PharmacoEconomics
https://doi.org/10.1007/s40273-023-01300-8

LEADING ARTICLE

Using Age-Specific Values for Pediatric HRQoL in Cost-Effectiveness Analysis: Is There a Problem to Be Solved? If So, How?

Nancy J. Devlin[1] · Tianxin Pan[1] · Mark Sculpher[2] · Mark Jit[3] · Elly Stolk[4,5] · Donna Rowen[6] · Barend van Hout[6] · Richard Norman[7]

Accepted: 21 June 2023

Abstract

Value sets for the EQ-5D-Y-3L published to date appear to have distinctive characteristics compared with value sets for corresponding adult instruments: in many cases, the value for the worst health state is higher and there are fewer values < 0. The aim of this paper is to consider how and why values for child and adult health differ; and what the implications of that are for the use of EQ-5D-Y-3L values in economic evaluations to inform healthcare resource allocation decisions. We posit four potential explanations for the differences in values: (a) The wording of severity labels may mean the worst problems on the EQ-5D-Y-3L are descriptively less severe than those on the EQ-5D-5L; (b) Adults may genuinely consider that children are less badly affected than adults by descriptively similar health issues. That is, for any given health problem, adult respondents in valuation studies consider children's overall health-related quality of life (HRQoL) on average to be higher than that for adults; (c) Values are being sought by eliciting adults' stated preferences for HRQoL in another person, rather than in themselves (regardless of whether the 'other person' concerned is a child); and (d) The need to elicit preferences for child HRQoL that are anchored at dead = 0 invokes special considerations regarding children's survival. Existing evidence does not rule out the possibility that (c) and (d) exert an upward bias in values. We consider the implications of that for the interpretation and use of values for pediatric HRQoL. Alternative methods for valuing children's HRQoL in a manner that is not 'age specific' are possible and may help to avoid issues of non-comparability. Use of these methods would place the onus on health technology assessment bodies to reflect any special considerations regarding child quality-adjusted life-year gains.

Key Points for Decision Makers

Value sets for EQ-5D-Y-3L published to do date have distinctive characteristics, which may be caused by multiple factors operating independently and in combination.

Differences in the length of the value scale for child and adult health states means they cannot necessarily be considered equivalent, limiting the ability to compare quality-adjusted life-years and incremental cost-effectiveness ratios generated from them.

We distinguish between age-specific and age-invariant values for pediatric health-related quality of life. Age-invariant values could avoid problems of non-comparability but require special societal considerations regarding children to be reflected elsewhere in the health technology assessment process.

✉ Nancy J. Devlin
nancy.devlin@unimelb.edu.au

1 Health Economics Unit, Melbourne School of Population and Global Health, University of Melbourne, 207 Bouverie St, Melbourne, VIC 3010, Australia

2 Centre for Health Economics, University of York, York, UK

3 Department of Infectious Disease Epidemiology, Faculty of Epidemiology and Population Health, London School of Hygiene and Tropical Medicine, London, UK

4 EuroQol Research Foundation, Rotterdam, Rotterdam, The Netherlands

5 Erasmus School of Health Policy and Management, Erasmus University Rotterdam, Rotterdam, The Netherlands

6 School of Health and Related Research, University of Sheffield, Sheffield, UK

7 School of Population Health, Curtin University, Perth, WA, Australia

Published online: 13 July 2023

1 Introduction

Measurement of health-related quality of life (HRQoL) is a vital input to a cost-effectiveness analysis (CEA) of healthcare. EQ-5D is a preferred instrument to measure HRQoL in this context [1]. A key feature of EQ-5D is the availability of preference weights (values) for all HRQoL states, reflecting how good or bad those states are in a manner that allows the estimation of quality-adjusted life-years (QALYs). Until 2010, the EQ-5D was only available for use in adults, but demand for its use in pediatric populations motivated the development of a version of EQ-5D, the EQ-5D-Y-3L, for use in children [2]. The EQ-5D-Y-3L covers the same dimensions of health as the EQ-5D but with the wording of dimensions and severity levels adapted to be age appropriate for children. The EQ-5D-Y-3L can be self-completed by children 8–15 years of age [3]. Properties of the values that can be attached to health states described by the EQ-5D-Y-3L are the focus of this paper.

An initial study of values for EQ-5D-Y-3L, undertaken in Germany, Spain, England, and the Netherlands, found that adults assigned somewhat different values to health states described by the (adult) EQ-5D-3L and those described by the EQ-5D-Y-3L [4]. Overall, values for EQ-5D-3L were lower than those for EQ-5D-Y-3L states. This was attributable in part to differences in the perspective adult respondents were asked to adopt, i.e., ‘own health’ in the former, versus ‘a 10-year-old child’ in the latter. The study concluded that existing value sets for the EQ-5D-3L could therefore not be directly applied to patient data collected for the EQ-5D-Y-3L, suggesting a need for value sets specific to children’s HRQoL states described by it. This in turn led to a focus on methods development aimed at developing and testing the methods to use in generating age-specific values for the EQ-5D-Y-3L.

