

**Work in progress: please do not quote without authors' permission**

## **Measuring and valuing health benefits for economic evaluation in adolescence: an assessment of the practicality and validity of the Child Health Utility 9D in the Australian adolescent population**

**Katherine Stevens<sup>1</sup>, BSc, MSc, PhD, Julie Ratcliffe<sup>2</sup>, BA, MSc, PhD**

<sup>1</sup>Health Economics and Decision Science, School of Health and Related Research, University of Sheffield, UK

<sup>2</sup> Flinders Clinical Effectiveness, Flinders University, Adelaide, South Australia

### **Address for correspondence:**

Katherine Stevens

Health Economics and Decision Science

School of Health and Related Research, University of Sheffield,

Regent Court, 30 Regent Street

Sheffield S1 4DA, UK

Telephone: 0114 222 0841

Fax: 0114 272 4095

Email: [K.Stevens@Sheffield.ac.uk](mailto:K.Stevens@Sheffield.ac.uk)

**Acknowledgments:** This study was supported by a Flinders University seeding grant

**Ethics statement:** This study has been approved by the Social and Behavioural Research Ethics Committee, Flinders University, project number: 4701

**Key words:** adolescent; utility; health outcomes; validation

**Running title:** Validation of the CHU9D in adolescents

## Abstract

**Objectives:** The Child Health Utility 9D (CHU9D) is a new generic preference based measure of health related quality of life developed for children age 7-11 years. There is increasing interest in the potential for application in adolescents and previous research has demonstrated it shows good construct validity here. This paper further examines the practicality and validity in adolescents by comparing with the KIDSCREEN10, a short generic measure for assessing children and adolescents' health-related quality of life and well-being.

**Methods:** A web based survey was administered to 961 consenting adolescents, including the CHU9D, a general health question, questions on the presence of long standing illness, disability or medical conditions, socio demographic variables and the KIDSCREEN10. The practicality, face and construct validity of the CHU9D were examined and the CHU9D and KIDSCREEN10 were compared in terms of their coverage, correlations between dimensions and overall scores.

**Results:** Both measures demonstrated good practicality and validity. The strongest degree of correlation was found with the only dimension in common for the CHU9D and the KIDSCREEN (sad). The lowest correlations were found between all the CHU9D dimensions and the 'have you had enough time for yourself' dimension of the KIDSCREEN10.

**Conclusion:** The findings from this study provide further support for the practicality and validity of the application of the CHU9D in the economic evaluation of adolescent health care and public health programmes.

Further research to test the psychometric performance of the CHU9D in more diverse clinical samples of adolescents is desirable including tests of reliability.

## Introduction

Economic evaluation of child and adolescent health care and public health interventions is receiving increasing interest for example in the prevention of obesity and the role that economic evaluation can play [1] and until recently there has been a paucity of reliable and valid instruments available to measure self-reported health status in this age group, in particular preference based measures (PBM) which allow the calculation of quality adjusted life years (QALYs) [2]. PBMs offers a key advantage over many other existing measures of paediatric and adolescent quality of life in that they allow the calculation of quality adjusted life years (QALYs), enabling use in cost utility analysis which allows for comparison both within and across clinical areas [3]. Those that do not allow the calculation of QALYs such as the PedsQoL [4] and the KIDSCREEN10 [5], although widely validated, are more limited in terms of their usefulness in economic evaluation and health care resource allocation decision making. Historically, no PBM has been demonstrated to be valid and reliable in children of all ages, meaning that often a wide variety of measures and approaches have been used, including the use of adult measures, proxy measures or even expert opinion [2]. Proxy measures are not desirable as the ideal in health status measurement is for the patient to self complete as it is increasingly recognized in clinical trials and health services research that descriptions of the experience of a health state should be elicited from the patients in order to reflect the actual experience of the disease and its treatment [6]. In addition, if a child can provide reliable and valid data, then self report is the optimal [7]. Adult preference based measures such as the EQ5D [8] or the SF6D [9] may not be suitable for young people as they may contain dimensions which are irrelevant for this age group or may not include those that are relevant and hence not capture important health related quality of life (HRQoL) effects of treatments or interventions [10]. There is a growing recognition that the use of adult instruments may not be appropriate in children and adolescents, and recently work has been undertaken to develop new instruments specifically for the population in question [2].

