAS scores⁹. This evidence suggested the conceptual overlapping of the dimensions across the two instruments and thus the feasibility for mapping algorithms. The aim of this study was to assess associations amongst Y-5L responses, PedsQL responses, and

summary scores among children and adolescents (mostly healthy individuals) in Hong Kong, and to explore the feasibility of mapping PedsQL scores onto Y-5L responses.

Methods

Study sample

We sampled school-attending children and adolescents from primary or secondary schools in Hong Kong Special Administrative Region, China between May and July 2018. Thirty schools (Hong Kong Island: 7, Kowloon West: 4, Kowloon East: 6, New Territories West: 5, and New Territories East: 8) that participated in a population-based programme to promote and monitor physical fitness among Hong Kong students, the School Physical Fitness Award Scheme, were invited¹⁰. Among those schools, twelve primary and seven secondary schools consented to participate in this study, and sought research consent from parents/guardians of their students. Students were eligible if they (1) provided their parents/guardians' written consent for their participation in this study, (2) provided date of birth and gender, (3) were aged 6 to 17 years, (4) were studying at a primary or secondary school at the time of the interview, and (5) completed both the Y-5L and PedsQL questionnaires.

Data collection

The traditional Chinese (Hong Kong) versions of the Y-5L and PedsQL questionnaires were administered. Two versions of the survey were designed for students in primary school grades 1 to 3, and grades 4 to 6 and secondary school, respectively. Hardcopies of the survey were delivered to the schools and then distributed to the students by teachers in class. The questionnaires were completed by self-administration. Teachers were available at the time of interview to answer questions about the survey raised by students without influencing the answers. Data quality assurance process included double data entry and independent checking. Furthermore, inconsistent Y-5L responses and EQ VAS scores were excluded from the analysis if the student reported no problems in all five dimensions but the EQ VAS score was < 10, reported at least one extreme problem but EQ VAS score was 100, or the level sum score was ≥ 20 but the EQ VAS score was ≥ 90 .

Study Instruments

EQ-5D-Y-5L

The EQ-5D-Y is a child-specific and age-appropriate instrument for measuring HRQoL in children and adolescents². Its descriptive system consists of five dimensions: “mobility”, “looking after myself”, “doing usual activities”, “having pain or discomfort”, and “feeling worried, sad or unhappy”. The instrument has been shown to have high feasibility, good convergent validity, and moderate test–retest reliability among young population^{2,11}. Each dimension of the five-level Y-5L has a response scale from 1 to 5 (L1 = no problems, L2 = a little bit of problems, L3 = some problems, L4 = a lot of problems, L5 = cannot/extreme problems), and each of the 3,125 possible Y-5L health states can be represented using a five-digit code composed of the five response levels. For example, the health states 11111 and 55555 represent the full health and the worst health, respectively. The Y-5L instrument has been shown to be valid, reliable and responsive among pediatric patients in Hong Kong¹²⁻¹⁴. It also includes the EQ Visual Analogue Scale (EQ VAS), which assesses the overall health status of the respondents. It is a vertical scale from 0 to 100, with 0 labelled as “the worst health you can imagine” and 100 labelled as “the best health you can imagine”. Both the descriptive system and EQ VAS assess the health status of respondents on the day of the survey.

PedsQL

The PedsQL is a modular instrument for measuring HRQoL in children and adolescents, with proven validity and reliability¹⁵. The Chinese version of the PedsQL has been validated in Hong Kong¹⁶. The PedsQL 4.0 Generic Core Scales consist of four functioning scales with a total of 23 items: physical functioning (8 items), emotional functioning (5 items), social functioning (5 items), and school functioning (5 items). The instrument has three summary scores: physical health summary score, psychosocial health summary score, and total score. In the PedsQL questionnaire, respondents are asked about the frequency of problems encountered in the last month for the 23 items. A five-point Likert response scale ranging from 0 to 4 is used for each PedsQL item (0 = never a problem; 1 = almost never a problem; 2 = sometimes a problem; 3 = often a problem; 4 = almost always a problem). As the items are then reversely scored and linearly transformed to a scale from 0 to 100, a higher score indicates better HRQoL. Two different age-specific modules of PedsQL were used in the study. Students in primary school

grades 1 to 3 were given the PedsQL GCS Child version, while students in primary school grades 4 to 6 and secondary school were given the PedsQL GCS Adolescent version.

Statistical analysis

Mean PedsQL summary scores by Y-5L levels were estimated for each of the five dimensions. Trends in PedsQL summary scores across Y-5L levels were tested using the Jonckheere–Terpstra test. A valid mapping algorithm requires conceptual coherence between the source instrument (PedsQL) and target instrument (Y-5L)¹⁷, so the Y-5L and PedsQL responses were expected to be correlated for items in similar dimensions and domains. Spearman rank correlation was used to measure the associations between Y-5L responses and PedsQL scores and items. The strength of correlation was indicated by the significant correlation coefficient: low for 0.10–0.29, moderate for 0.30–0.49, and high for 0.50–1.00¹⁸.

Response mapping of Y-5L responses (dependent variables) from physical and psychosocial health summary scores (independent variables) was estimated using ordered logistic regression models. Regressions of the five levels of Y-5L on PedsQL summary scores were estimated separately for each dimension. As the Y-5L value set is not yet developed, calculation of predicted utility for model assessment was not viable in this study. Model performance of response mapping was reported by the proportion of correct predictions. Y-5L level with the highest predicted probability was taken as the predicted level. We also fitted two other models, one adjusting for demographic characteristics of age and sex, and one using PedsQL item responses as independent variables instead of the summary scores. In addition, all mapping models were fitted separately for each age group to account for potential heterogeneity across age groups. Direct mapping of the utility score from PedsQL summary scores was also not possible due to the absence of Y-5L value set.

All statistical analyses were performed using STATA/MP 18.0. All significance tests were two-tailed, where P-value <0.05 was considered statistically significant.

Ethical approval

Approval was obtained from the Institutional Review Board of the University of Hong Kong—the Hospital Authority Hong Kong West Cluster (reference number: UW 18-139).

Results

Descriptive statistics

Among the 8,225 students recruited in this study, 1,127 were excluded: 513 had missing data for either date of birth or gender, 68 reached the age of 18, 48 did not report their schools, 371 did not complete the Y-5L questionnaire, 76 did not provide PedsQL responses for physical health summary score or psychosocial health summary score, 51 had Y-5L responses inconsistent with EQ VAS scores. The final sample included 7,098 respondents, with 24.1% aged 6-8, 33.2% aged 9-11, and 42.6% aged 12-17. The mean age of the respondents was 11.7 years (SD: 3.2) and 44.4% were boys (Table 1). Differences in age and sex distributions were observed between our sample and the general Hong Kong population during the study period (Supplementary Table 1). Y-5L responses, EQ VAS scores, and PedsQL summary scores were different across primary schools and secondary schools (Supplementary Figure 1).

Distribution of Y-5L levels and PedsQL summary scores

Proportions of respondents with no problems were high in “mobility” (MO) (93.4%), “looking after myself” (SC) (95.7%) and “doing usual activities” (UA) (92.7%), while fewer participants reported no problems in “having pain or discomfort” (PD) (71.9%) and “feeling worried, sad or unhappy” (AD) (57.5%). Younger age groups reported fewer health problems in SC, UA, PD, and AD, while the distribution of severity levels was comparable across age groups for MO. The proportion of respondents with health state 11111 was 46.7%, and the mean EQ VAS score was 82.8 (SD: 18.0). Respondents in older age groups had less health state 11111 and lower mean EQ VAS score than the younger age groups. Mean PedsQL total score was 80.9 (SD: 13.3), mean physical health summary score was 85.2 (SD: 13.9), and mean psychosocial health summary score was 72.9 (SD: 19.8). Students aged 12-17 had lower total scores, physical health summary scores, and psychosocial health summary scores than the age 6-8 and age 9-11 groups (Table 1).