The subsequent protocol describing those methods [5] led to the rapid development of value sets for EQ-5D-Y-3L [6, 7]. The protocol recommends use of a combination of two methods in producing EQ-5D-Y-3L value sets: a discrete choice experiment (DCE) and a composite time trade-off. The DCE task comprises a series of pairwise comparisons of EQ-5D-Y-3L states, with respondents asked to indicate which they prefer when considering their views about a hypothetical 10-year-old child. In the composite time trade-off task, respondents are asked to indicate what they prefer for a hypothetical 10-year-old child: living in an impaired health state for 10 years or in full health for a shorter duration. The amount of time the child would live in full health is varied until indifference is reached. By describing the health states in terms of EQ-5D-Y-3L profiles, and by referring to a 10-year-old child in the valuation task, the resulting value sets are specific both to its descriptive system and to the specific age of children whose health is being described.

In principle, the availability of the EQ-5D-Y-3L and age-specific value sets to accompany it means they can be used to support economic evaluations of interventions primarily aimed at children—an important step forward given the substantial evidence gaps noted in health technology assessments (HTAs) of pediatric interventions [8, 9]. Yet a number of these value set studies sound a note of caution about the extent to which the values for child HRQoL they report, and QALYs estimated from them, can be compared to values for adult HRQoL and adults’ QALYs [10, 11]. From the value sets now available for the EQ-5D-Y-3L and results from related methodological studies, a picture has begun to emerge that the empirical characteristics of EQ-5D-Y-3L values differ from those of value sets for adult EQ-5D instruments.

The aim of this paper is to consider how and why values for child and adult health differ; and what the implications of that are for the use of EQ-5D-Y-3L values in economic evaluations to inform healthcare resource allocation decisions. At the heart of this paper is the question: are the observed differences in adults’ preferences for child HRQoL and adult HRQoL a problem that needs to be solved? If so, how?

2 What Are the Characteristics of EQ-5D-Y-3L Values and Why Do They Differ from Values for EQ-5D Adult Instruments?

The principal characteristics of EQ-5D-Y-3L value sets can be summarized in terms of scale length (i.e., the difference between full health and the value for the worst health state) and order of dimension importance. These can be compared with the properties of value sets for the EQ-5D-5L [12]. While value sets exist for both the three-level and five-level versions of adult EQ-5D instruments, values for the EQ-5D-5L arguably provide a better basis for comparisons with values for the EQ-5D-Y-3L. The methods used to construct EQ-5D-Y-3L value sets resemble more closely the methods used to produce value sets for EQ-5D-5L than for EQ-5D-3L, albeit with some differences (for a summary of the differences in the EQ-VT protocol applied to EQ-5D-5L and EQ-5D-Y-3L, see Devlin et al. [5]). Further, EQ-5D-5L value sets tend to have been developed more recently than values for EQ-5D-3L, so the preferences they reflect are a more contemporaneous comparison with the preferences for EQ-5D-Y-3L, all value sets for which have been published since 2021. Values for both instruments have been obtained from representative samples of the adult general public.

Table 1 Potential causes of differences in the characteristics of values between EQ-5D-Y-3L and EQ-5D-5L

Potential factors	Interpretation?
The wording of severity labels may mean the worst problems on the EQ-5D-Y-3L are descriptively less severe than those on the EQ-5D-5L, e.g., in the functional dimensions, 'cannot do/unable to' for the EQ-5D-5L compared with 'a lot of problems' in the EQ-5D-Y-3L	Difference in values arguably legitimate
Adults may genuinely consider that children are less badly affected than adults by descriptively similar health issues. That is, for any given health problem, adult respondents in valuation studies consider children's overall HRQoL on average to be higher than that for adults	Difference in values arguably legitimate
Values are being sought by eliciting adults' stated preferences for HRQoL in another person, rather than in themselves (regardless of whether the 'other person' concerned is a child)	The difference in perspective limits the comparability of values
The need to elicit preferences for child HRQoL that are anchored at dead = 0 invokes special considerations regarding children's survival, which are not related to the health problems being valued per se, making child and adult values qualitatively different	Special considerations about children limits the comparability of values

A detailed comparison of EQ-5D-Y-3L and EQ-5D-5L value sets is reported by Roudijk et al. [13]. In summary, in all countries but one (Spain) the value range is shorter for EQ-5D-Y-3L value sets [10, 11, 14–19] (i.e., the value for the worst health state is higher) than that for EQ-5D-5L value sets in the corresponding countries. In eight of the nine cases where both EQ-5D-Y-3L and EQ-5D-5L value sets exist, the percentage of states with a value < 0 (i.e., considered by respondents on average to be 'worse than dead') is lower in EQ-5D-Y-3L than in EQ-5D-5L. These differences between the value scale for child and adult health states are especially marked in Asian countries [7].

The relative importance of dimensions also differs. All EQ-5D-Y-3L value sets published to date rate problems with pain/discomfort as the worst aspect of child HRQoL, and all except one rate problems with self-care as the least important aspect of HRQoL [13]. This order of dimension importance differs from that for EQ-5D-5L value sets [12]. These results may have intuitive explanations in adults' views about children, for example, adult respondents in valuation studies may perceive 'self-care' to have a different meaning when considered for children rather than for themselves; adults may consider functional issues around self-care and usual activities in children as better able to be mitigated through adult support.

Interpreting these differences between values for child and adult HRQoL is complex. There are differences in the descriptive systems being valued, combined with multiple differences in the valuation methods and how they are framed. The observed narrower scale of values may be attributable to a number of factors operating independently and in combination. These are summarized in Table 1.

