

Title: Estimating age and sex specific utility values from the CHU9D associated with child and adolescent BMIz

Running Heading: Utility values by BMIz in children

Authors in order:

1) Anagha Killedar

The University of Sydney, Faculty of Medicine and Health, School of Public Health, NSW 2006, Australia

ORCID: 0000-0003-2363-8827

2) Thomas Lung

The University of Sydney, Faculty of Medicine and Health, School of Public Health, NSW 2006, Australia

ORCID: 0000-0001-9978-6311

The George Institute for Global Health, University of New South Wales, NSW 2042, Australia

3) Stavros Petrou

Nuffield Department of Primary Care Health Sciences, University of Oxford, Oxford OX2 6GG, United

Kingdom

ORCID: 0000-0003-3121-6050

4) Armando Teixeira-Pinto

The University of Sydney, Faculty of Medicine and Health, School of Public Health, NSW 2006, Australia

ORCID: 0000-0002-9389-2516

5) Alison Hayes

The University of Sydney, Faculty of Medicine and Health, School of Public Health, NSW 2006, Australia

ORCID: 0000-0002-8679-6040

Corresponding Author: Anagha Killedar

Mailing Address: Edward Ford Building A27, University of Sydney, NSW 2006, Australia

Email Address: anagha.killedar@sydney.edu.au

Telephone: +61 2 9351 3307

## Abstract

**Objective:** To identify age- and sex-specific utilities for children and adolescents by BMI z-score.

**Methods:** We used data from 6822 participants and 12094 observations from two cohorts and two waves of interviews from the Longitudinal Study of Australian Children (LSAC). We fit linear models using generalised estimating equations with a compound-symmetry working correlation matrix to investigate associations between Child Health Utility 9D (CHU9D) and BMI z-score in girls and boys aged 10 to 17 years. We initially fit models in each sex fully adjusted for known predictors of HRQoL including socioeconomic position, long-term medical condition and maternal smoking status and also included an interaction between age and BMI z-score to examine age-specific effects. Finally, we derived a minimal model for each sex by eliminating interaction terms with  $P > 0.01$  and predictors with  $P > 0.05$ .

**Results:** Our adjusted results show different utility patterns in girls and boys. In girls, utility decrements for each unit increase in BMI z-score changed with age ( $P < 0.01$  for interaction between age and BMI z-score). At age 10, the mean utility decrement for every unit increase in BMI z-score was 0.002 (95% CI 0.011 decrement to 0.006 increment), but by age 17, this utility decrement was 0.023 (95% CI 0.013 to 0.032 decrement). In boys, small non-significant decrements were found in utility for every unit increase in BMI z-score with no observable change with age.

**Conclusion:** Our analyses demonstrate that age and sex should be considered when attributing utility values and decrements to BMI z-scores.

## Key Points for Decision Makers

- Limited studies demonstrate small decrements in preference-based quality of life in children with overweight and obesity compared to those with healthy weight, but these studies lack the resolution to assess how this is affected by age and sex.
- Our analyses show that the size of these decrements are age and sex-specific in children aged 10 to 17. At age 10, the decrements are similar between boys and girls, but by age 17, the decrements in girls are higher than in boys.

- Our results imply that economic evaluations of obesity interventions may misrepresent cost-effectiveness if age and sex are not taken into account when estimating quality of life outcomes. Our models can be used to estimate the appropriate values and facilitate rigorous evaluations.

## **Estimating age and sex specific utility values from the CHU9D associated with child and adolescent BMIz**

### **1. Introduction**

Cost-utility analyses are an established part of decision-making in many settings [1-3]. They are particularly pertinent in the evaluation of interventions addressing overweight and obesity in childhood and adolescence where health-related quality of life (HRQoL) is a more relevant outcome than mortality and clinical morbidity [4-6]. Trials of such interventions, however, rarely measure HRQoL for calculation of quality-adjusted life years (QALYs) [7]. Moreover, decision-makers are often interested in the accrued longer term effects of childhood interventions on QALYs which cannot be identified in the short time-frames of most clinical trials. A common approach to both predicaments is to assign utilities obtained from the literature to defined health states, such as change in BMI z-score or weight status, which have either been measured within a trial or modelled through to a policy-relevant timepoint.