Mean PedsQL summary scores across Y-5L levels

Mean total PedsQL score decreased as the Y-5L level increased in MO (L1: 81.8, L2: 69.7, L3: 66.7, L4: 54.7, L5: 53.5; $P<0.001$), SC (L1: 81.3, L2: 73.2, L3: 70.6, L4: 58.2, L5: 60.2; $P<0.001$), UA (L1: 81.9, L2: 69.4, L3: 63.9, L4: 49.5, L5: 53.3; $P<0.001$), PD (L1: 83.9, L2: 74.6, L3: 66.3, L4: 60.4, L5: 60.0; $P<0.001$), and AD (L1: 85.9, L2: 76.8, L3: 68.4, L4: 62.7, L5: 58.7; $P<0.001$). Similar trends were observed for physical and psychosocial health summary scores (Table 2).

Correlations between Y-5L and PedsQL summary scores

Spearman correlations showed moderate associations between PD and the total PedsQL score (correlation [ρ]: -0.37), physical health summary score (ρ : -0.34), and psychosocial health summary score (ρ : -0.34). AD was moderately associated with the total PedsQL score (ρ : -0.48), physical health summary score (ρ : -0.34), and psychosocial health summary score (ρ : -0.49) in overall, and strongly associated with the total PedsQL score (ρ : -0.50) and psychosocial health summary score (ρ : -0.53) among those aged 12-17. No moderate or strong associations were observed between MO, SC, UA and the summary scores (Table 3).

Mapping of Y-5L dimensions from physical and psychosocial health summary scores

Mapping of Y-5L dimensions from physical and psychosocial health summary scores had more accurate predictions in all five dimensions for L1 (MO: 99.8%, SC: 100.0%, UA: 99.7%, PD: 95.2%, AD: 83.1%) and less accurate predictions for L2-L5 (MO: 0.6%, SC: 0.0%, UA: 0.8%, PD: 13.5%, AD: 33.3%). Prediction rates for L2-L5 increased with age for PD (age 6-8: 6.6%, age 9-11: 15.0%, age 12-17: 17.2%) and AD (age 6-8: 10.7%, age 9-11: 29.7%, age 12-17: 44.8%), while the prediction rates for L2-L5 were comparable across age groups for MO, SC, and UA (Table 4). Mapping models further adjusting for age and sex did not increase predictive rates, and models using PedsQL item responses as independent variables also had similar predictive performance (Supplementary Table 2).

Discussion

The aim of this study was to evaluate the associations between Y-5L and PedsQL and assess the feasibility of mapping PedsQL summary scores onto Y-5L responses. Trends in PedsQL

summary scores across Y-5L levels were observed. This suggested discriminatory power for utilizing summary scores to distinguish Y-5L levels. However, only PD and AD had moderate associations with the summary scores. Response mapping models for all the five dimensions had high prediction rates for L1 but inadequate predictive performance for L2 to L5. In fact, the estimated models for MO, SC, UA, and PD could not map the summary scores to L4-L5. Predictive performance did not improve when the models further included age and sex, or included the PedsQL item responses instead of the summary scores.

Although a large sample of 7,098 children and adolescents was used in this study, the estimated response mapping did not perform well. The ceiling effect (46.7%) in our sample of school-attending children and adolescents was comparable to the 40.3% observed in a sample of general pupil population in China¹⁹. The sample did not include sufficient observations of children with impaired health profile, especially for MO, SC, and UA, which had 93.4%, 95.7%, and 92.7% reporting no problems, respectively. The Y-5L level data lacked variability and could not produce meaningful mapping. Population of children with acute or chronic diseases would have a greater span in severity levels for better predictive performance in response mapping models, but they were not likely included in this healthy sample. Further recruitment efforts should focus on pediatric patients with acute or chronic conditions in clinical settings. Occasionally, pediatric patients with severe health conditions may not be able to complete the questionnaires independently. A common practice is to employ the proxy-reported version of Y-5L as an alternative, where parents or caregivers serve as proxies to report on the children's health status²⁰. Clinical samples of pediatric patients will clarify the associations between the two instruments and improve mapping quality.

Collaborative data sharing is an efficient option for mapping studies when recruitment alone could not provide sufficient samples. Researchers may combine data from different studies to enrich the dataset and ensure coverage of the dimensions and levels. Specifically, future mapping studies may apply pooled sampling of international data. Anonymized datasets in the centralized repository shall contain responses for the source and target instruments, demographic characteristics of the respondents including age, sex, country or region, and the disease or patient groups of the respondents. Details on the instruments used and data collection, such as the version, language, and mode of recruitment and administration (self-completed or

proxy-reported), shall also be included. These data are necessary for the identification and filling of data gaps which will help the synergy between existing datasets and additional data collection. This collaborative approach, while enhancing dataset comprehensiveness and representativeness, requires careful monitoring to ensure adherence to data transfer agreements, archiving standards, secure storage, and controlled access for collaborative utilization of valuable research data.

Direct mapping of PedsQL onto Y-5L utility score was not viable in this study due to the absence of Y-5L value set. When the Y-5L value set becomes available, predicted Y-5L utility score can be generated using the expected value method on the predicted probability of the levels for the five dimensions²¹, which can be used to assess model performance for response mapping. Direct mapping to the Y-5L utility score can also be estimated by various regression models. The most commonly used direct mapping model is the OLS model, where predicted utility scores greater than one are set to one. Other commonly used models included generalized linear models, Tobit models, and censored least absolute deviation model⁵. A previous study explored the application of machine learning methods in both direct and indirect mappings of the EQ-5D-5L from the non-preference-based Patient-Reported Outcomes Measurement Information System Global Health 10 (PROMIS-GH10), and reported that the least absolute shrinkage and selection operator (LASSO) out-performed other machine learning and econometric models²². The Y-5L value set will also facilitate other mapping studies, particularly the crosswalk model between Y-3L and Y-5L.

There were several limitations on the feasibility of mapping algorithms for Y-5L from PedsQL. First, there was a difference in the recall periods between Y-5L and PedsQL. In Y-5L, the respondents were asked to rate their health for the current day on a severity scale, while in PedsQL, they were asked about the frequency of problems encountered in the last month. Second, the psychosocial summary score consists of the social and school functioning subscales, which were not explicitly covered in the Y-5L descriptive system (Examples of “doing usual activities” include “going to school” and “doing things with family or friends”). The differences in recall periods and domains covered limited the conceptual overlapping necessary for an accurate mapping algorithm. Third, while age-specific modules of PedsQL were used in this study, potential heterogeneity in comprehension and developmental differences across

primary and secondary schools remained unaddressed. Fourth, data of Y-3L responses were not collected in this study. Otherwise, the Y-3L utility scores could be generated using the Y-3L value set for China²³, which would facilitate both direct and response mappings of Y-3L from PedsQL.

In this healthy cohort sampled from the school-attending children and adolescent in Hong Kong, our mapping algorithms demonstrated inadequate predictive performance in differentiating those with and without health problems, primarily due to insufficient clinical samples with worse health states. Future development of the Y-5L value set will facilitate direct mapping of the Y-5L utility score from the PedsQL. Study findings call for the collaborative efforts in collecting both PedsQL and Y-5L responses from clinical samples with impaired health profiles for future mapping studies. Nevertheless, researchers are recommended to collect data using preference-based instruments even if a mapping algorithm exists. Mapping can only serve as an alternative when data collection using preference-based instruments is not feasible.

Funding

This study was supported by the General Research Fund of the Research Grants Council (#17119518). No funding organization had any role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; and preparation of the manuscript.

References

1. Kwon J, Freijser L, Huynh E, et al. Systematic Review of Conceptual, Age, Measurement and Valuation Considerations for Generic Multidimensional Childhood Patient-Reported Outcome Measures. *Pharmacoeconomics* 2022; **40**(4): 379-431.
2. Kreimeier S, Greiner W. EQ-5D-Y as a Health-Related Quality of Life Instrument for Children and Adolescents: The Instrument's Characteristics, Development, Current Use, and Challenges of Developing Its Value Set. *Value Health* 2019; **22**(1): 31-7.
3. Varni JW, Seid M, Rode CA. The PedsQL: measurement model for the pediatric quality of life inventory. *Med Care* 1999; **37**(2): 126-39.
4. Brazier JE, Yang Y, Tsuchiya A, Rowen DL. A review of studies mapping (or cross walking) non-preference based measures of health to generic preference-based measures. *Eur J Health Econ* 2010; **11**(2): 215-25.