First, the differences in values could (a) reflect differences in what is being valued in each case, i.e., differences between the descriptive systems of the child and adult instruments. The process of adapting the EQ-5D-Y to be appropriate for describing children's health states may have resulted in health state descriptions that are considered less severe by the adult respondents *valuing* the states. For example, in the functional dimensions (mobility, self-care, and usual activities), the worst problems are 'unable to' for the EQ-5D-5L compared with 'a lot of problems' in the EQ-5D-Y-3L. The dimension 'anxiety/depression' in the EQ-5D-5L is 'worried, sad or unhappy' in the EQ-5D-Y-3L, and the worst level of problem in the later is 'very,' compared with 'extremely' anxious or depressed in the EQ-5D-5L. If values for EQ-5D-Y-3L are higher for this reason, this is arguably a legitimate reflection of relevant social preferences. Note that the effect on QALY estimates will be a product both of the measurement properties of EQ-5D-Y-3L, and the properties of values for it. There is at least a possibility that the less severe level descriptors mean that children or their proxies are more likely to report health problems using the most severe levels; taken in combination with higher *values* for those states, the overall effect on HRQoL (compared with that for similar underlying health issues in adults) may partially cancel out.

The higher values for EQ-5D-Y-3L could also (b) reflect the genuinely held views of some adult respondents that, for any given level of health state description, children are less badly affected in terms of their overall HRQoL than adults, for example, because they think that children can 'bounce back' or more rapidly adapt to poor health. Powell et al. report mixed views of adults on whether health problems are better, worse, or about the same when experienced by children as opposed to adults. At least *some* of the respondents felt children to be more resilient and benefited from having more existing support [20]. Whether there are sufficient numbers of people with such views to drive the difference in the value scale of most EQ-5D-Y-3L value sets is unknown.

In contrast, there is evidence to suggest that (c) the particular perspective adult respondents are asked to adopt when valuing EQ-5D-Y-3L health states exerts an influence on values [21–24]. That is, being asked to evaluate a health state in someone else, as opposed to in oneself, may result in less willingness to trade-off life-years, regardless of whether the other person whose health is being considered is a child. The difference in perspective taken in valuing EQ-5D-Y-3L and EQ-5D-5L means that, even if these values are a legitimate reflection of respondents' preferences in each case, the values that are produced are not comparable. Research appears to suggest that the consideration of ill health in others as opposed to in oneself does exert an effect on values, but that there is also a distinct effect on values of the states under consideration being those of a child.

If the higher values for EQ-5D-Y-3L were known to be attributable to (a) and (b), this might suggest there is no problem to be solved in health state valuations: the observed values are capturing relevant preference information. The higher HRQoL values arising from (b) may be surprising or even seem abhorrent (e.g., where they lead to lower estimated QALY gains for improving HRQoL in children than in adults, de-prioritizing these interventions *ceteris paribus*). However, if the values accurately reflect adults' preferences regarding child health, and providing we accept the normative basis for eliciting *adults'* preferences for child HRQoL, they arguably represent a legitimate basis for decision making. Whereas if values are biased as in (c) and (d), this suggests a need to mitigate that bias.

The difficulty in empirically disentangling the various factors above mean that interpreting the 'child effect' on values remains ambiguous. However, to date, there has arguably been inadequate consideration of the role of (d). The use of HRQoL values for computation of QALYs requires values be anchored at dead = 0[1], which may invoke special considerations regarding survival and premature death in children that are not related to children's HRQoL per se. Given evidence suggesting adult respondents are 'unwilling to trade' off life-years in time trade-off (TTO) valuations of children's HRQoL [15–18, 25], how this should affect the interpretation and use of values for child HRQoL?

3 What Are Respondents' Underlying Preferences for Health and Survival in Children?

A widely accepted minimum requirement of values to be used in a CEA is that they are anchored at 0 (dead) and 1 (full health) [26, 27]. Methods that meet this requirement include TTO, standard gamble, and variants of DCE incorporating dead or duration in choice sets. All three methods value HRQoL (the thing to be valued, or 'valuand') by identifying the amount of something else (the 'numeraire') that respondents are willing to give up. In the case of TTO, the numeraire is life-years; in standard gamble, it is probability of survival. People have preferences both regarding the valuand *and* the numeraires in which value is expressed, for example, time preference and risk attitudes, which are known to affect the values from TTO and standard gamble, respectively. Considering that children and adults are in different phases of their lives, it is plausible that the higher values for child health not only reflect differences in perceived HRQoL but also different preferences about the numeraire.