The choice of which utilities to apply in this context, however, can greatly affect the outcome of the economic evaluation. Hence, it is imperative to ensure that high quality data sources and rigorous methodology are used to obtain utilities as highlighted in the recent guidelines for reporting of utilities in systematic reviews [8]. Among other recommendations, the authors emphasised the importance of using age-specific utilities, echoing concerns raised by others about how utilities are applied in paediatric populations. The use of adult-specific utilities [9, 10], parent-proxy measurements [11] and adult valuations of multi-attribute utility instruments aimed at children [11, 12] have been criticised for being inappropriate for economic evaluations of paediatric interventions.

In a recent systematic review [13], significant heterogeneity was found in existing published studies investigating weight related utilities in childhood and adolescence, with most studies showing a small association, but some showing no evidence of an association between weight status and utilities. Due to the small number of studies identified, mean utilities could only be synthesised for broad weight status categories and age groups (above or below age 10). The evidence in adolescents and older adolescents was particularly limited.

The objective of the present study was to identify age and sex specific utilities associated with BMI z-scores in children and adolescents. We used data from two nationally representative cohorts in a large, ongoing longitudinal study of Australian children in which preference-based HRQoL outcomes, using the Child Health Utility 9D (CHU9D), and anthropometric data were directly measured in the two most recent waves of the study.

We aimed to identify utilities that can be used in both trial-based and modelled economic evaluations of interventions in children and adolescents.

## 2. Methods

### 2.1 Study Population

Two nationally representative cohorts from the Longitudinal Study of Australian Children (LSAC), the Birth cohort and the Kindergarten cohort, were used in this analysis. These cohorts were recruited to the LSAC using a two-stage clustered sampling design [14]. The first wave of data collection was conducted in 2004 with interviews conducted biennially since then. In our analysis, we used data from the two most recent waves, six and seven, the only waves to measure HRQoL using the CHU9D. In wave six, interviews were conducted between March 2014 and February 2015 when children from the Birth (B) cohort were aged between 10.1 and 11.8 years and children from the Kindergarten (K) cohort were aged between 14.0 and 15.8 years. In wave seven, interviews were conducted between April 2016 and July 2017 when children from the B cohort were aged between 12.1 and 14.3 years and children from the K cohort were aged between 15.9 and 18.3 years. Ethics approval for our study using data from LSAC was obtained from the University of Sydney Human Ethics Committee (Project Number 2018/726).

### 2.2 Measures

#### 2.2.1 CHU9D

The outcome in this analysis was utility derived from the CHU9D, a generic, preference-based measure of HRQoL targeted at adolescent populations. This instrument elicits responses on nine dimensions of HRQoL: feeling worried, sadness, pain, tiredness, annoyance, difficulty with schoolwork, sleep problems, difficulty with daily routine and difficulty in joining in activities. Participants were asked, on a five-point scale, how severe their function was on each of these dimensions on the day of the interview. All responses were self-reported by the child with both child and parent permission. The valuation of the responses was based on a best-worst scaling study conducted in an Australian community adolescent population [15].

### 2.2.2 BMI z-score

The study factor of this analysis was BMI z-score using World Health Organisation (WHO) growth standards. Anthropometric measurements of height and weight were taken at each interview [16]. Two measurements of height in centimetres were taken using a laser stadiometer at each interview, and if they differed by 0.5cm or more a third was taken. The mean of the two closest measurements was recorded. Weight in kilograms was measured using the Tanita body fat scales. To maintain confidentiality of the data, extreme measures of height and weight were top-coded. Age in days was calculated as the difference between the date of interview and the date of birth for each child. Using the age, sex, height and weight of each child, a BMI z-score was calculated using a SAS macro published by WHO [17]. We omitted BMI z-scores less than -5 and more than 5 as these are considered biologically implausible [18].

### 2.2.3 Predictors of HRQoL

Important known predictors of HRQoL were considered in this analysis as controlling covariates or stratifying variables. These included age, sex (male/female), long-term socioeconomic position (high/low) [19], indigenous status (indigenous/not indigenous) [20], mother's smoking status (smoker, non-smoker) [21], and child's long term medical condition or disability (has condition, does not have condition) [22]. Age in years at time of interview was included as a continuous variable accurate to the day. A time-varying measure of socioeconomic position (SEP) incorporating family income, parental occupation and parental education was recorded for each participant in LSAC. We calculated a long-term measure of SEP by taking the average of the recorded SEP z-score from waves two to seven and categorising this into low and high SEP corresponding to below and above the median z-score of 0 respectively. As mother's smoking status was not recorded in wave seven for the K cohort, we used their wave six measurements as this was likely the strongest predictor for current smoking status. Child's long-term medical condition or disability was defined as whether the child had at least one medical condition or disability lasting six months or longer, and was collected in both waves used in this analysis.