5. Dakin H, Abel L, Burns R, Koleva-Kolarova R, Yang Y. HERC database of mapping studies, Version 9.0. 1st December 2023. <http://www.herc.ox.ac.uk/downloads/herc-database-of-mapping-studies> (accessed 2nd February 2024).
6. Khan KA, Petrou S, Rivero-Arias O, Walters SJ, Boyle SE. Mapping EQ-5D utility scores from the PedsQL generic core scales. *Pharmacoeconomics* 2014; **32**(7): 693-706.
7. Wille N, Badia X, Bonsel G, et al. Development of the EQ-5D-Y: a child-friendly version of the EQ-5D. *Qual Life Res* 2010; **19**(6): 875-86.
8. Kreimeier S, Astrom M, Burstrom K, et al. EQ-5D-Y-5L: developing a revised EQ-5D-Y with increased response categories. *Qual Life Res* 2019; **28**(7): 1951-61.
9. Verstraete J, Scott D. Comparison of the EQ-5D-Y-5L, EQ-5D-Y-3L and PedsQL in children and adolescents. *J Patient Rep Outcomes* 2022; **6**(1): 67.
10. The Hong Kong Childhealth Foundation. Physical Education-School Physical Fitness Award Scheme. 2023. https://www.edb.gov.hk/en/curriculum-development/kla/pe/references_resource/spfas/ (accessed 2nd February 2024).
11. Ravens-Sieberer U, Wille N, Badia X, et al. Feasibility, reliability, and validity of the EQ-5D-Y: results from a multinational study. *Qual Life Res* 2010; **19**(6): 887-97.
12. Wong CKH, Cheung PWH, Luo N, Cheung JPY. A head-to-head comparison of five-level (EQ-5D-5L-Y) and three-level EQ-5D-Y questionnaires in paediatric patients. *Eur J Health Econ* 2019; **20**(5): 647-56.
13. Wong CKH, Cheung PWH, Luo N, Lin J, Cheung JPY. Responsiveness of EQ-5D Youth version 5-level (EQ-5D-5L-Y) and 3-level (EQ-5D-3L-Y) in Patients With Idiopathic Scoliosis. *Spine* 2019; **44**(21): 1507-14.
14. Xu RH, Zhu L, Sun R, Tan RL, Luo N, Zou S, Dong D. Investigating the psychometric properties of the EQ-5D-Y-3L, EQ-5D-Y-5L, CHU-9D, and PedsQL in children and adolescents with osteogenesis imperfecta. *Eur J Pediatr* 2022; **181**(12): 4049-58.
15. Varni JW, Seid M, Kurtin PS. PedsQL 4.0: reliability and validity of the Pediatric Quality of Life Inventory version 4.0 generic core scales in healthy and patient populations. *Med Care* 2001; **39**(8): 800-12.
16. Chan LF, Chow SM, Lo SK. Preliminary validation of the Chinese version of the Pediatric Quality of Life Inventory. *Int J Rehabil Res* 2005; **28**(3): 219-27.
17. Round J, Hawton A. Statistical Alchemy: Conceptual Validity and Mapping to Generate Health State Utility Values. *Pharmacoecon Open* 2017; **1**(4): 233-9.
18. Cohen J. Statistical power analysis for the behavioral sciences: Academic press; 2013.
19. Pei W, Yue S, Zhi-Hao Y, Ruo-Yu Z, Bin W, Nan L. Testing measurement properties of two EQ-5D youth versions and KIDSCREEN-10 in China. *The European Journal of Health Economics* 2021; **22**(7): 1083-93.
20. Lin J, Wong CKH, Cheung JPY, Cheung PWH, Luo N. Psychometric performance of proxy-reported EQ-5D youth version 5-level (EQ-5D-Y-5L) in comparison with three-level (EQ-5D-Y-3L) in children and adolescents with scoliosis. *Eur J Health Econ* 2022; **23**(8): 1383-95.
21. Le QA, Doctor JN. Probabilistic mapping of descriptive health status responses onto health state utilities using Bayesian networks: an empirical analysis converting SF-12 into EQ-5D utility index in a national US sample. *Med Care* 2011; **49**(5): 451-60.
22. Aghdaee M, Parkinson B, Sinha K, Gu Y, Sharma R, Olin E, Cutler H. An examination of machine learning to map non-preference based patient reported outcome measures to health state utility values. *Health Econ* 2022; **31**(8): 1525-57.
23. Yang Z, Jiang J, Wang P, et al. Estimating an EQ-5D-Y-3L Value Set for China. *Pharmacoeconomics* 2022; **40**(Suppl 2): 147-55.

Characteristics	Total (N=7,098)	Age 6-8 (N=1,713)	Age 9-11 (N=2,359)	Age 12-17 (N=3,026)	P-value
Age, mean (SD)	11.7 (3.2)	7.7 (0.7)	10.5 (0.8)	14.8 (1.7)	<0.001*
Gender, % (n)					
Male	44.4% (3,152)	47.7% (817)	47.2% (1,114)	40.4% (1,221)	<0.001*
Female	55.6% (3,946)	52.3% (896)	52.8% (1,245)	59.6% (1,805)	
EQ-5D-Y-5L					
Mobility, % (n)					
L1	93.4% (6,629)	93.5% (1,601)	93.6% (2,209)	93.2% (2,819)	0.794
L2	5.5% (392)	4.7% (81)	5.4% (127)	6.1% (184)	
L3	0.8% (56)	1.3% (22)	0.7% (17)	0.6% (17)	
L4	0.2% (15)	0.4% (6)	0.2% (5)	0.1% (4)	
L5	0.1% (6)	0.2% (3)	0.0% (1)	0.1% (2)	
Looking after myself, % (n)					
L1	95.7% (6,793)	87.9% (1,506)	97.0% (2,288)	99.1% (2,999)	<0.001*
L2	3.3% (237)	9.3% (160)	2.5% (59)	0.6% (18)	
L3	0.6% (43)	1.6% (27)	0.4% (9)	0.2% (7)	
L4	0.2% (15)	0.8% (13)	0.0% (1)	0.0% (1)	
L5	0.1% (10)	0.4% (7)	0.1% (2)	0.0% (1)	
Doing usual activities, % (n)					
L1	92.7% (6,583)	90.5% (1,551)	92.0% (2,170)	94.6% (2,862)	<0.001*
L2	6.0% (428)	7.1% (122)	6.9% (163)	4.7% (143)	
L3	0.9% (61)	1.3% (22)	1.0% (23)	0.5% (16)	
L4	0.2% (17)	0.7% (12)	0.0% (1)	0.1% (4)	
L5	0.1% (9)	0.4% (6)	0.1% (2)	0.0% (1)	
Having pain or discomfort, % (n)					
L1	71.9% (5,104)	76.2% (1,306)	73.5% (1,734)	68.2% (2,064)	<0.001*
L2	24.4% (1,733)	19.4% (332)	23.3% (550)	28.1% (851)	
L3	3.0% (211)	3.2% (55)	2.5% (60)	3.2% (96)	
L4	0.5% (32)	0.5% (9)	0.5% (11)	0.4% (12)	
L5	0.3% (18)	0.6% (11)	0.2% (4)	0.1% (3)	
Feeling worried, sad or unhappy, % (n)					
L1	57.5% (4,084)	68.8% (1,179)	59.6% (1,406)	49.5% (1,499)	<0.001*
L2	32.0% (2,272)	22.8% (390)	31.1% (733)	38.0% (1,149)	
L3	7.3% (518)	4.8% (83)	7.0% (165)	8.9% (270)	
L4	1.9% (135)	1.9% (32)	1.3% (30)	2.4% (73)	
L5	1.3% (89)	1.7% (29)	1.1% (25)	1.2% (35)	
Health state 11111, % (n)	46.7% (3,318)	53.6% (918)	49.6% (1,169)	40.7% (1,231)	<0.001*
EQ VAS score, mean (SD)	82.8 (18.0)	89.3 (17.5)	85.3 (16.9)	77.2 (17.4)	<0.001*
PedsQL, mean (SD)					
Total PedsQL score	80.9 (13.3)	83.1 (13.9)	83.9 (12.1)	77.3 (12.8)	<0.001*
Physical health summary score	85.2 (13.9)	86.6 (15.0)	88.2 (12.3)	82.2 (13.8)	<0.001*
Psychosocial health summary score	72.9 (19.8)	76.7 (20.4)	75.6 (19.0)	68.7 (19.2)	<0.001*
Emotional sub-scale score	84.6 (16.8)	85.7 (17.9)	87.1 (15.7)	82.0 (16.6)	<0.001*
Social sub-scale score	78.4 (16.4)	81.3 (17.2)	82.3 (14.7)	73.7 (16.0)	<0.001*
School sub-scale score	78.6 (14.8)	81.2 (15.4)	81.7 (13.7)	74.8 (14.3)	<0.001*