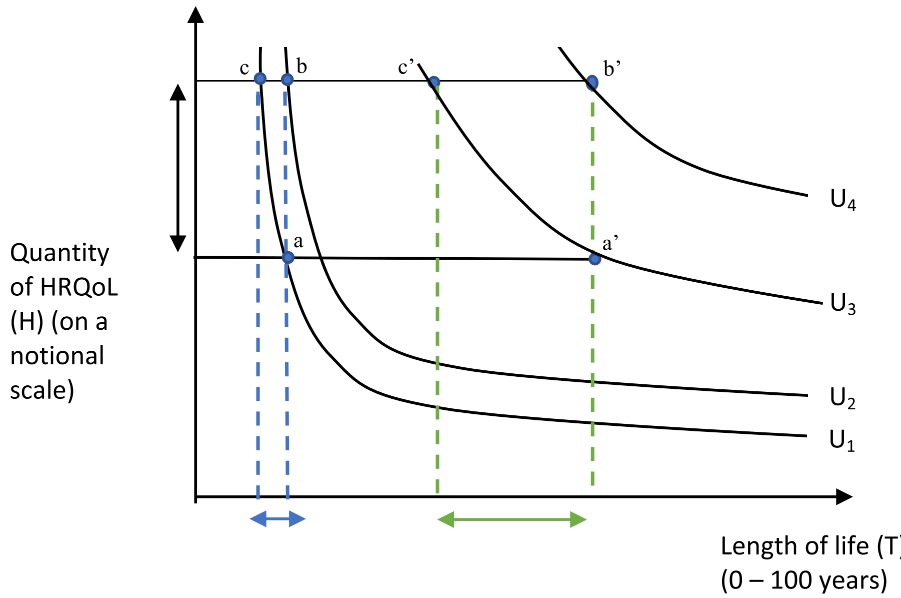
The impact of time preferences on TTO values for adult and child HRQoL states can be explained more formally using utility theory. Differences in adult respondents' willingness to trade off life-years observed in TTO valuations of child and adult HRQoL can be thought of in terms of the characteristics of respondents' utility functions that the TTO tasks aims to identify. Each respondent has (or constructs, in reaction to the tasks we present them) a utility function for HRQoL, U(H), and for life-years, U(T). The TTO establishes the value of a given state Hi on a scale anchored at 0 (dead) and 1 (full health) by the trade-offs respondents state they are willing to make between changes in HRQoL (H) and years of life (T). The task attempts to find the T in full health (Hf) that has the same utility (lies on the same indifference curve) as 10 years in Hi. Setting the years spent in full health at a value of U(Hf) = 1 means that once this point of indifference is identified, the value for Hi can be inferred. The conventional TTO is analogous to a Hicksian *compensating variation for a gain* [28].

A key thing to realize is that the values TTO produces depend on the nature of respondents' underlying utility functions in both H and T: the marginal rate of substitution between H and T ($MRS_{H,T}$) determines the TTO value which, in turn, is determined by the ratio of the marginal utility (MU) in each (MU_T/HU_H). This presents no issue provided one can assume that MU_T is constant across all observations being compared. In contrast, to the extent that, over some range of years constituting 'childhood', the marginal utility (from the adult general public's point of view) is higher for an extension in the length of life in children than for an increase in health status in children (i.e., $MU_T > MU_H$), considerations about survival will dominate trade-offs between the two, and the resulting $MRS_{H,T}$ will be higher than when adult health states are being considered.

This is illustrated in Fig. 1, which shows an identical improvement in H for children (shown by distance a–b) and adults (distance a′–b′), in each case increasing total utility (from U_1 to U_2 in the case of a child, and U_3 to U_4 in the case of an adult). Given this increase in H, the TTO aims to find the (smaller) amount of T that is considered to be equivalent in utility terms. If a respondent considers the MU_T is very high for children relative to MU_H, a relatively small amount of T, shown by distance b–c, needs to be sacrificed

[1] Whether the estimation of QALYs *requires* HRQoL values to be anchored at dead = 0 is itself debatable, but is an issue relevant to values for both adult and child HRQoL.

Fig. 1 Indifference curves representing the stated preferences of adult members of the general public regarding health-related quality of life (HRQoL) and life-years (by age)



to return from U_2 to U_1. In the case of adult states, the MU_T may be closer to MU_H, and the amount of T that needs to be traded off to return from U_4 to the original indifference curve U_3, shown by distance b′–c′, is greater. This leads to lower TTO values for adult health states, compared with those for children.

Note that even if respondents' MU_H is *completely constant with respect to age,* a diminishing MU_T will produce a higher $MRS_{H,T}$ in younger ages, and therefore produce higher TTO values for child HRQoL. To put it bluntly: even if improvements in HRQoL are considered by the adult general public to be special and more important in children than in adults, that could be masked (or completely offset) if the length of life in children is considered by the adult general public to be of even greater importance, resulting in higher values for child HRQoL. The stated preference tasks we use do not allow us directly to observe these underlying preferences in H and T, so the role of preferences regarding T and their effect on values is generally not explicitly identified.

An important implication is that we should not directly compare TTO values for adults and children in the absence of knowledge about preferences for time. A value of 0.7 for adults and 0.7 for children may not refer to the same level of HRQoL. This complicates the use of adult and child values in applications such as an HTA, as they are in effect measured in different currencies.

An issue is that U(T) is not observed, which means that U(T) and U(H) are perfectly confounded. This means that we do not know the appropriate interpretation of the observation that TTO values for child health states exceed TTO values for adult health states: do the same changes in health affect overall HRQoL less in children than adults, or are the differences due to variations in U(T)? Pertinent questions regarding the interpretation of differences in values for the health of children and adults therefore concern the shape of MU_T. Is total utility in H different in younger ages compared to adulthood? How does the shape of MU_H compare with the shape of MU_T?

The average views of the adult general public regarding the marginal utility of successive improvements in HRQoL (MU_H), independent of time, in children *compared to adults* could, in principle, be lower, the same, or higher. It is plausible that special concerns with children's well-being mean that members of the general public consider the marginal utility of improvements in HRQoL in children to be higher than that of corresponding HRQoL improvements in adults. It is also plausible that there is no generalizable difference in *overall* utility in HRQol between children and adults, but rather, more subtle differences in which aspects (dimensions) of HRQoL matter more/less for each.