## 2.3 Statistical Analysis

Separate analyses were conducted for boys and girls due to the clear sex-specific associations between BMI z-score and CHU9D score (Figure 1), and age and CHU9D score (Appendix Figure A1) identified in descriptive analyses. These descriptive analyses also supported the use of linear models. Therefore, for both analyses, we examined the association between BMI z-score and CHU9D utility score by fitting linear models using generalised estimating equations (GEE) with a compound-symmetry working correlation matrix. Within each sex, we combined the data from the B and K cohorts as the similarity in the recruitment and sampling strategies, and the distribution of values for HRQoL predictors (Table 1), minimised the potential for cohort effects. We also combined data from each wave, which facilitated the investigation of age-specific effects of BMI z-score on CHU9D utility score from ages 10 to 17. GEE was used to account for the intra-individual correlation arising from the repeated measurements in each participant within B and K cohorts.

For both girls and boys, we initially fitted unadjusted linear models assessing the association between BMI z-score and CHU9D score. We then fitted models that were fully adjusted for known predictors of HRQoL (age, SEP, long-term medical condition, indigenous status, mother's smoking status, and child's long term medical condition or disability) and included an interaction between age and BMI z-score to examine age-specific effects. Finally, we derived a minimal model using a backwards stepwise procedure starting with eliminating interaction terms with  $P > 0.01$ , then predictors, other than BMI z-score, with  $P > 0.05$ . All analyses were conducted in (StataCorp. 2017. Stata Statistical Software: Release 15. College Station, TX: StataCorp LLC).

#### 2.4 Estimated utilities by BMI z-score weight status

The final 'minimal' models for girls and boys were used to predict mean utilities specific to sex, age and BMI z-score adjusted for all other covariates in the model using the margins command in STATA. Mean values within the study observations were specified for all covariates other than sex, age and BMI z-score. For the dichotomous variables, indigenous status, long-term medical condition and mother's smoking status, the proportion of observations with the outcome was specified. Utilities for boys and girls, at BMI z-scores -2 to 3 and ages 10 to 17 were calculated. We tabulated utilities for BMI z-scores of -2, 1 and 2, which correspond to weight status boundaries as defined by the WHO [17]. Children with a BMI z-score under -2 are classified 'thin', between -2 and 1 are classified as 'normal' (i.e. healthy weight), between 1 and 2 are classified as 'overweight' and above 2 are classified

as ‘obese’. We also tabulated the utility for a BMI z-score of 3 as the majority of observations within the obese category lay between a BMI z-score of 2 and 3 (86%).

## 2.5 Descriptive analysis by CHU9D dimension

To gain further insight into the results of our primary analysis, we analysed the association between sex, weight status and each individual dimension of CHU9D. We categorised BMI z-score into two groups: healthy weight and overweight groups corresponding to a BMI z-score less than or equal to 1, and above 1 respectively. For each sex, we then calculated the difference in mean scores (overweight group minus healthy weight group) for each dimension.

## 2.6 Sensitivity Analyses

The impact of a number of analysis decisions on the results was assessed in sensitivity analyses. In separate analyses, we refitted the minimal models derived for boys and girls and 1) used BMI z-scores based on Centers for Disease Control and Prevention (CDC) growth standards [23]; 2) excluded underweight children from the dataset (that is, children with BMI z-score less than -2 based on WHO growth standards); 3) used the weight status groups healthy and underweight, overweight and obese instead of continuous BMI z-score; and 4) recalculated CHU9D utility using a tariff value set based on the preferences of adults from the United Kingdom [24]. For each of these sensitivity analyses, sex, age and BMI z-score or weight status specific utilities, adjusted for all other covariates, were predicted using the margins command in STATA.