Whilst a number of non- preference based measures of HRQoL for children and adolescents (for example the PEDSQoL [4]) have been developed and are widely utilised internationally until recently the only PBM suitable for application in the calculation of QALYs for economic evaluation for this age group was the Health Utilities Index 2 (HUI2) [10]. Since then 3 more PBM have been developed for application in children and young people, the EQ5DY [11],



the AQoL6D [12] and the Child Health Utility 9D (CHU9D) [10]. The EQ5DY has been derived from the existing adult measure, the EQ5D [8] and conceptually contains the same 5 dimensions, the only difference being that the language has been adapted so that it can be understood by young people [11]. The developers recognise that it may not be applicable to apply the existing adult preference weights in order to calculate QALYs and they also recognise that work needs to be done in order to ensure the content validity for this age group i.e. whether the measure contains all aspects of HRQoL that are important to children and adolescents [11]. It is essential to demonstrate that the descriptive system is valid and reliable prior to the development of preference weights [11]. The AQoL6D was also adapted from an adult PBM. A previous article in this journal reported upon this adaptation process (which included some reduction of dimensions by removing those deemed to be unsuitable for adolescents and slight changes in phraseology to others) and the development of reference weights for the AQoL6D directly from adolescents [12].

The CHU9D is a new generic preference based measure of HRQoL for application in the economic evaluation of health care interventions in young people. In contrast to other PBMs for application in this age group the CHU9D was developed from scratch rather than representing an adaptation of an existing instrument and was developed exclusively with children [10]. It contains 9 dimensions, each with 5 levels and is designed to be self-completed by the young person [13]. It was originally developed for application with children in the 7-11 year old age group [10] but there is increasing interest in the potential for its application with children in older age groups. Recently, work has been undertaken to test its validity in an Australian adolescent population in relation to the existing widely used utility measure, the Health Utilities Index 2 (HUI2) [14]. It was found that the CHU9D demonstrated good practicality, face and construct validity in this population.

The KIDSCREEN10 is a generic measure of HRQoL and well-being for application with young people aged 8 to 18 years which was developed simultaneously in several different countries. It is designed as a self-report measure which is applicable for both health and chronically ill young people. Due to its development and testing in many countries [16], it is viewed as a cross-national measure. The KIDSCREEN10 index was derived from the longer 27 item version which was in turn derived from the 52 item version. It allows the calculation of a global HRQoL score and usually takes 5 minutes for completion. The measure has

demonstrated good internal consistency reliability and test-retest reliability and has been shown to be able to differentiate between different groups [17].

Although the KIDSCREEN10, in its present form is not suitable for application in economic evaluation, it has been widely validated across several European countries with children and adolescents aged 8-18 years with a sample size of 22,830 [16]. This means that it is a useful instrument for comparison when assessing the validity of other instruments which may be applicable in this age group.

In this paper, we continue the work in testing the validity of the CHU9D in an adolescent population undertaken by Ratcliffe et al [14] by further examining the practicality and validity of the CHU9D in relation to the KIDSCREEN10. We also compare the CHU9D and KIDSCREEN10 in terms of their coverage and correlation.

## **Methods**

A web based survey was developed for administration to a community based sample of adolescents living in Australia, aged 11-17 years. The survey was administered in collaboration with an independent market research company who have an existing on-line panel of parents who have given approval for their children to participate in research studies and have provided information on their home address, age and gender. Parent and adolescent dyad consent was required prior to participation in the survey. Once consent from both parties was obtained, the adolescent was asked to complete a survey which included the CHU9D, a self reported general health question with 5 response options (excellent, very good, good, fair and poor), socio demographic variables (gender, age and socio economic status as measured by the Family Affluence Scale), the KIDSCREEN10 instrument and whether they had a long standing disability, illness or medical condition. In addition, participants were asked to answer a series of Best Worst choice discrete choice experiment questions which we have reported previously in a separate paper [18]. The time taken to complete the survey was recorded as well as the respondents' rating of how difficult they found the survey was to complete.