330 Table 2. Descriptive statistics of PedsQL summary scores by Y-5L levels

Y-5L	N	Total PedsQL score			Physical health summary score			Psychosocial health summary score		
		Mean	SD	P-value	Mean	SD	P-value	Mean	SD	P-value
MO										
L1	6,629	81.8	12.6		86.1	13.2		79.5	14.2	
L2	392	69.7	14.6		73.8	16.5		67.6	16.0	
L3	56	66.7	17.3	<0.001	69.6	18.9	<0.001	65.1	18.6	<0.001
L4	15	54.7	21.3		58.4	27.2		52.8	22.4	
L5	6	53.5	16.9		65.1	18.7		47.1	23.0	
SC										
L1	6,793	81.3	13.0		85.7	13.6		79.0	14.5	
L2	237	73.2	13.5		77.2	15.3		71.0	15.0	
L3	43	70.6	19.3	<0.001	73.5	19.1	<0.001	69.1	21.2	<0.001
L4	15	58.2	20.6		66.7	21.7		53.6	24.4	
L5	10	60.2	19.9		65.0	17.0		57.6	24.6	
UA										
L1	6,583	81.9	12.5		86.2	13.2		79.7	14.0	
L2	428	69.4	14.3		74.5	15.9		66.8	15.8	
L3	61	63.9	16.9	<0.001	71.7	16.9	<0.001	59.9	19.4	<0.001
L4	17	49.5	25.8		56.3	29.7		45.9	28.9	
L5	9	53.3	21.6		61.2	25.4		49.1	24.8	
PD										
L1	5,104	83.9	12.0		88.0	12.7		81.7	13.5	
L2	1,733	74.6	12.3		79.3	13.5		72.0	13.7	
L3	211	66.3	14.5	<0.001	72.5	15.8	<0.001	63.1	16.7	<0.001
L4	32	60.4	19.4		66.6	21.1		57.0	22.3	
L5	18	60.0	26.8		64.4	32.5		57.6	28.2	
AD										
L1	4,084	85.9	11.2		88.8	12.5		84.4	12.2	
L2	2,272	76.8	11.2		82.1	13.0		74.0	12.2	
L3	518	68.4	12.3	<0.001	76.6	14.6	<0.001	64.0	13.5	<0.001
L4	135	62.7	16.4		74.7	19.6		56.4	17.2	
L5	89	58.7	18.7		70.6	22.4		52.3	21.4	

331

332 Notes: P-value for trend was calculated by the Jonckheere–Terpstra test.

333 Table 3. Correlations between Y-5L and PedsQL scores and responses in overall and by age
334 groups

Overall (N=7,098)	MO	SC	UA	PD	AD
Total PedsQL score	-0.21**	-0.13**	-0.23**	-0.37**	-0.48**
Physical health summary score	-0.21**	-0.13**	-0.21**	-0.34**	-0.34**
Hard to walk; 100 m	-0.16**	-0.06**	-0.13**	-0.16**	-0.19**
Hard to run	-0.17**	-0.03*	-0.13**	-0.22**	-0.24**
Hard to participate in sport/exercise	-0.15**	-0.02	-0.15**	-0.21**	-0.23**
Hard to lift something heavy	-0.12**	-0.07**	-0.13**	-0.16**	-0.19**
Hard to bath/shower myself	-0.10**	-0.31**	-0.11**	-0.04*	-0.03*
Hard to do household chores	-0.13**	-0.16**	-0.16**	-0.11**	-0.12**
Pain or aches	-0.18**	-0.05**	-0.15**	-0.42**	-0.31**
Low energy levels	-0.15**	-0.07**	-0.17**	-0.30**	-0.32**
Psychosocial health summary score	-0.19**	-0.12**	-0.22**	-0.34**	-0.49**
Emotional sub-scale score	-0.16**	-0.09**	-0.18**	-0.33**	-0.53**
Afraid or scared	-0.13**	-0.08**	-0.14**	-0.25**	-0.43**
Sad	-0.12**	-0.04*	-0.13**	-0.28**	-0.49**
Angry	-0.12**	-0.07**	-0.13**	-0.23**	-0.36**
Trouble sleeping	-0.13**	-0.08**	-0.14**	-0.24**	-0.30**
Worry about what will happen to me	-0.11**	-0.04*	-0.11**	-0.25**	-0.45**
Social sub-scale score	-0.17**	-0.11**	-0.21**	-0.25**	-0.37**
Trouble getting along with other kids/teenagers	-0.13**	-0.08**	-0.15**	-0.20**	-0.32**
Other kids/teenagers do not want to be my friend	-0.14**	-0.08**	-0.16**	-0.19**	-0.29**
Other kids/teenagers tease me	-0.13**	-0.08**	-0.16**	-0.22**	-0.30**
Cannot do things others my age can do	-0.13**	-0.11**	-0.18**	-0.18**	-0.27**
Hard to keep up with others	-0.15**	-0.09**	-0.16**	-0.19**	-0.28**
School sub-scale score	-0.16**	-0.09**	-0.18**	-0.28**	-0.33**
Hard to pay attention in class	-0.14**	-0.07**	-0.17**	-0.24**	-0.31**
Forget things	-0.11**	-0.04*	-0.11**	-0.21**	-0.29**
Trouble keeping up with my schoolwork	-0.14**	-0.09**	-0.15**	-0.19**	-0.24**
Miss school because of not feeling well	-0.08**	-0.05**	-0.11**	-0.16**	-0.16**
Miss school to go doctor or hospital	-0.10**	-0.06**	-0.10**	-0.15**	-0.14**
Age 6-8 (N=1,713)	MO	SC	UA	PD	AD
Total PedsQL score	-0.23**	-0.26**	-0.28**	-0.35**	-0.40**
Physical health summary score	-0.21**	-0.25**	-0.24**	-0.32**	-0.29**
Hard to walk; 100 m	-0.19**	-0.10**	-0.15**	-0.22**	-0.23**
Hard to run	-0.18**	-0.14**	-0.17**	-0.19**	-0.17**
Hard to participate in sport/exercise	-0.19**	-0.12**	-0.19**	-0.15**	-0.15**
Hard to lift something heavy	-0.10**	-0.13**	-0.12**	-0.16**	-0.14**
Hard to bath/shower myself	-0.07*	-0.37**	-0.11**	-0.06*	-0.07*
Hard to do household chores	-0.14**	-0.21**	-0.14**	-0.12**	-0.11**
Pain or aches	-0.17**	-0.10**	-0.19**	-0.34**	-0.26**
Low energy levels	-0.17**	-0.16**	-0.19**	-0.28**	-0.24**
Psychosocial health summary score	-0.21**	-0.24**	-0.26**	-0.33**	-0.40**
Emotional sub-scale score	-0.19**	-0.20**	-0.22**	-0.33**	-0.44**