# 3. Results

## 3.1 Characteristics of Study Populations

There were 6165 observations for girls and 6525 observations for boys that had complete information for the CHU9D and BMI z-score. The characteristics of these populations are displayed in Table 1. In both boys and girls, the distribution of known predictors of CHU9D are similar in the B and K cohort at wave six. Despite losses between waves six and seven of between 9 and 12%, the distribution of values for predictors remained broadly stable in each cohort and sex. Within observations with complete data for the CHU9D utility score and BMI z-score,



there were minimal missing data (less than 6%) for other characteristics. The only exception was mother's smoking status, which was not collected in wave seven for the K cohort.

### 3.2 Unadjusted Analyses

In both girls and boys, the unadjusted association between BMI z-score and CHU9D utility score was small but significant (Table 2). In girls, utility decreased by 0.014 (95% CI -0.019 to -0.009), and for boys, utility decreased by 0.005 (95% CI -0.009 to -0.0004) for every one unit increase in BMI z-score.

### 3.3 Adjusted Analyses

The coefficients and statistics for the fully adjusted and minimal models in girls and boys are presented in table 3, respectively. In girls, the minimally adjusted model included BMI z-score, age, long-term medical condition, mother's smoking status and the interaction between BMI z-score and age. The association between BMI z-score and CHU9D was affected by age ( $P < 0.01$  for BMI z-score by age interaction). For every year increase in age, the coefficient for BMI z-score decreased by 0.003 (95% CI 0.001 to 0.005). In contrast, in boys, there was no evidence that age affected the association between BMI z-score and CHU9D ( $P = 0.48$  for BMI z-score by age interaction in the full model). Therefore, the interaction term was excluded from the minimally adjusted model. Variables that were included were BMI z-score, age, indigenous status, long-term medical condition, and mother's smoking status. There was only weak evidence of a BMI z-score – CHU9D association in boys at any age between 10-17 years ( $P = 0.07$  for BMI z-score coefficient).

### 3.4 Trends in adjusted mean utilities and decrements

Girls and boys displayed very different trends in mean utilities adjusted for other covariates as per the minimal models (Figure 2). While adjusted utility decreased with increasing BMI z-score in both sexes, in girls, these decrements became more extreme with increasing age. For example, at age 10, the utility decrement for every unit increase in BMI z-score was 0.002 (95% CI -0.011 to 0.006), but by age 17, the decrement increased to 0.023 (95% CI -0.013 to -0.032). In boys, the decrement for a unit increase in BMI z-score was 0.004 (95% CI -0.0003 to 0.008) at all ages, but adjusted mean utility increased with age. Overall, boys had higher utilities than girls.

### 3.5 Utility values for economic evaluation

Table 4 presents the adjusted utility values and their 95% confidence intervals by sex, BMI z-score and age and at mean values for other covariates in the minimal models. These may be used in economic evaluations of interventions addressing childhood and adolescent obesity.

### 3.6 Descriptive analysis by CHU9D dimension

The difference in dimension scores between overweight and healthy weight groups, by sex, is presented in Figure A2. A difference of above 0 indicates that the mean score in the overweight group is higher (i.e. worse) than in the healthy weight group. This difference was greater in girls than in boys in six out of the nine dimensions.

### 3.6 Sensitivity Analyses

In all sensitivity analyses, the trends identified and conclusions made were the same as those from the base case analysis (Appendix C). All analyses indicated lower utility scores with higher BMI z-scores or weight status, smaller decrements in boys than in girls, increasing strength of the association of utility and weight with age in girls, and small and consistent utility decrements at all ages between 10 to 17 among boys. However, the magnitude of the coefficients and standard errors changed in each of the analyses, and, in some cases, this changed the P-values.

## 4. Discussion

In this study we used longitudinal data from 6822 participants and 12094 observations to estimate utility values for children and adolescents, based on BMI z-scores. To the best of our knowledge, it is the first study to identify age, sex and BMI z-score specific utilities in adolescents. Importantly, the estimates are informed by high quality data. We used a large, national dataset with direct measurements of height and weight, rather than self-reported measurements. Furthermore, the instrument used to derive utilities, the CHU9D, was specifically designed for children [25] and used responses from the child participants, not parent proxies. The valuation of the instrument was derived from an Australian adolescent population encompassing a similar age range as our study population (11-17 years) [15], facilitating an age-appropriate measure of utility. The analysis accounted for repeated measurements over two years, and formally assessed age effects on the relationship between BMI z-score and utility while adjusting for known predictors of HRQoL.