The individual responses to the 9 CHU9D questions were converted to utilities (on the 0-1 dead perfect health QALY scale) using the existing UK adult general population algorithm



developed by Stevens [19] which is based on the standard gamble method of health state valuation. The responses to the KIDSCREEN10 questions were converted to a score by assigning Rasch person parameters to each possible sum score. The person parameters were transformed into values with a mean of 50 and a standard deviation of approximately 10 [20]. A low score indicates a poor HRQoL and a high score indicates good HRQoL.

The practicality and face validity of the CHU9D were examined by assessing the time taken to complete the survey, the response and completions rates and how difficult respondents found the survey. Construct validity was examining the extent to which the CHU9D was able to discriminate between groups with known health differences. We tested to see if the CHU9D could discriminate by the adolescent's overall rating of their general health, whether it could discriminate between those who had a long standing disability, illness or medical condition or not and whether it could discriminate between those who had a KIDSCREEN10 score lower and higher than the median by comparing the mean utility of these two groups. Those respondents who rated their health higher on the general health rating question would be expected to have higher utility scores on the CHU9D. For example, those rating their health as excellent should have a higher score than those rating their health as good. This measure of self rated health has been previously used in large surveys of adolescent health and has been shown to be a valid measure of subjective health [21]. Similarly, adolescents who have a long term disability, illness or medical condition would be expected to have a lower utility score than those who don't and the group who have a KIDSCREEN10 score lower than the median would be expected to have a lower mean utility. We tested all these differences to see if they were statistically significant.

The relationship between the CHU9D and KIDSCREEN10 and a number of socio-demographic variables (gender, age and socio-economic status) was also investigated. Respondents were split according to age in years and a one way Anova test was performed to test for significant difference in CHU9D and KIDSCREEN10 scores. Differences in CHU9D and KIDSCREEN10 by gender were tested for by using a Mann Whitney U test.

Socio-economic status was measured by the Family Affluence Scale (FAS). This is a validated measure of socioeconomic position which is designed for self report by adolescents aged 11-17 years. The instrument includes four items relating to family affluence: (1) "Does your family own a car, van or truck?" (no/yes one/yes two or more), (2)

“Do you have your own bedroom for yourself?” (no/yes), (3) “During the past 12 months how many times did you travel away on holiday (vacation) with your family?” (not at all/once/twice/more than twice), (4) “How many computers does your family own?” (none/one/two/more than two). The score is calculated on a 0-7 point scale with 1 point each for having one car, one computer and one room and one extra point each for having more than one car, holiday or computer. A lower score represents a lower level of affluence and vice versa [22]. The FAS score was grouped into 3 categories – those less than or equal to 3, those with FAS=4 or 5 and those with FAS greater than or equal to 6. We tested for differences between the groups and expected to find a difference as there is evidence that affluence level is related to quality of life [23].

We tested to see if there was any difference in dimension levels between those who had and those who hadn't got a long term disability, illness or medical condition and between those who had a KIDSCREEN10 score less than and greater than or equal to the median. The differences were tested using a Mann Whitney test.

The relationship between the CHU9D and the KIDSCREEN10 was explored by comparing the scores derived from the two instruments by graphical means (scatter plot) and looking at the correlation between the two scores using Spearman's Rho. The content and coverage of the two instruments was compared by examining the correlation between individual CHU9D dimensions and KIDSCREEN10 dimensions. In addition, we looked at the correlations between each CHU9D dimension and overall general health. Expected correlations between CHU9D and KIDSCREEN10 were as follows:

KIDSCREEN10 variable

Felt sad

Felt full of energy

Able to do the things you want

Got on well at school

CHU9D variable

Sad

Tired

Able to join in activities

School Work/Homework

## Results

961 adolescents consented to take part in the survey. Both instruments exhibited good completion rates with 630 (65.6%) complete responses for the KIDSCREEN10 and 634 (66%) complete responses for the CHU9D. Of those who started to complete the CHU9D (n=636), only 2 (0.3%) did not complete it. The median time to complete the survey was 7 minutes. This time includes both the CHU9D, the KIDSCREEN10 and the best worst discrete choice experiment exercise.

Table 1 shows the characteristics of the sample and Table 2 shows how difficult the respondents found the survey to complete.