Afraid or scared	-0.14**	-0.17**	-0.15**	-0.24**	-0.35**
Sad	-0.15**	-0.13**	-0.16**	-0.29**	-0.38**
Angry	-0.15**	-0.15**	-0.14**	-0.20**	-0.27**
Trouble sleeping	-0.15**	-0.13**	-0.16**	-0.25**	-0.25**
Worry about what will happen to me	-0.13**	-0.14**	-0.15**	-0.22**	-0.37**
Social sub-scale score	-0.20**	-0.20**	-0.25**	-0.25**	-0.30**
Trouble getting along with other kids/teenagers	-0.15**	-0.16**	-0.17**	-0.17**	-0.22**
Other kids/teenagers do not want to be my friend	-0.14**	-0.15**	-0.18**	-0.16**	-0.20**
Other kids/teenagers tease me	-0.16**	-0.13**	-0.17**	-0.20**	-0.25**
Cannot do things others my age can do	-0.16**	-0.19**	-0.22**	-0.15**	-0.21**
Hard to keep up with others	-0.16**	-0.20**	-0.19**	-0.17**	-0.22**
School sub-scale score	-0.16**	-0.20**	-0.21**	-0.24**	-0.25**
Hard to pay attention in class	-0.12**	-0.17**	-0.21**	-0.22**	-0.23**
Forget things	-0.12**	-0.18**	-0.16**	-0.15**	-0.17**
Trouble keeping up with my schoolwork	-0.16**	-0.19**	-0.16**	-0.18**	-0.17**
Miss school because of not feeling well	-0.10**	-0.11**	-0.12**	-0.14**	-0.13**
Miss school to go doctor or hospital	-0.08*	-0.09**	-0.09**	-0.13**	-0.12**
Age 9-11 (N=2,359)	MO	SC	UA	PD	AD
Total PedsQL score	-0.22**	-0.17**	-0.26**	-0.37**	-0.48**
Physical health summary score	-0.22**	-0.17**	-0.22**	-0.36**	-0.36**
Hard to walk; 100 m	-0.15**	-0.10**	-0.14**	-0.17**	-0.16**
Hard to run	-0.21**	-0.12**	-0.15**	-0.23**	-0.22**
Hard to participate in sport/exercise	-0.17**	-0.12**	-0.16**	-0.21**	-0.22**
Hard to lift something heavy	-0.14**	-0.10**	-0.14**	-0.16**	-0.21**
Hard to bath/shower myself	-0.13**	-0.30**	-0.14**	-0.05*	-0.06*
Hard to do household chores	-0.08**	-0.14**	-0.17**	-0.11**	-0.13**
Pain or aches	-0.17**	-0.09**	-0.16**	-0.43**	-0.30**
Low energy levels	-0.17**	-0.13**	-0.19**	-0.31**	-0.32**
Psychosocial health summary score	-0.20**	-0.15**	-0.25**	-0.34**	-0.49**
Emotional sub-scale score	-0.18**	-0.13**	-0.21**	-0.31**	-0.50**
Afraid or scared	-0.13**	-0.12**	-0.15**	-0.22**	-0.40**
Sad	-0.12**	-0.11**	-0.16**	-0.25**	-0.45**
Angry	-0.12**	-0.07**	-0.16**	-0.20**	-0.35**
Trouble sleeping	-0.17**	-0.10**	-0.16**	-0.23**	-0.28**
Worry about what will happen to me	-0.13**	-0.10**	-0.15**	-0.25**	-0.42**
Social sub-scale score	-0.17**	-0.15**	-0.23**	-0.25**	-0.38**
Trouble getting along with other kids/teenagers	-0.11**	-0.12**	-0.17**	-0.18**	-0.29**
Other kids/teenagers do not want to be my friend	-0.15**	-0.10**	-0.18**	-0.17**	-0.28**
Other kids/teenagers tease me	-0.13**	-0.11**	-0.20**	-0.21**	-0.29**
Cannot do things others my age can do	-0.12**	-0.14**	-0.18**	-0.20**	-0.27**
Hard to keep up with others	-0.17**	-0.11**	-0.18**	-0.19**	-0.25**
School sub-scale score	-0.17**	-0.11**	-0.21**	-0.30**	-0.34**
Hard to pay attention in class	-0.16**	-0.13**	-0.20**	-0.25**	-0.29**
Forget things	-0.11**	-0.07**	-0.11**	-0.21**	-0.28**
Trouble keeping up with my schoolwork	-0.13**	-0.09**	-0.18**	-0.20**	-0.23**
Miss school because of not feeling well	-0.09**	-0.06*	-0.11**	-0.18**	-0.17**

Miss school to go doctor or hospital	-0.11**	-0.08**	-0.13**	-0.17**	-0.15**
Age 12-17 (N=3,026)	MO	SC	UA	PD	AD
Total PedsQL score	-0.20**	-0.07**	-0.24**	-0.36**	-0.50**
Physical health summary score	-0.20**	-0.06**	-0.22**	-0.32**	-0.32**
Hard to walk; 100 m	-0.16**	-0.03	-0.14**	-0.12**	-0.16**
Hard to run	-0.14**	-0.01	-0.16**	-0.20**	-0.21**
Hard to participate in sport/exercise	-0.13**	-0.01	-0.19**	-0.20**	-0.21**
Hard to lift something heavy	-0.12**	0.01	-0.16**	-0.16**	-0.19**
Hard to bath/shower myself	-0.12**	-0.15**	-0.09**	-0.02	-0.03
Hard to do household chores	-0.17**	-0.11**	-0.15**	-0.11**	-0.14**
Pain or aches	-0.20**	-0.07**	-0.17**	-0.44**	-0.30**
Low energy levels	-0.14**	-0.02	-0.18**	-0.29**	-0.33**
Psychosocial health summary score	-0.18**	-0.06**	-0.22**	-0.34**	-0.53**
Emotional sub-scale score	-0.13**	-0.04*	-0.17**	-0.32**	-0.57**
Afraid or scared	-0.13**	-0.04*	-0.15**	-0.27**	-0.48**
Sad	-0.11**	-0.01	-0.14**	-0.28**	-0.54**
Angry	-0.10**	-0.04*	-0.10**	-0.25**	-0.40**
Trouble sleeping	-0.09**	-0.04*	-0.13**	-0.22**	-0.34**
Worry about what will happen to me	-0.09**	-0.03	-0.11**	-0.25**	-0.47**
Social sub-scale score	-0.15**	-0.06**	-0.19**	-0.24**	-0.39**
Trouble getting along with other kids/teenagers	-0.13**	-0.06**	-0.17**	-0.21**	-0.35**
Other kids/teenagers do not want to be my friend	-0.14**	-0.06*	-0.15**	-0.21**	-0.32**
Other kids/teenagers tease me	-0.11**	-0.06**	-0.16**	-0.21**	-0.31**
Cannot do things others my age can do	-0.13**	-0.06*	-0.18**	-0.17**	-0.27**
Hard to keep up with others	-0.14**	-0.04*	-0.17**	-0.18**	-0.30**
School sub-scale score	-0.16**	-0.06**	-0.20**	-0.26**	-0.33**
Hard to pay attention in class	-0.16**	-0.05*	-0.18**	-0.23**	-0.31**
Forget things	-0.11**	0.00	-0.12**	-0.22**	-0.30**
Trouble keeping up with my schoolwork	-0.14**	-0.07**	-0.16**	-0.17**	-0.24**
Miss school because of not feeling well	-0.06**	-0.04*	-0.11**	-0.15**	-0.15**
Miss school to go doctor or hospital	-0.10**	-0.07**	-0.11**	-0.14**	-0.13**

335 Note: *P<0.05 and **P<0.001 (2-tailed). Association was considered low for 0.10–0.29,
336 moderate for 0.30–0.49 (cells highlighted yellow), and high for 0.50–1.00 (cells highlighted
337 orange). Correlations with Y-5L items are negative as a higher PedsQL score indicates a better
338 HRQoL.