Members of the adult general public taking part in stated preference studies may hold views that children are different from adults, for many reasons: they are vulnerable, and it is generally accepted that adults and society have a moral duty to protect and care for children. In principle, this could manifest in additional value for extending life (U(T)), or in improving quality of life (U(H)) in children, or both. These could cancel out in a TTO task if the ratio between MU_T and MU_H remains constant. However, children may be considered *particularly* deserving of opportunities to *extend* life because they have had less opportunity to experience it than adults. If this is the case, the additional value of life extension in children would manifest as a reduced willingness to trade off life-years, or, put another way, a greater willingness

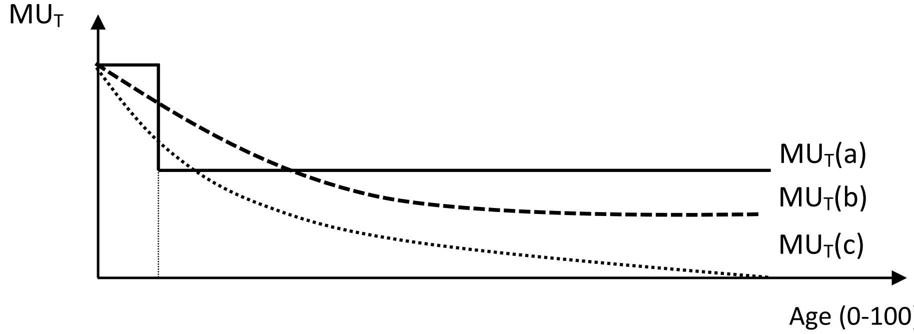


Fig. 2 Marginal utility in additional life-years in full health by age

to tolerate poor HRQoL in children, yielding relatively high HRQoL values. Strong preferences regarding child survival impacting on values for child HRQoL is therefore one possible explanation for the characteristics of values for EQ-5D-Y-3L. However, it is also possible that other explanations (a)–(c) noted in the previous section also play a role. A key point is that we cannot know which is the case, until we understand how U(T) is reflected in the observed values.

A further key point is that, regardless of whether child HRQoL values are high because of the underlying strength of preference regarding child survival or because adult respondents genuinely feel that health problems 'matter less' for children than adults, the impact is a discontinuity in HRQoL values. If a child patient's health is being monitored longitudinally using EQ-5D-Y-3L and that child reaches an age threshold where the EQ-5D-Y-3L instrument is replaced with the adult EQ-5D instrument, the patient's preference-weighted health profile could change significantly at that timepoint, even if their state remains *descriptively* similar.

It is important to note that this 'child versus adult' dichotomy is somewhat artificial. The step change in values that is observed is determined by the age ranges for which the HRQoL instruments in question have been validated for use, and to which the relevant value sets have been applied, rather than by a priori characteristics of the general public's preferences[2]. For example, we could hypothesize that rather than MU_T being 'kinked' about a single age ('childhood versus adulthood'), as in MU_T(a) in Fig. 2, it could be diminishing and then become constant as shown in MU_T(b), or always diminishing in age as shown in MU_T(c). For the formalization of the problem we are addressing, the exact shape of MU_T is not material; the crucial point is that MU_T not being constant has a bearing on how values may or may not be interpreted.

[2] In the case of EQ-5D-Y-3L, the step change in values is a product both of the age range for which the instrument was designed, and the fact that the values for the instrument are based on adults' stated preferences for a 10-year-old child.

While this exposition has focused on the TTO, the same challenge is expected to arise with *any* method for valuing child HRQoL that involves anchoring at dead = 0. At its core, the observed TTO values are an expression of underlying beliefs about what health states are worth living. If the numeraire was to change (to risk, for example), the direction or size of the discontinuity may change, but changing the numeraire is unlikely to resolve the fundamental issue. Indeed, evidence suggests this is the case: a study testing a range of methods for anchoring child HRQoL values at dead = 0 found that child values exceeded adult values in all cases [29]. The incommensurability of child and adult values can thus be regarded as a fundamental property associated with the scale on which values are measured, rather than a property of one specific method.

To summarize: the observed differences in HRQoL values noted in Sect. 2 may arise at least in part from special considerations that adults bring to bear on trade-offs between survival and HRQoL when considering children, which they do *not* bring to bear when valuing HRQoL for themselves. This effect may be difficult to discern from other factors, including the possibility that adults on average genuinely feel that poor health in children causes less disutility than poor health in adults.

4 What Are the Implications of Non-comparability of Values for the Application of Age-Specific Values for HRQoL in CEA?

If the characteristics of values for child HRQoL are at least partly a product of the fundamental difference in perspective adults respondents are asked to adopt, and confounding between U(H) and U(T) in their responses to the task, this would have consequences for how the resulting values may be applied in the context of economic evaluations of healthcare technologies. The characteristics of EQ-5D-Y-3L value sets would *not* present an issue for their use in CEAs only under the circumstances where (a) decision makers hold ring-fenced budgets for child health (and opportunity costs are limited to QALYs foregone from that same budget) and (b) the benefits from treatments being evaluated are experienced only during childhood.