The mean CHU9D utility was 0.85 and the mean KIDSCREEN10 score was 43.65. No significant difference was found between either the CHU9D and KIDSCREEN10 scores by gender. There was a significant difference in utility for the CHU9D by age ( $p=0.006$ ) with utility generally decreasing as age increased. A similar trend was found for the KIDSCREEN10, although this was not significant. Table 3 shows the mean CHU9D and KIDSCREEN10 scores by FAS groups. The higher the level of affluence, the higher the utility score and KIDSCREEN10 score. These differences were significant for both measures at  $p<0.01$ .

Table 4 shows the mean CHU9D utility score and KIDSCREEN10 score by the level of self reported general health. The differences between the levels of self reported general health were significant for both the CHU9D and the KIDSCREEN10 ( $p<0.001$ ) demonstrating evidence of construct validity for both the CHU9D and the KIDSCREEN10.

Table 5 shows the mean CHU9D and KIDSCREEN10 scores by whether the participant had a long standing disability, illness, or medical condition or not. These differences were significant for both the CHU9D ( $p=0.003$ ) and the KIDSCREEN10 ( $p=0.001$ ) demonstrating that both measures are able to distinguish between groups with known health differences, again giving evidence of construct validity.

Table 6 shows the mean and median utility values for those who had a KIDSCREEN less than the median value and those who had a KIDSCREEN10 score greater than or equal to

the median value. The median KIDSCREEN10 score for this population was 42.27. The difference between these groups was significant ( $p < 0.005$ ).

Table 7 shows the distributions across the levels of the CHU9D dimensions by whether the adolescent had reported that they had a long standing disability, illness or medical condition or not and whether there were significant differences between the 2 groups. All dimensions apart from the 3 emotional ones (worried, sad, annoyed) showed significant differences. All dimensions showed significant differences between groups for whether the KIDSCREEN10 was higher or lower than the median. The distributions are not reported.

Table 8 shows the mean KIDSCREEN10 score by each level of each CHU9D dimension. In general, mean KIDSCREEN index scores corresponded well with the dimensions of the CHU9D with increasing levels of severity on each dimension being associated with lower mean KIDSCREEN index scores.

### **Correlations**

Figure 1 shows a graphical comparison of CHU9D and KIDSCREEN10 scores. The CHU9D and the KIDSCREEN10 show a substantial degree of correlation (Spearman's  $\rho = 0.61$ ), significant at  $p = 0.01$ .

Table 9 shows the correlation of each of the CHU9D dimensions with the KIDSCREEN10 score. All correlations were significant at the 0.01 level. The correlations are negative as the CHU9D dimensions are labelled such that 1 is the highest level and 5 is the lowest level, hence you would expect a negative correlation with the KIDSCREEN10 score. All dimensions apart from pain, tired and School Work/Homework show a moderate degree of correlation.

### **Correlations between CHU9D and KIDSCREEN10 dimensions**

The strongest degree of correlation was found with the only dimension in common for the CHU9D and the KIDSCREEN (sad) (Spearman's  $\rho = -0.496$ ,  $p < 0.001$ ) which is classed as moderate agreement [24]. A moderate degree of correlation was also found between similar dimensions for both instruments; 'school work' for the CHU9D and 'have you got on well at school' for the KIDSCREEN10 (Spearman's  $\rho = -0.416$ ,  $p < 0.001$ ); 'sad' for the CHU9D and 'have you felt lonely' for the KIDSCREEN10 (Spearman's  $\rho = -0.444$ ,  $p < 0.001$ ); 'schoolwork'

for the CHU9D and 'have you been able to pay attention' for the KIDSCREEN10 (Spearman's  $\rho = -0.433$ ,  $p < 0.001$ ). The lowest correlations were between all the CHU9D dimensions and the 'have you had enough time for yourself' dimension of the KIDSCREEN10. All Spearman's  $\rho$  were  $< -0.173$ . Other KIDSCREEN10 variables where there was consistently low agreement across all CHU9D dimensions included 'have your parents treated you fairly' (all Spearman's  $\rho < -0.271$ ) and 'have you been able to do the things you want to do in your free time' (all Spearman's  $\rho < -0.27$ ).