339

340 Table 4. Prediction rates of ordered logistic regression of Y-5L levels on physical and
341 psychosocial health summary scores in overall and by age groups

Overall (N=7,098)				Age 6-8 (N=1,713)			Age 9-11 (N=2,359)			Age 12-17 (N=3,026)		
MO	N	Correct prediction		N	Correct prediction		N	Correct prediction		N	Correct prediction	
		n	%		n	%		n	%		n	%
L1	6,629	6,614	99.8%	1,601	1,599	99.9%	2,209	2,204	99.8%	2,819	2,814	99.8%
L2	392	3	0.8%	81	1	1.2%	127	1	0.8%	184	2	1.1%
L3	56	0	0.0%	22	0	0.0%	17	0	0.0%	17	0	0.0%
L4	15	0	0.0%	6	0	0.0%	5	0	0.0%	4	0	0.0%
L5	6	0	0.0%	3	0	0.0%	1	0	0.0%	2	0	0.0%
Total	7,098	6,617	93.2%	1,713	1,600	93.4%	2,359	2,205	93.5%	3,026	2,816	93.1%
SC	N	Correct prediction		N	Correct prediction		N	Correct prediction		N	Correct prediction	
		n	%		n	%		n	%		n	%
L1	6,793	6,793	100.0%	1,506	1,498	99.5%	2,288	2,286	99.9%	2,999	2,999	100.0%
L2	237	0	0.0%	160	0	0.0%	59	0	0.0%	18	0	0.0%
L3	43	0	0.0%	27	0	0.0%	9	0	0.0%	7	0	0.0%
L4	15	0	0.0%	13	0	0.0%	1	0	0.0%	1	0	0.0%
L5	10	0	0.0%	7	0	0.0%	2	0	0.0%	1	0	0.0%
Total	7,098	6,793	95.7%	1,713	1,498	87.4%	2,359	2,286	96.9%	3,026	2,999	99.1%
UA	N	Correct prediction		N	Correct prediction		N	Correct prediction		N	Correct prediction	
		n	%		n	%		n	%		n	%
L1	6,583	6,565	99.7%	1,551	1,541	99.4%	2,170	2,165	99.8%	2,862	2,852	99.7%
L2	428	4	0.9%	122	0	0.0%	163	3	1.8%	143	4	2.8%
L3	61	0	0.0%	22	0	0.0%	23	0	0.0%	16	0	0.0%
L4	17	0	0.0%	12	0	0.0%	1	0	0.0%	4	0	0.0%
L5	9	0	0.0%	6	0	0.0%	2	0	0.0%	1	0	0.0%
Total	7,098	6,569	92.5%	1,713	1,541	90.0%	2,359	2,168	91.9%	3,026	2,856	94.4%
PD	N	Correct prediction		N	Correct prediction		N	Correct prediction		N	Correct prediction	
		n	%		n	%		n	%		n	%
L1	5,104	4,859	95.2%	1,306	1,269	97.2%	1,734	1,637	94.4%	2,064	1,934	93.7%
L2	1,733	268	15.5%	332	27	8.1%	550	94	17.1%	851	164	19.3%
L3	211	1	0.5%	55	0	0.0%	60	0	0.0%	96	1	1.0%
L4	32	0	0.0%	9	0	0.0%	11	0	0.0%	12	0	0.0%
L5	18	0	0.0%	11	0	0.0%	4	0	0.0%	3	0	0.0%
Total	7,098	5,128	72.2%	1,713	1,296	75.7%	2,359	1,731	73.4%	3,026	2,099	69.4%
AD	N	Correct prediction		N	Correct prediction		N	Correct prediction		N	Correct prediction	
		n	%		n	%		n	%		n	%
L1	4,084	3,392	83.1%	1,179	1,117	94.7%	1,406	1,215	86.4%	1,499	1,094	73.0%
L2	2,272	965	42.5%	390	54	13.8%	733	271	37.0%	1,149	656	57.1%
L3	518	30	5.8%	83	0	0.0%	165	10	6.1%	270	24	8.9%
L4	135	0	0.0%	32	0	0.0%	30	0	0.0%	73	1	1.4%
L5	89	8	9.0%	29	3	10.3%	25	2	8.0%	35	3	8.6%
Total	7,098	4,395	61.9%	1,713	1,174	68.5%	2,359	1,498	63.5%	3,026	1,778	58.8%

343 Supplementary Table 1. Distributions of age and sex among respondents included in the study
 344 and the general population in Hong Kong

	Sample (N=7,098)	General population in Hong Kong (mid-year) (N=686,200)	SMD
Age, mean (SD)	11.7 (3.2)	11.3 (3.5)	0.10
Age group, % (n)			
6-8	24.1% (1,713)	28.0% (192,000)	
9-11	33.2% (2,359)	24.6% (168,600)	0.19
12-17	42.6% (3,026)	47.4% (325,600)	
Gender, % (n)			
Male	44.4% (3,152)	51.6% (354,100)	
Female	55.6% (3,946)	48.4% (332,100)	0.14

345

346 Notes: SD = standard deviation; SMD = standardized mean difference

347 Numbers of people in the Hong Kong general population were rounded to the nearest thousand.

348 Source: https://www.censtatd.gov.hk/en/web_table.html?id=110-01002

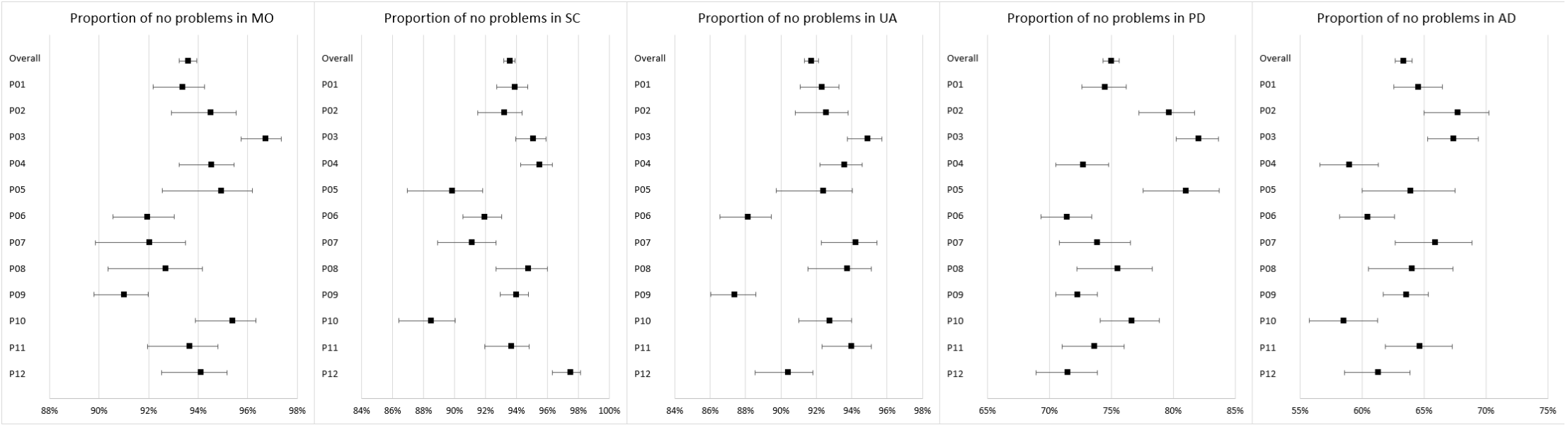
349 Supplementary Table 2. Prediction rates of ordered logistic regression of Y-5L levels on (i)
350 physical and physical health summary score, psychosocial health summary score, age, and sex,
351 and (ii) all PedsQL responses