These conditions are arguably rarely likely to hold. The improvements in quality or length of life that result from the avoidance of poor health in childhood often extend beyond childhood into adulthood. Poor child health can cause developmental delays that limits longer term HRQoL. Preventing premature death in children yields additional years of life that may extend into adulthood. Many CEAs of child health interventions therefore entail lifetime cost-effectiveness modeling, requiring parameters to consistently capture HRQoL spanning childhood and adulthood. At some point in these models, where health states shift from being described using a child-specific HRQoL instrument to an adult HRQOL instrument, the different properties of the values accompanying these instruments could cause an artefactual 'cliff edge' in HRQoL, which is unrelated to any change in underlying health status, but arises purely from the different characteristics of the value sets.

Further, in principle, the use of CEA evidence (e.g., in HTA processes) involves resource allocation decisions across the full age range of users and potential users of the healthcare system. If the relevant health system maximand can be characterized as QALY maximization, this requires comparability of outcome measurement and valuation such that estimates of QALYs gained and QALYs foregone can be interpreted like for like. If the values for child and adult health lack direct comparability, this becomes problematic. Evaluation of opportunity cost becomes more challenging and questions about intergenerational fairness may be raised. Economists would need to consider how to accommodate values with such characteristics in economic evaluations.

In practice, the recommendations made by HTA bodies are not based solely on the cost effectiveness of technologies. Other factors, including equity, and special considerations concerning the characteristics of the patients or condition (e.g., severity, rarity, end of life) may be considered relevant. Health technology assessment bodies may well wish to prioritize QALY gains for children, because "Children hold a special place in any society. They are among its most vulnerable members and children's health and well-being affects both their lifelong opportunities and the future for us all. One of the basic responsibilities of any society is to care for children in the best and most compassionate manner possible" [30]. Health technology assessment bodies have the ability to take special considerations regarding children into account via their deliberative processes or by applying explicit weights to child QALY gains. However, the ability to do so in a clear and unambiguous manner is compromised if special considerations about child survival are already exerting an effect on values for child HRQoL because of adult respondents' unwillingness to trade off life-years when considering child health problems.

5 Do We Need Age-Specific Values? What Are the Alternatives?

The issues outlined above arise because, in valuing EQ-5D-Y-3L, adults' stated preferences are sought, and in eliciting them we are both describing health states using an age-specific tool and asking them to value the states by imagining them to be experienced by a child. The resulting EQ-5D-Y-3L values obtained from adults are therefore 'age specific' in both *what* is being valued, and *how* it is valued. There are, however, potentially useful alternatives to the valuation of EQ-5D-Y-3L that do not require eliciting age-specific values.

A wide variety of different approaches to the challenge of valuation have been adopted by other pediatric HRQoL instrument developers and researchers. For example, the UK and Dutch value sets for the CHU9D are based on adults being asked to evaluate health states described by the instrument, but as they imagine these to be experienced from their own (adult) perspective [31, 32]. In effect, adult respondents are 'blinded' to the states under consideration being descriptions of child health states, so special considerations relating to child survival are avoided. However, if adults genuinely believe that poor health in children has less effect on utility (argument (b) in Table 1), such approaches would in effect be overriding and ignoring of that preference.

Values for the Health Utilities Preschool (HuPS) instrument for preschool children are obtained by mapping HuPS to the HUI3 instrument and using the value set that exists for the latter [33]. This enables both HuPS and HUI3 descriptive systems to be linked to a single age-invariant set of values relevant to both children and adults, avoiding the issues with age-specific values we have identified.

In the case of both the Dutch and UK values for CHU9D, and the use of HUI3 values to preference-weight HuPS, values are not age specific. This facilitates a consistent basis for preference-weighting child and adult HRQoL and ensures comparability of QALYs. However, the desirability of using age-invariant values would need to be explored: they do not reflect societal views about HRQoL in children and this may raise normative questions related to what we are trying to maximize with economic evaluations in healthcare. It leaves to HTA bodies the task of addressing any special societal considerations about child QALYs via their deliberative processes, or perhaps via QALY weights.

All the alternatives above rely on the use of *adults'* preferences to value child HRQoL instruments. This is also the case with the protocol for valuing the EQ-5D-Y-3L valuation protocol: adults' preferences are sought for both normative and pragmatic reasons [5, 34]. A further option is to ask children to value age-specific health state descriptions from their own perspective. For example, the Australian values for CHU9D are based on children's responses to stated preference tasks, evaluated from a 'self' perspective [35]. Research has established the feasibility of seeking stated preferences of adolescents to establish the relative importance of HRQoL dimensions [36, 37]. However, there are practical and ethical challenges with obtaining stated preferences from children, particularly around the tasks required to anchor values at dead = 0. The valuation of child states by children may appeal to other normative principles ('noting about me, without me'). Further, the use of a child 'own health' perspective, similar to the 'own health' perspective adopted by adults in valuing adult HRQoL states, could potentially reduce the gaps between child and adult values caused by (c) in Table 1.

There is no consensus on the question of whose preferences should be sought in valuing child HRQoL [34, 38, 39], and the theoretical foundations of CEA offer little guidance on that, or on the merits of age-specific versus age-invariant value sets. In the absence of any consensus on these fundamental questions, the basis for evidence on measuring and valuing child HRQoL currently being submitted to HTA varies in important ways that HTA bodies may not be fully aware of. The values produced from such disparate efforts will have very different properties and characteristics and may have quite varying interpretations and comparability (or lack thereof) with values for adult health states—introducing additional uncertainty and complexity for decision makers attempting to use this evidence.