## Discussion

The CHU9D has demonstrated good practicality and feasibility as an online HRQoL utility measure with an excellent completion rate and low time to complete. The mean time to complete compares favourably with other measures [25] however it should be noted that the mean time includes the completion time for both the CHU9D and the KIDSCREEN 10 and also the choice experiment reported in a previous paper [18] hence this time is an overestimate of the time required to complete the CHU9D.

The online survey was designed such that respondents could opt out of the survey at any stage. However, respondents were also required to answer a question before being able to move onto the next question. Therefore those respondents who only partially completed or did not complete the CHU9D did not move onto the section asking for socio economic data therefore we are unable to assess whether there are any differences in characteristics between the completers and non-completers of the survey. However of those respondents that chose to start answering the CHU9D, only 2 out of 636 failed to complete the CHU9D fully, indicating that the CHU9D is amenable to completion by this age group.

Just over half of the respondents said that they had no difficulty in completing the survey (again this included the CHU9D, the KIDSCREEN10 and the choice experiment) and around 80% said they had either no difficulty or found it slightly difficult. This also supports the feasibility and acceptability of the CHU9D demonstrating that the majority of respondents have very little difficulty in completing it.

The CHU9D demonstrated good construct validity in this population, as it was able to discriminate between groups with known health differences (those with a long standing



illness, disability or medical condition, by level of self reported health and by KIDSCREEN score). At the dimension level, the CHU9D again performed well, demonstrating differences between groups with known health differences. Interestingly, when comparing those with and without a long standing illness or disability, only the non-emotional dimensions reflected this difference, although there was still a significant difference overall in terms of mean CHU9D utility. This may be due to the nature of the long standing illness in that adolescents adapt to their condition over time which may mean less impact on the emotional dimensions however the more physical and social dimensions (such as sleep, pain, activities and daily routine) still retain these differences.

The distribution across the response options for the CHU9D indicated a fairly healthy population, with a higher prevalence at the upper levels. This was not surprising given the nature of the sample recruited. Nevertheless, the CHU9D was able to discriminate between groups with known differences in health. The CHU9D was also able to discriminate between groups with differences in FAS status. It is known that affluence is related to HRQoL [23] therefore these differences were expected to be present. Unlike previous research with the KIDSCREEN10 [26], our data did not show a significant difference by gender, however it did show a statistically significant difference by age for the CHU9D with the same pattern for the KIDSCREEN10 (although not significant). This is a similar finding to that by Erhart et al [26] who found a statistically significant decrease in KIDSCREEN10 score by increasing age.

Some evidence of convergent validity was demonstrated by the CHU9D showing good correlations with a previously validated measure of child and adolescent HRQoL and well-being (the KIDSCREEN10). Generally there was a moderate degree of correlation between dimension levels and KIDSCREEN10 scores and all were significant. In general, the CHU9D shows a higher degree of convergent validity with the KIDSCREEN10 than the EQ5DY does [11]. Whilst the degree of convergence between the CHU9D and KIDSCREEN10 is moderate, there are some exceptions within the dimensions. This is due to the difference in scope and coverage of the two measures. The CHU9D is focused on HRQoL whereas the KIDSCREEN10 contains some dimensions which are broader in scope, such as “have you had enough time for yourself” and “have your parents treated you fairly” which would not be typically included in a measure of HRQoL designed for application in economic evaluation. The KIDSCREEN 10 measure by definition is broader encompassing well-being as well as HRQoL, hence why there are these differences.

Whilst we have undertaken testing of the practicality and validity of the CHU9D in this study, the data collected did not allow for the testing of reliability, another key psychometric criteria. This is something which should be tested in future research. It would also be desirable to have more objective measures of the health of the individuals rather than relying on a self reported observation from the individual in order to further test validity. Finally, further testing on longitudinal clinical samples would also be desirable as this would allow for testing responsiveness/sensitivity to change and also populations with more diverse clinical conditions.

## **Conclusions**

The findings from this study provide further support for the practicality and validity of the CHU9D for application in economic evaluation of health care interventions with adolescents aged 11-17 years. When compared to the KIDSCREEN10 a well validated measure of HRQoL and well-being for children and adolescents, the CHU9D performs well. Whilst there is a moderate degree of correlation between the 2 measures there are some key differences in the coverage of the dimensions with the KIDSCREEN10 being broader in scope and the CHU9D focusing on a narrower definition of HRQoL. Further research to test the validity and reliability of the CHU9D in more diverse clinical samples of adolescents is desirable.