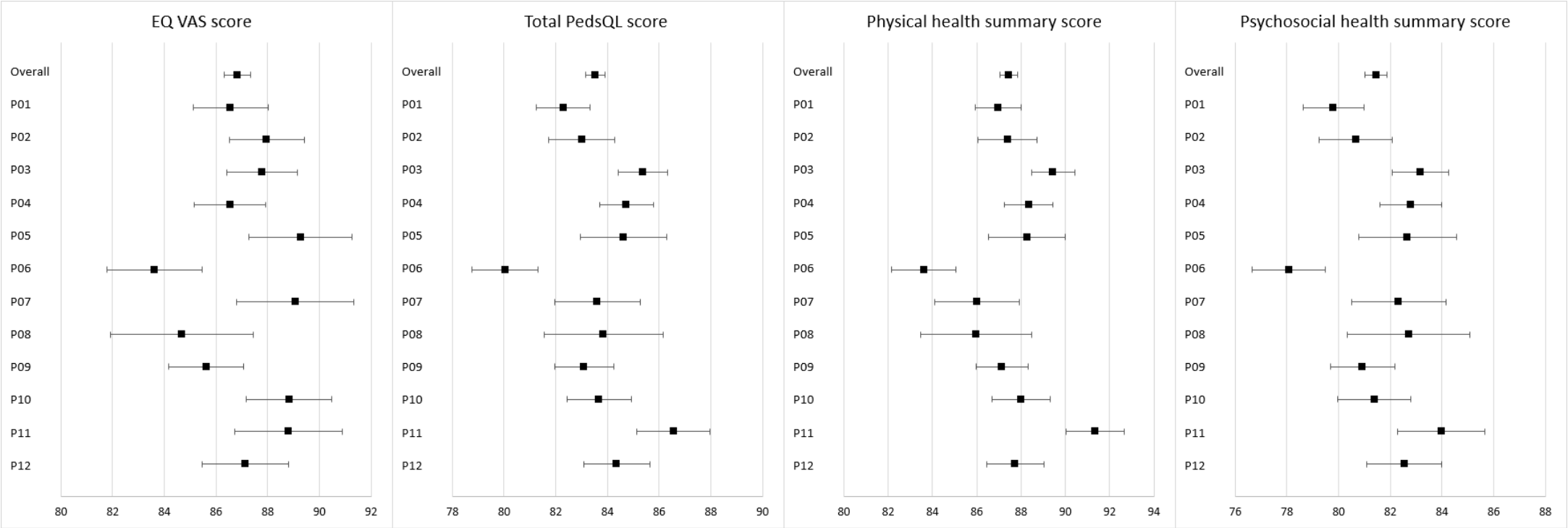
(i) Physical health summary score, psychosocial health summary score, age, and sex												
Overall (N=7,098)				Age 6-8 (N=1,713)			Age 9-11 (N=2,359)			Age 12-17 (N=3,026)		
MO	N	Correct prediction		N	Correct prediction		N	Correct prediction		N	Correct prediction	
		n	%		n	%		n	%		n	%
L1	6,629	6,610	99.7%	1,601	1,600	99.9%	2,209	2,205	99.8%	2,819	2,812	99.8%
L2	392	3	0.8%	81	1	1.2%	127	1	0.8%	184	3	1.6%
L3	56	0	0.0%	22	0	0.0%	17	0	0.0%	17	0	0.0%
L4	15	0	0.0%	6	0	0.0%	5	0	0.0%	4	0	0.0%
L5	6	0	0.0%	3	0	0.0%	1	0	0.0%	2	0	0.0%
Total	7,098	6,613	93.2%	1,713	1,601	93.5%	2,359	2,206	93.5%	3,026	2,815	93.0%
SC	N	Correct prediction		N	Correct prediction		N	Correct prediction		N	Correct prediction	
		n	%		n	%		n	%		n	%
L1	6,793	6,781	99.8%	1,506	1,496	99.3%	2,288	2,286	99.9%	2,999	2,999	100.0%
L2	237	0	0.0%	160	0	0.0%	59	0	0.0%	18	0	0.0%
L3	43	0	0.0%	27	0	0.0%	9	0	0.0%	7	0	0.0%
L4	15	0	0.0%	13	0	0.0%	1	0	0.0%	1	0	0.0%
L5	10	0	0.0%	7	0	0.0%	2	0	0.0%	1	0	0.0%
Total	7,098	6,781	95.5%	1,713	1,496	87.3%	2,359	2,286	96.9%	3,026	2,999	99.1%
UA	N	Correct prediction		N	Correct prediction		N	Correct prediction		N	Correct prediction	
		n	%		n	%		n	%		n	%
L1	6,583	6,557	99.6%	1,551	1,541	99.4%	2,170	2,162	99.6%	2,862	2,852	99.7%
L2	428	11	2.6%	122	0	0.0%	163	5	3.1%	143	7	4.9%
L3	61	0	0.0%	22	0	0.0%	23	0	0.0%	16	0	0.0%
L4	17	0	0.0%	12	1	8.3%	1	0	0.0%	4	0	0.0%
L5	9	0	0.0%	6	0	0.0%	2	0	0.0%	1	0	0.0%
Total	7,098	6,568	92.5%	1,713	1,542	90.0%	2,359	2,167	91.9%	3,026	2,859	94.5%
PD	N	Correct prediction		N	Correct prediction		N	Correct prediction		N	Correct prediction	
		n	%		n	%		n	%		n	%
L1	5,104	4,857	95.2%	1,306	1,269	97.2%	1,734	1,639	94.5%	2,064	1,928	93.4%
L2	1,733	266	15.3%	332	36	10.8%	550	92	16.7%	851	165	19.4%
L3	211	1	0.5%	55	0	0.0%	60	0	0.0%	96	1	1.0%
L4	32	0	0.0%	9	0	0.0%	11	0	0.0%	12	0	0.0%
L5	18	0	0.0%	11	0	0.0%	4	0	0.0%	3	0	0.0%
Total	7,098	5,124	72.2%	1,713	1,305	76.2%	2,359	1,731	73.4%	3,026	2,094	69.2%
AD	N	Correct prediction		N	Correct prediction		N	Correct prediction		N	Correct prediction	
		n	%		n	%		n	%		n	%
L1	4,084	3,407	83.4%	1,179	1,115	94.6%	1,406	1,219	86.7%	1,499	1,084	72.3%
L2	2,272	976	43.0%	390	69	17.7%	733	277	37.8%	1,149	650	56.6%
L3	518	31	6.0%	83	0	0.0%	165	11	6.7%	270	25	9.3%
L4	135	0	0.0%	32	0	0.0%	30	0	0.0%	73	1	1.4%
L5	89	7	7.9%	29	3	10.3%	25	2	8.0%	35	3	8.6%
Total	7,098	4,421	62.3%	1,713	1,187	69.3%	2,359	1,509	64.0%	3,026	1,763	58.3%