Arguably, the responsibility for what approach to take in valuing child HRQoL goes beyond the remit of instrument developers or indeed of health economists. Economists do not have a mandate to make the necessary value judgments although can (and arguably, have a responsibility to) "help those who are charged to make such judgements (and) who need to understand the options and their consequences" [40] (p. 333). A lack of attention to these fundamental questions is an important shortcoming of the existing literature on valuing pediatric HRQoL. In raising these issues, we recognize the risks of appearing to undermine the progress that has been made with developing value sets for the EQ-5D-Y-3L and other pediatric HRQoL instruments. This paper can be read as a 'call to arms' to examine these issues more thoroughly, and to ensure that values, regardless of which instrument they apply to, are relevant and useful to decision makers.

This in turn points to the importance of engagement with decision makers and other appropriate authorities about these methods choices and their implications. The EuroQol Group already recommends stakeholder engagement as a prerequisite for EQ-5D-Y-3L value set studies [6]. However, decision makers may lack familiarity with the relevant issues, and may lack clear views or underlying principles on which to form them. Further, different stakeholders may have conflicting views, making it difficult to achieve a consensus on methods choices. Consultation therefore requires a deep commitment to informing and engaging, and helping stakeholders to form carefully considered positions. There may be merit in bringing decision makers and other stakeholders together to consider the issues, via an international HTA roundtable, although ultimately, methods choices will need to reflect the judgments and needs of local decision makers and the populations and health systems they represent. Researchers need to be committed to informing users of the implications of methods choices and value set characteristics for decision making, and, notwithstanding the appeal of standardized protocols, to remain open to the possibility that different methods choices may be appropriate in some decision contexts.

In the meantime, given the EQ-5D-Y-3L value sets now available for use in a CEA, research is needed to isolate the extent to which differences in adults' valuations of child health reflect health state descriptions being perceived as less severe, and to explore the way that differences in description and valuation of health states combine in the assessment of overall HRQoL using child and adult instruments. Further, given the potential for an artefactual cliff-edge change in the assessment of HRQoL in the transition between child and adult instruments in, for example, longitudinal data collection or in cost-effectiveness models, methods for addressing that (e.g., smoothing the transitions in values, or establishing an 'exchange rate' or adjustment factor between child and adult HRQoL) could be developed and tested. However, these ex-post adjustments should not replace a thorough examination of the premise for valuing children's HRQoL states.

Acknowledgements The authors are grateful to Aki Tsuchiya for detailed comments on an earlier version of this paper, and to three *Pharmacoeconomics* reviewers for helpful comments on an earlier version. ND is grateful to Theodore Aspinall-Devlin for reminding her why this topic is so important.

Funding Open Access funding enabled and organized by CAUL and its Member Institutions.

Declarations

Funding Work on this paper was supported by EuroQol Research Foundation Grant 202-2020RA. Opinions expressed in the paper are not necessarily those of the EuroQol Group or of the Younger Populations Working Group. Mark Jit was also supported by the National Institute for Health Research (NIHR) Health Protection Research Unit in Modelling and Health Economics (NIHR200908) and in Immunisation (NIHR200929).

Conflicts of interest/competing interests Nancy J. Devlin, Richard Norman, Donna Rowen, Elly Stolk, and Barend van Hout are members of the EuroQol Group. Tianxin Pan, Mark Sculpher, and Mark Jit have no conflicts of interest that are directly relevant to the content of this article.

Ethics approval Not applicable.

Consent to participate Not applicable.

Consent for publication Not applicable.

Availability of data and material Not applicable.

Code availability Not applicable.

Authors' contributions ND was responsible for conceiving this work, securing funding, and leading the writing. ES and RN contributed input at every stage of writing. MS, TP, MJ, DR, and BvH provided detailed input at two stages of drafting. All authors have seen and approved submission of the final manuscript.