## References

1. John J, Wenig CM, Wolfensetter SB. Recent economic findings on childhood obesity: cost-of-illness and cost-effectiveness of interventions. *Curr Opin Clin Nutr Metab Care* 2010; 13(3):305-13.
2. Ungar W. Challenges in Health State Valuation in Paediatric Economic Evaluation: Are QALYs Contraindicated? *Pharmacoeconomics* 2011; 29 (8): 641-652.
3. Brazier JE, Ratcliffe J, Salomon J et al. *Measuring and Valuing Health Benefits for Economic Evaluation*. Oxford University Press, 2007.
4. Fitzpatrick R, Davey C, Buxton M, Jones DR. Evaluating patient based outcome measures for use in clinical trials. *Health Technology Assessment* 1998; 2 (14): 1-74.
5. Matza LS, Swensen AR, Flood EM, Secnik K & Leidy NK. Assessment of Health Related Quality of Life in Children: A Review of Conceptual, Methodological, and Regulatory Issues. *Value in Health* 2004; 7(1) 79-92.
6. Dolan P. Modelling valuations for EuroQol Health States. *Medical Care* 1997; 35(11), 1095-1108.
7. Brazier JE, Roberts J & Deverill M. The estimation of a preference based measure of health from the SF-36. *Journal of Health Economics* 2002; 21 (2), 271-292.
8. Stevens KJ. Working With Children to Develop Dimensions for a Preference-Based, Generic, Pediatric Health-Related Quality-of-Life Measure. *Qualitative Health Research* 2010a; 20, 340 – 351.
9. [http://www.pedsqol.org/about\\_pedsqol.html](http://www.pedsqol.org/about_pedsqol.html) Accessed 19/08/2011
10. Torrance G, Feeny D, Furling W et al. Multiattribute utility function for a comprehensive health status classification system. *Health Utilities Index Mark 2*. *Medical Care* 1996; 34, 702-22.
11. Ravens-Sieberer U, Wille N, Badia X, Bonsel G, Burström K, Cavrini G, Devlin N et al. Feasibility, reliability, and validity of the EQ-5D-Y: results from a multinational study. *Quality of life research* 2010a; 19(6), 887-897.
12. Moodie M, Richardson J, Rankin B, Sinha K, Lezzi A. Predicting time trade-off health state valuations of adolescents in four pacific countries using the AQoL-6D instrument. *Value in Health* 2010; 13: 1014-1027.
13. Stevens KJ. Assessing the performance of a new generic measure of health related quality of life for children and refining it for use in health state valuation. *Applied Health Economics and Health Policy* 2011; 9(3); 157-169.
14. Ratcliffe J, Stevens K, Flynn T Brazier J, Sawyer M. An assessment of the construct validity of the CHU9D in the Australian adolescent general population. Accepted in *Quality of Life Research*. 2011a.