We have derived two prediction models, one for girls and one for boys, which can be used to calculate the appropriate utilities to incorporate into health economic models for the calculation of QALYs. In both estimation models, utility decreased with increasing BMI z-score which is consistent with most other published studies in a similar age group [26, 13, 27]. One exception is a study in which no difference between utilities of high school students, aged 13-17 across weight statuses was found, but this study likely suffered from insufficient power [28]. In studies that did demonstrate decreasing utility with increasing BMI z-score or weight status, including systematic reviews, the mean utility and decrements reported [26, 13, 27] are slightly larger than in our study. Reported utility decrements tended to be most similar with the older girls in our sample in whom the weight-utility association was strongest.

Another finding of our study was the very different relationship between BMI z-scores and utility values for girls and boys. In girls, both overall utilities and the decrements in utility with higher BMI z-scores were age-specific. In boys, however, overall utilities were age-specific but the decrements with higher BMI z-score were consistently small at all ages between 10 and 17 years. The decrement for a one unit increase in BMI z-score (0.004) in boys was similar to that in the youngest girls (0.002 in 10 year old girls) but much smaller than that in older girls (0.023 in 17 year old girls). This somewhat surprising finding is consistent with the breadth of research on weight stigma in children and adolescents. Qualitative and quantitative work has repeatedly shown that adolescent girls are particularly susceptible to weight-related bullying and poor self-esteem with considerable negative impacts on social and psychological health outcomes [29-34]. These psychosocial effects appear to be the basis for our results, as the most prominent differences between boys and girls in the individual dimensions of the CHU9D instrument (Appendix Figure A2) were in 'feeling worried', 'feeling sad' and 'having difficulty joining in activities'. Moreover, the adolescent valuation of the CHU9D that was used in our study assigns a high weighting to the mental health related dimensions [15, 35].

While most other studies investigating the association between weight and utility in children control for sex, only two to our knowledge have assessed whether the association is different between boys and girls [27, 36]. Our results are consistent with one such study that included a relatively large sample of school students aged between 11 and 18 from the state of Victoria, Australia [27]. The authors found that girls who had overweight or obesity had larger decrements in utility than boys when compared to their healthy weight counterparts. In a smaller study with a similar age group, an interaction between gender and weight status was found, but no overall

association between weight status and HRQoL was identified [36]. Our finding that the association between increasing BMI z-score and utility became stronger with age in girls, is in line with our earlier study assessing the association between weight status and a non-preference based measure of HRQoL, PedsQL™ [37]. We only identified one study that examined whether age modified the association between weight status and utility [36] and, interestingly, the authors found no evidence of an interaction effect. This could be explained by the study's limited power and an analysis that did not stratify by sex. Whilst these findings can be explained from a psychosocial perspective, they do have unexpected implications when considering their use in economic evaluations: an intervention that elicits the same change in BMI z-score and costs between the sexes could have better cost-utility in girls than in boys.

The results of our analysis have broader implications to disease areas beyond obesity. Descriptive (Figure 1) and adjusted (Figure 2) analyses both showed an overall pattern of girls' utility decreasing and boys' utility increasing with age, irrespective of BMI z-score. These underlying opposing associations will likely mean that age and sex have modifying effects in conditions other than obesity, in which case these factors should be accounted for in the cost-utility analyses of other types of interventions. Further research is needed to determine if these modifying effects arise in other disease areas.

A caveat to our findings is that the study combined observations from two cohorts to make assertions about age effects on the BMI z-score-utility association. However, the two cohorts have been sampled in an analogous manner and, as shown in Table 1, have similar distributions of characteristics at wave six which did not change meaningfully after some losses by wave seven. Also, we were unable to account for LSAC survey weights while simultaneously controlling for repeated measurements using GEE, but our sample was large and diverse enough that the associations are likely still generalisable to the child and adolescent population in Australia. It's also possible that the preference structures underpinning the CHU9D Australian tariff set might not be generalisable to childhood populations in other countries. However, the results of our sensitivity analysis using the UK adult valuation of the CHU9D instrument suggests that the trends found in the base case analysis are likely applicable to the United Kingdom (Appendix C). Further research using external datasets from other high income countries should be conducted to confirm the generalisability of our results. Another caveat of our models is that they should not be used to estimate utilities outside the age range of 10 to 17 years and for those in the underweight category as our study population did not include sufficient observations in these groups. Further work should be directed towards these

population groups to allow the estimation of utilities in all children and adolescents. It should also be noted, that despite our controlled analysis, the observational nature of the dataset used to derive the models does not guarantee causality: the identified association between BMI z-score and utility may be a result of some third unobserved factor.