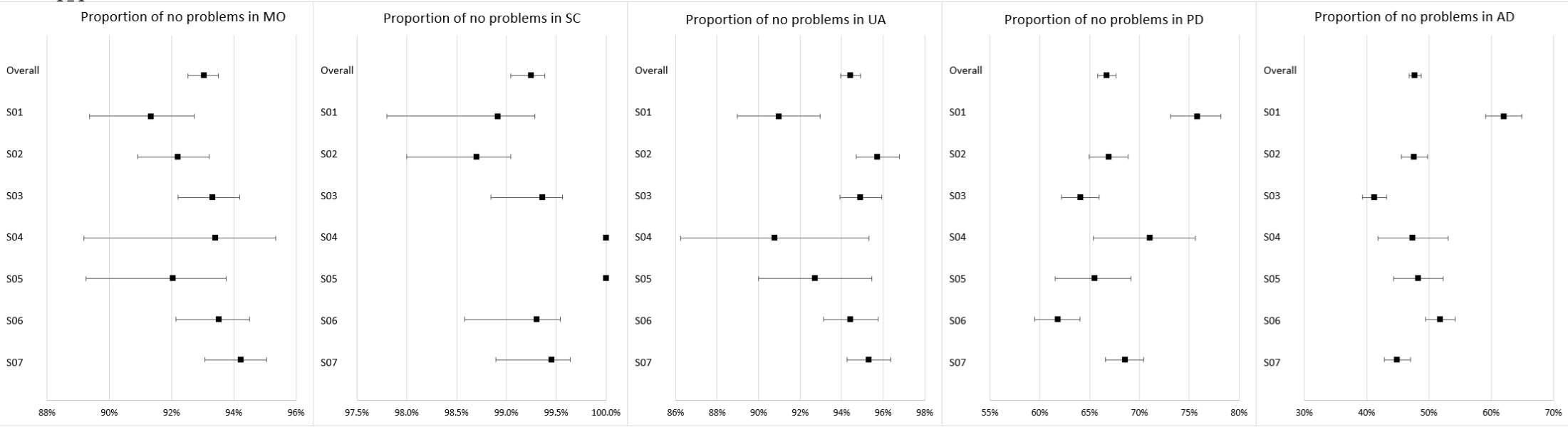
(ii) All PedsQL responses

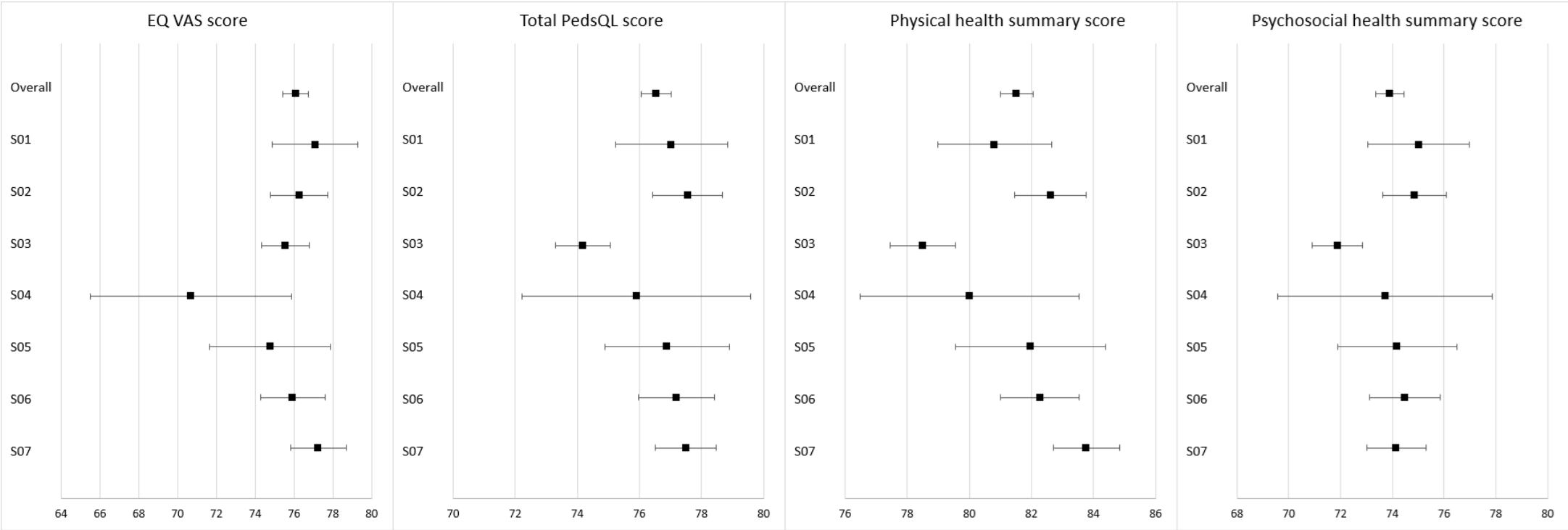
Overall (N=6,777)												
Age 6-8 (N=1,564)				Age 9-11 (N=2,250)				Age 12-17 (N=2,963)				
MO	N	Correct prediction		N	Correct prediction		N	Correct prediction		N	Correct prediction	
		n	%		n	%		n	%		n	%
L1	6,329	6,316	99.8%	1,463	1,461	99.9%	2,106	2,098	99.6%	2,760	2,752	99.7%
L2	378	7	1.9%	76	1	1.3%	121	4	3.3%	181	8	4.4%
L3	52	0	0.0%	18	0	0.0%	17	0	0.0%	17	0	0.0%
L4	13	0	0.0%	5	0	0.0%	5	0	0.0%	3	0	0.0%
L5	5	0	0.0%	2	0	0.0%	1	0	0.0%	2	0	0.0%
Total	6,777	6,323	93.3%	1,564	1,462	93.5%	2,250	2,102	93.4%	2,963	2,760	93.1%
SC	N	Correct prediction		N	Correct prediction		N	Correct prediction		N	Correct prediction	
		n	%		n	%		n	%		n	%
L1	6,493	6,468	99.6%	1,375	1,354	98.5%	2,181	2,176	99.8%	2,937	2,937	100.0%
L2	223	7	3.1%	149	12	8.1%	57	6	10.5%	17	0	0.0%
L3	40	0	0.0%	24	0	0.0%	9	0	0.0%	7	0	0.0%
L4	12	0	0.0%	10	0	0.0%	1	0	0.0%	1	0	0.0%
L5	9	0	0.0%	6	1	16.7%	2	0	0.0%	1	0	0.0%
Total	6,777	6,475	95.5%	1,564	1,367	87.4%	2,250	2,182	97.0%	2,963	2,937	99.1%
UA	N	Correct prediction		N	Correct prediction		N	Correct prediction		N	Correct prediction	
		n	%		n	%		n	%		n	%
L1	6,293	6,270	99.6%	1,422	1,414	99.4%	2,071	2,063	99.6%	2,800	2,792	99.7%
L2	408	13	3.2%	112	4	3.6%	154	5	3.2%	142	10	7.0%
L3	55	0	0.0%	17	0	0.0%	22	0	0.0%	16	0	0.0%
L4	13	0	0.0%	8	0	0.0%	1	0	0.0%	4	0	0.0%
L5	8	0	0.0%	5	0	0.0%	2	0	0.0%	1	0	0.0%
Total	6,777	6,283	92.7%	1,564	1,418	90.7%	2,250	2,068	91.9%	2,963	2,802	94.6%
PD	N	Correct prediction		N	Correct prediction		N	Correct prediction		N	Correct prediction	
		n	%		n	%		n	%		n	%
L1	4,878	4,542	93.1%	1,196	1,144	95.7%	1,658	1,559	94.0%	2,024	1,849	91.4%
L2	1,656	508	30.7%	304	48	15.8%	520	167	32.1%	832	318	38.2%
L3	194	2	1.0%	45	1	2.2%	57	1	1.8%	92	2	2.2%
L4	31	0	0.0%	8	0	0.0%	11	1	9.1%	12	0	0.0%
L5	18	0	0.0%	11	1	9.1%	4	0	0.0%	3	0	0.0%
Total	6,777	5,052	74.5%	1,564	1,194	76.3%	2,250	1,728	76.8%	2,963	2,169	73.2%
AD	N	Correct prediction		N	Correct prediction		N	Correct prediction		N	Correct prediction	
		n	%		n	%		n	%		n	%
L1	3,897	3,280	84.2%	1,086	1,019	93.8%	1,340	1,145	85.4%	1,471	1,138	77.4%
L2	2,186	1,100	50.3%	357	87	24.4%	702	316	45.0%	1,127	698	61.9%
L3	492	58	11.8%	73	2	2.7%	155	18	11.6%	264	47	17.8%
L4	119	0	0.0%	23	0	0.0%	28	0	0.0%	68	5	7.4%
L5	83	16	19.3%	25	2	8.0%	25	6	24.0%	33	4	12.1%
Total	6,777	4,454	65.7%	1,564	1,110	71.0%	2,250	1,485	66.0%	2,963	1,892	63.9%

353 Supplementary Figure 1. Distributions of respondents with no problems in Y-5L dimensions, EQ VAS score, and PedsQL summary scores by (a)
354 primary schools and (b) secondary schools
355 (a) Primary schools









371 Notes: Error bars represent 95% confidence intervals for the proportions of respondents with no problems in the Y-5L dimensions, the mean EQ
372 VAS score, and mean PedsQL summary scores.

373 P01 to P12 and S01 to S07 represent the twelve primary schools and seven secondary schools, respectively.