15. Ravens-Sieberer U, & the European KIDSCREEN Group. The KIDSCREEN Questionnaires - Quality of life questionnaires for children and adolescents – Handbook. Lengerich: Pabst Science Publisher. 2006.
16. Ravens-Sieberer U, Erhart M, Rajmil, L, Herdman M, Auquier P, Bruil J et al. Reliability, construct and criterion validity of the KIDSCREEN-10 score: a short measure for children and adolescents' well-being and health-related quality of life. *Quality of Life Research* 2010b; 19(10), 1487-500.
17. <http://www.kidscreen.de/> Accessed 19/08/2011
18. Ratcliffe J, Flynn T, Terlich F, Stevens K, Brazier J, Sawyer M. A pilot study to apply best worst scaling discrete choice experiment methods to obtain adolescent specific values for the Child Health Utility 9D. Flinders Centre for Clinical Change Working paper 2011/1, Flinders University, South Australia. 2011b.
19. Stevens KJ. Valuation of the Child Health Utility Index 9D (CHU9D). *Health Economics and Decision Science Discussion Paper* 10/07. 2010b.
20. Ravens-Sieber U and the European KIDSCREEN Group (2006). The KIDSCREEN Questionnaires-Quality of life questionnaires for children and adolescents – handbook. Lengerich: Pabst Science publisher.
21. Idler, EL & Benyamini. Self rated health and mortality: A review of twenty-seven community studies. *Journal of Health and Social Behaviour* 1997; 38, 21-37.
22. Currie C, Molcho M, Boyce W et al. Researching health inequalities in adolescents: the development of the Health Behaviour in School-Aged Children (HBSC) family affluence scale. *Social Science and Medicine* 2008; 66(6): 1429-1436.
23. Vvan Lenthe FJ, de Bourdeaudhuij I, Klepp KI et al. Preventing socioeconomic inequalities in health behaviour in adolescents in Europe: background, design and methods of project TEENAGE. *BMC Public Health* 2009; 8 (9) :125.
24. Landis JR & Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977; 33, 159-174.
25. Landgraf JM. Practical considerations in the measurement of HRQoL in child/adolescent clinical trials. In Fayers P & Hays R. (Eds.), *Assessing Quality of Life in Clinical Trials*. 2nd edition. Oxford University Press, 2005.
26. Erhart M, Ottova V, Gaspar T, Jerick H et al. Measuring mental health and well being of school children in 15 European countries using the KIDSCREEN-10 index. *International Journal of Public Health* 2009; 54 S160-S166.

**Table 1: Characteristics of the sample**

Mean age (standard deviation) (n=590)	14.5 years (2.0)
Gender (n=592)	
Male %	54.7%
Female %	45.3%
Self rated health (n=625)	
excellent health %	24.6%
very good health %	44.6%
good health %	22.4%
fair health %	6.9%
poor health %	1.4%
Long standing illness or disability (n=592)	
Yes	11.3%
No	88.7%
Family Affluence Level (n=588)	
Low (FAS $\leq$ 3) %	9.4%
Medium (FAS = 4 or 5) %	37.9%
High (FAS $\geq$ 6) %	52.7%

**Table 2: Level of difficulty in completing the survey**

Level of difficulty	Frequency (%)
Not difficult	306 (51.9)
Slightly difficult	173 (29.3)
Moderately difficult	86 (14.6)
Very difficult	25 (4.2)

**Table 3: CHU9D and KIDSCREEN10 scores by FAS group**

FAS group	Mean CHU9D utility	Mean KIDSCREEN10 score
Low (FAS≤3)	0.81	40.62
Medium (FAS=4 or 5)	0.84	42.86
High (FAS≥6)	0.87	44.95

**Table 4: CHU9D and KIDSCREEN10 scores by self reported general health**

General health	Mean CHU9D utility	Mean KIDSCREEN10 score
Excellent	0.93	50.12
Very good	0.86	43.85
Good	0.80	39.33
Fair	0.73	36.52
Poor	0.67	32.30

**Table 5: CHU9D and KIDSCREEN10 scores by whether long standing health problem or not**

Long standing disability, illness, or medical condition	Mean CHU9D utility	Mean KIDSCREEN10 score
Yes	0.80	40.81
No	0.86	44.11

**Table 6: CHU9D utility by KIDSCREEN10 score.**

Group	n	Mean Utility	Median Utility
KIDSCREEN10 score <median	340	0.904	0.919
KIDSCREEN10 score ≥ median	281	0.789	0.796

This difference was significant ( $p < 0.005$ )

**Table 7: Distribution across levels of the CHU9D dimensions by long standing disability, illness or medical condition or not.**

Dimension	Level	Long standing disability, illness or medical condition (valid % reported)		Significant Difference (at p<0.05)
		Yes (n=67)	No (n=525)	
Worried	No	61.2	61.3	No
	A little bit	22.4	21.3	
	A bit	10.4	12.6	
	Quite	6.0	4.0	
	Very	0.0	0.8	
Sad	No	67.2	73.1	No
	A little bit	13.4	15.8	
	A bit	14.9	8.0	
	Quite	3.0	2.9	
	Very	1.5	0.2	
Pain	No	59.7	70.7	Yes
	A little bit	23.9	19.0	
	A bit	7.5	7.2	
	Quite a lot	7.5	2.3	
	A lot	1.5	0.8	
Tired	No	16.4	29.3	Yes
	A little bit	40.3	41.7	
	A bit	26.9	16.0	
	Quite	11.9	10.1	
	Very	4.5	2.9	
Annoyed	No	52.2	58.7	No
	A little bit	23.9	25.1	
	A bit	11.9	11.2	
	Quite	6.0	3.4	
	Very	6.0	1.5	