## 5. Conclusion

The results and final models developed in this study will be a valuable resource for health economists tasked with conducting trials of relevant interventions, but lacking direct measurements of preference-based HRQoL outcomes, or when cost-utility is modelled from childhood to late adolescence. Our results have the flexibility to be used when effects on either BMI z-score, based on WHO growth standards, or weight status have been measured or modelled. The models can be used to calculate an estimated utility based on the individual characteristics of the population of interest including age, sex, BMI z-score and, optionally, indigenous status, the presence of a long-term medical condition, and mother's smoking status. Where these latter characteristics are unknown, average values from our study's population can be used. We suggest that health economists who need to estimate utility for their study population indirectly should consider the age and sex of the participants when calculating cost-utility for interventions that prevent or treat obesity, and perhaps other conditions, in children or adolescents.

## Data Availability Statement

The data that support the findings of this study are available from the Department of Social Services (DSS), Australian Government but restrictions apply to the availability of these data, which were used under license for the current study, and so are not publicly available. Data are however available upon application to the DSS. The code used to analyse the data is available on request to the corresponding author.

## Acknowledgements

We thank the Department of Social Services, Australian Government for providing access to data collected from the Longitudinal Study of Australian Children. We also thank Emma Frew and the International Health Economics Association's Economics of Obesity Special Interest Group for their comments and feedback on an earlier draft of this paper.

## **Compliance with Ethical Standards**

*Conflicts of Interest* None of the authors have any conflicts of interest to declare.

*Funding Sources* Anagha Killedar is supported by the Kassulke Scholarship and the National Health and Medical Research Council (NHMRC) Scholarship (APP1169039) for PhD study. Thomas Lung is supported by an NHMRC Early Career Fellowship (APP1141392) and a National Heart Foundation Postdoctoral Fellowship (award ID 101956)

*Author Contributions* Anagha Killedar, Alison Hayes, Thomas Lung and Armando Teixeira-Pinto contributed to the study design. Anagha Killedar analysed the data and wrote the initial draft of the manuscript. Anagha Killedar, Alison Hayes, Thomas Lung and Stavros Petrou interpreted the analyses, and all authors commented on and made revisions to manuscript drafts. Anagha Killedar will act as the overall guarantor.

*Ethical Approval* Ethics approval for our study was obtained from the University of Sydney Human Ethics Committee (Project Number 2018/726).

*Informed Consent* This analysis used secondary data from the Longitudinal Study of Australian Children.

## **References**

1. National Institute for Health and Care Excellence. Guide to the methods of technology appraisal 2013. In: NICE, editor. London 2013.
2. Australian Government. Guidelines for preparing a submission to the Pharmaceutical Benefits Advisory Committee (Version 5.0). In: Department of Health, editor. 2016.
3. Centre for Epidemiology and Evidence. Commissioning Economic Evaluations: A Guide. In: NSW Ministry of Health, editor. North Sydney, NSW 2017.
4. Varni JW, Burwinkle TM, Seid M, Skarr D. The PedsQL™ 4.0 as a pediatric population health measure: feasibility, reliability, and validity. *Ambulatory pediatrics*. 2003;3(6):329-41.
5. McGlynn EA, Halfon N. Overview of issues in improving quality of care for children. *Health services research*. 1998;33(4 Pt 2):977-1000.