School Work/Homework	No problems	46.3	53.5	Yes
	A few problems	19.4	28.2	
	Some problems	16.4	12.0	
	Many problems	13.4	3.0	
	Can't do	4.5	3.2	
Sleep	No problems	43.3	60.0	Yes
	A few problems	25.4	26.5	
	Some problems	19.4	9.1	
	Many problems	10.4	3.6	
	Couldn't sleep	1.5	0.8	
Daily Routine	No problems	59.7	71.8	Yes
	A few problems	23.9	21.9	
	Some problems	6.0	5.0	
	Many problems	10.4	0.8	
	Can't do	0.0	0.6	
Activities	Any	37.3	59.0	Yes
	Most	25.4	23.2	
	Some	10.4	10.3	
	A few	16.4	5.0	
	No	10.4	2.5	

**Table 8: Mean KIDSCREEN10 scores by each level of each CHU9D dimension.**

Dimension	Level	n	Mean KIDSCREEN10 score
Worried	I don't feel worried today	38 9	45.99
	I feel a little bit worried today	13 2	41.21
	I feel a bit worried today	75	39.12
	I feel quite worried today	24	36.53
	I feel very worried today	5	35.85
Sad	I don't feel sad today	45 5	45.60
	I feel a little bit sad today	97	40.06
	I feel a bit sad today	53	37.28
	I feel quite sad today	17	35.81
	I feel very sad today	3	32.23
Pain	I don't have any pain today	44 0	45.02
	I have a little bit of pain today	11 8	41.61
	I have a bit of pain today	45	39.35
	I have quite a lot of pain today	17	37.63
	I have a lot of pain today	5	38.37
Tired	I don't feel tired today	17 3	47.96
	I feel a little bit tired today	26 3	43.05
	I feel a bit tired today	10 6	41.62
	I feel quite tired today	63	39.53
	I feel very tired today	19	40.00
Annoyed	I don't feel annoyed today	36 6	46.13



	I feel a little bit annoyed today	15 2	41.74
	I feel a bit annoyed today	70	38.48
	I feel quite annoyed today	22	37.35
	I feel very annoyed today	14	38.64
School Work /Homework	I have no problems with my schoolwork/homework today	32 6	46.19
	I have a few problems with my schoolwork/homework today	17 0	42.60
	I have some problems with my schoolwork/homework today	78	39.26
	I have many problems with my schoolwork/homework today	28	38.78
	I can't do my schoolwork/homework today	22	37.82
Sleep	Last night, I had no problems sleeping	36 3	46.12
	Last night, I had a few problems sleeping	16 2	41.52
	Last night, I had some problems sleeping	64	38.88
	Last night, I had many problems sleeping	27	36.99
	Last night, I couldn't sleep at all	5	42.34
Daily Routine	I have no problems with my daily routine today	43 9	45.64
	I have a few problems with my daily routine today	13 6	40.11
	I have some problems with my daily routine today	32	37.00
	I have many problems with my daily routine today	11	35.25
	I can't do my daily routine today	3	34.67
Able to join in activities	I can join in with any activities today	35 4	46.40
	I can join in with most activities today	14 6	42.25
	I can join in with some activities today	62	38.56
	I can join in with a few activities today	38	38.01
	I can join in with no activities today	21	35.09

**Table 9: Correlation between CHU9D dimensions and KIDSCREEN10 score**

CHU9D Dimension	Correlation with KIDSCREEN 10 score (Spearman's Rho)
Worried	-.450
Sad	-.478
Pain	-.30
Tired	-.346
Annoyed	-.434
Schoolwork/Homework	-.367
Sleep	-.420
Daily routine	-.441
Able to join in activities	-.478

**Figure 1:**

**Scatter plot comparison of CHU9D utility and KIDSCREEN10 scores**