6. Schwimmer JB, Burwinkle TM, Varni JW. Health-related quality of life of severely obese children and adolescents. *JAMA*. 2003;289(14):1813-9. doi:10.1001/jama.289.14.1813.
7. Doring N, Mayer S, Rasmussen F, Sonntag D. Economic Evaluation of Obesity Prevention in Early Childhood: Methods, Limitations and Recommendations. *Int J Environ Res Public Health*. 2016;13(9). doi:10.3390/ijerph13090911.
8. Brazier J, Ara R, Azzabi I, Busschbach J, Chevrou-Severac H, Crawford B et al. Identification, Review, and Use of Health State Utilities in Cost-Effectiveness Models: An ISPOR Good Practices for Outcomes Research Task Force Report. *Value in health: the journal of the International Society for Pharmacoeconomics and Outcomes Research*. 2019;22(3):267-75. doi:10.1016/j.jval.2019.01.004.
9. Kwon J, Kim SW, Ungar WJ, Tsiplova K, Madan J, Petrou S. Patterns, trends and methodological associations in the measurement and valuation of childhood health utilities. *Qual Life Res*. 2019. doi:10.1007/s11136-019-02121-z.
10. Montgomery SM, Kusel J. The prevalence of child-specific utilities in NICE appraisals for paediatric indications: rise of the economic orphans? *Expert Rev Pharmacoecon Outcomes Res*. 2016;16(3):347-50. doi:10.1080/14737167.2016.1179116.
11. Wolstenholme JL, Bargo D, Wang K, Harnden A, Raisanen U, Abel L. Preference-based measures to obtain health state utility values for use in economic evaluations with child-based populations: a review and UK-based focus group assessment of patient and parent choices. *Qual Life Res*. 2018;27(7):1769-80. doi:10.1007/s11136-018-1831-6.
12. Griebisch I, Coast J, Brown J. Quality-adjusted life-years lack quality in pediatric care: a critical review of published cost-utility studies in child health. *Pediatrics*. 2005;115(5):e600-14. doi:10.1542/peds.2004-2127.
13. Brown V, Tan EJ, Hayes AJ, Petrou S, Moodie ML. Utility values for childhood obesity interventions: a systematic review and meta-analysis of the evidence for use in economic evaluation. *Obesity reviews : an official journal of the International Association for the Study of Obesity*. 2018;19(7):905-16. doi:10.1111/obr.12672.
14. Soloff C, Lawrence D, Johnstone R. LSAC technical paper no. 1: Sample design. Melbourne, Australia: Australian Institute of Family Studies. 2005.
15. Ratcliffe J, Huynh E, Chen G, Stevens K, Swait J, Brazier J et al. Valuing the Child Health Utility 9D: Using profile case best worst scaling methods to develop a new adolescent specific scoring algorithm. *Social Science & Medicine*. 2016;157:48-59. doi:https://dx.doi.org/10.1016/j.socscimed.2016.03.042.

16. Australian Institute of Family Studies. The Longitudinal Study of Australian Children: An Australian Government Initiative Data User Guide–December 2018. In: Australian Bureau of Statistics, Department of Social Services, editors.: AIFS; 2018.
17. World Health Organization. Growth reference 5-19 years. Geneva, Switzerland. 2007.  
<https://www.who.int/growthref>. Accessed 1/05/2019.
18. World Health Organization. SAS macro package. Geneva, Switzerland. 2007.  
[https://www.who.int/growthref/tools/readme\\_sas.pdf](https://www.who.int/growthref/tools/readme_sas.pdf). Accessed 1/05/2019.
19. Otto C, Haller A-C, Klasen F, Hölling H, Bullinger M, Ravens-Sieberer U et al. Risk and protective factors of health-related quality of life in children and adolescents: Results of the longitudinal BELLA study. *PloS one*. 2017;12(12):e0190363.
20. Jansen PW, Mensah FK, Clifford S, Nicholson JM, Wake M. Bidirectional associations between overweight and health-related quality of life from 4-11 years: Longitudinal Study of Australian Children. *International Journal of Obesity*. 2013;37(10):1307-13. doi:10.1038/ijo.2013.71.
21. Vella SA, Magee CA, Cliff DP. Trajectories and Predictors of Health-Related Quality of Life during Childhood. *The Journal of Pediatrics*. 2015;167(2):422-7. doi:<https://doi.org/10.1016/j.jpeds.2015.04.079>.
22. Varni JW, Limbers CA, Burwinkle TM. Impaired health-related quality of life in children and adolescents with chronic conditions: a comparative analysis of 10 disease clusters and 33 disease categories/severities utilizing the PedsQL™ 4.0 Generic Core Scales. *Health and Quality of Life Outcomes*. 2007;5(1):43. doi:10.1186/1477-7525-5-43.
23. Centers for Disease Control and Prevention. Defining Childhood Obesity. CDC, USA. 2018.  
<https://www.cdc.gov/obesity/childhood/defining.html>. Accessed 19/02/2019.
24. Stevens K. Valuation of the Child Health Utility 9D Index. *Pharmacoeconomics*. 2012;30(8):729-47.  
doi:<https://dx.doi.org/10.2165/11599120-000000000-00000>.
25. Stevens KJ. Working With Children to Develop Dimensions for a Preference-Based, Generic, Pediatric, Health-Related Quality-of-Life Measure. *Qualitative Health Research*. 2010;20(3):340-51.  
doi:10.1177/1049732309358328.
26. Kwon J, Kim SW, Ungar WJ, Tsiplova K, Madan J, Petrou S. A Systematic Review and Meta-analysis of Childhood Health Utilities. *Medical Decision Making*. 2018;38(3):277-305. doi:10.1177/0272989x17732990.



27. Keating CL, Moodie ML, Richardson J, Swinburn BA. Utility-based quality of life of overweight and obese adolescents. *Value in health: the journal of the International Society for Pharmacoeconomics and Outcomes Research*. 2011;14(5):752-8. doi:10.1016/j.jval.2011.02.1181.
28. Chen G, Ratcliffe J, Olds T, Magarey A, Jones M, Leslie E. BMI, health behaviors, and quality of life in children and adolescents: a school-based study. *Pediatrics*. 2014;133(4):e868-74. doi:<https://dx.doi.org/10.1542/peds.2013-0622>.
29. Falkner NH, Neumark-Sztainer D, Story M, Jeffery RW, Beuhring T, Resnick MD. Social, educational, and psychological correlates of weight status in adolescents. *Obesity research*. 2001;9(1):32-42.
30. Griffiths LJ, Page AS. The impact of weight-related victimization on peer relationships: the female adolescent perspective. *Obesity*. 2008;16(S2):S39-S45.
31. Taylor NL. "Guys, She's Humongous!": Gender and Weight-Based Teasing in Adolescence. *Journal of Adolescent Research*. 2011;26(2):178-99. doi:10.1177/0743558410371128.
32. Puhl RM, King KM. Weight discrimination and bullying. *Best Practice & Research Clinical Endocrinology & Metabolism*. 2013;27(2):117-27. doi:<https://doi.org/10.1016/j.beem.2012.12.002>.
33. Hayden-Wade HA, Stein RI, Ghaderi A, Saelens BE, Zabinski MF, Wilfley DE. Prevalence, characteristics, and correlates of teasing experiences among overweight children vs. non-overweight peers. *Obesity research*. 2005;13(8):1381-92.
34. Faith MS, Leone MA, Ayers TS, Heo M, Pietrobelli A. Weight criticism during physical activity, coping skills, and reported physical activity in children. *Pediatrics*. 2002;110(2):e23.
35. Ratcliffe J, Huynh E, Stevens K, Brazier J, Sawyer M, Flynn T. Nothing About Us Without Us? A Comparison of Adolescent and Adult Health-State Values for the Child Health Utility-9D Using Profile Case Best-Worst Scaling. *Health Economics*. 2016;25(4):486-96. doi:<https://dx.doi.org/10.1002/hec.3165>.
36. Bolton K, Kremer P, Rossthorn N, Moodie M, Gibbs L, Waters E et al. The effect of gender and age on the association between weight status and health-related quality of life in Australian adolescents. *BMC Public Health*. 2014;14:898. doi:10.1186/1471-2458-14-898.
37. Killedar A, Lung T, Petrou S, Teixeira-Pinto A, Tan EJ, Hayes A. Weight status and health-related quality of life during childhood and adolescence: effects of age and socioeconomic position. Unpublished.



## Figure Legends

**Fig. 1** Lowess curve of CHU9D utility and BMI z-score in a) girls and b) boys. Lowess curves of all observations with complete data on CHU9D and BMI z-score from Waves 6 and 7 of the Longitudinal Study of Australian Children, separated by sex, were fitted using Stata Version 15. CHU9D utility is displayed on a scale from 0.65 to 0.87. *CHU9D* Child Health Utility 9D

**Fig. 2** Adjusted mean CHU9D utilities and 95% confidence intervals by BMI zscore and age for a) girls and b) boys. Mean utilities by BMI z-score, age and sex, adjusted for long-term medical condition and maternal smoking status for girls, and long-term medical condition, maternal smoking status and indigenous status for boys, were predicted using the derived minimal models. Adjusted mean CHU9D utility is displayed on a scale from 0.65 to 0.87. Vertical lines are placed at cut-off points for overweight (BMIz=1) and obese (BMIz=2) weight status categories as per World Health Organization guidelines. *CHU9D* Child Health Utility 9D