References

1. Kennedy-Martin M, Slaap B, Herdman M, van Reenen M, Kennedy-Martin T, Greiner W, et al. Which multi-attribute utility instruments are recommended for use in cost-utility analysis? A review of national health technology assessment (HTA) guidelines. Eur J Health Econ. 2020;21(8):1245–57. https://doi.org/10.1007/s10198-020-01195-8.
2. 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;19:887–97. https://doi.org/10.1007/s11136-010-9649-x.
3. Euroqol Research Foundation. EQ-5D-Y user guide. 2020. Version 2.0. https://euroqol.org/publications/user-guides/. Accessed 2 Jul 2024.
4. Kreimeier S, Oppe M, Ramos-Goñi JM, Cole A, Devlin N, Herdman M, et al. Valuation of EuroQol Five-Dimensional Questionnaire, Youth Version (EQ-5D-Y) and EuroQol Five-Dimensional Questionnaire, Three-Level Version (EQ-5D-3L) health states: the impact of wording and perspective. Value Health. 2018;21(11):1291–8. https://doi.org/10.1016/j.jval.2018.05.002.
5. Ramos-Goñi JM, Oppe M, Stolk E, Shah K, Kreimeier S, Rivero-Arias O, et al. International valuation protocol for the EQ-5D-Y-3L. Pharmacoeconomics. 2020;38(7):653–63. https://doi.org/10.1007/s40273-020-00909-3.
6. Devlin N, Pan T, Kreimeier S, Verstraete J, Stolk E, Rand K, Herdman M. Valuing EQ-5D-Y: the current state of play. Health Qual Life Outcomes. 2022;20:105. https://doi.org/10.1186/s12955-022-01998-8.
7. Devlin N, Roudijk B, Viney R, Stolk E. EQ-5D-Y-3L value sets, valuation methods and conceptual questions. Pharmacoeconomics. 2022;40(Suppl. 2):123–7. https://doi.org/10.1007/s40273-022-01226-7.
8. Hill H, Rowen D, Pennington B, Wong R, Wailoo A. A review of the methods used to generate utility values in NICE technology assessments for children and adolescents. Value Health. 2020;23(7):907–17.
9. Bailey C, Dalziel K, Cronin P, Devlin N, Viney R. How are child-specific utility instruments used in decision making in Australia? A review of Pharmaceutical Benefits Advisory Committee public summary documents. Pharmacoeconomics. 2022;40:157–82. https://doi.org/10.1007/s40273-021-01107-5.
10. Shiroiwa T, Ikeda S, Noto S, Fukuda T, Stolk E. Valuation survey of EQ-5D-Y based on the international common protocol: development of a value set in Japan. Med Decis Making. 2021;41(5):597–606. https://doi.org/10.1177/0272989X211001859.
11. Yang Z, Jiang J, Wang P, et al. Estimating an EQ-5D-Y-3L value set for China. Pharmacoeconomics. 2022. https://doi.org/10.1007/s40273-022-01216-9.
12. Devlin N, Roudijk B, Ludwig K, editors. Value sets for EQ-5D-5L. A compendium, comparative review and user guide. New York: Springer; 2022.
13. Roudijk B, Pan T, Abel-Olsen J, Devlin N. A comparative analysis of EQ-5D-Y-3L value sets and comparison to EQ-5D-5L value sets. Paper to be presented to the 2023 Rome Plenary meeting of the EuroQol Research Foundation.
14. Roudijk B, Sajjad A, Essers B, et al. A value set for the EQ-5D-Y-3L in the Netherlands. Pharmacoeconomics. 2022. https://doi.org/10.1007/s40273-022-01192-0.
15. Dewilde S, Roudijk B, Tollenaar NH, et al. An EQ-5D-Y-3L value set for Belgium. Pharmacoeconomics. 2022. https://doi.org/10.1007/s40273-022-01187-x.
16. Ramos-Goñi JM, Oppe M, Estévez-Carrillo A, Rivero-Arias O, IMPACT HTA HRQoL Group. Accounting for unobservable preference heterogeneity and evaluating alternative anchoring approaches to estimate country-specific EQ-5D-Y value sets: a case study using Spanish preference data. Value Health. 2022;25(5):835–43. https://doi.org/10.1016/j.jval.2021.10.013.
17. Kreimeier S, Mott D, Ludwig K, et al. EQ-5D-Y value set for Germany. Pharmacoeconomics. 2022. https://doi.org/10.1007/s40273-022-01143-9.
18. Rencz F, Ruzsa G, Bató A, et al. Value set for the EQ-5D-Y-3L in Hungary. Pharmacoeconomics. 2022. https://doi.org/10.1007/s40273-022-01190-2.
19. Fitriana TS, Roudijk B, Purba FD, et al. Estimating an EQ-5D-Y-3L value set for Indonesia by mapping the DCE onto TTO values. Pharmacoeconomics. 2022. https://doi.org/10.1007/s40273-022-01210-1.
20. Powell PA, Rowen D, Rivero-Arias O, Tsuchiya A, Brazier JE. Valuing child and adolescent health: a qualitative study on different perspectives and priorities taken by the adult general public. Health Qual Life Outcomes. 2021;19(1):222. https://doi.org/10.1186/s12955-021-01858-x.
21. Dewilde S, Janssen MF, Lloyd AJ, Shah K. Exploration of the reasons why health state valuation differs for children compared with adults: a mixed methods approach. Value Health. 2022;25(7):1185–95. https://doi.org/10.1016/j.jval.2021.11.1377.
22. Reckers-Droog V, Karimi M, Lipman S, Verstraete J. Why do adults value EQ-5D-Y-3L health states differently for themselves than for children and adolescents: a think-aloud study. Value

Health. 2022;25(7):1174–84. https://doi.org/10.1016/j.jval.2021.12.014.
23. Lipman SA, Essers BAB, Finch AP, et al. In a child's shoes: composite time trade-off valuations for EQ-5D-Y-3L with different proxy perspectives. Pharmacoeconomics. 2022;40(Suppl. 2):181–92. https://doi.org/10.1007/s40273-022-01202-1.
24. Lipman SA, Zhang L, Shah KK, Attema AE. Time and lexicographic preferences in the valuation of EQ-5D-Y with time trade-off methodology. Eur J Health Econ. 2023;24(2):293–305. https://doi.org/10.1007/s10198-022-01466-6. (**Erratum in: Eur J Health Econ. 2022;23(9):1613–1615**).
25. Yang Z, Jiang J, Wang P, Wu J, Fang Y, Feng D, et al. Using time trade-off values to estimate EQ-5D-Y value sets: an exploratory study. Value Health. 2023. https://doi.org/10.1016/j.jval.2023.03.003.
26. Devlin N, Finch A, Parkin D. Guidance to users of EQ-5D-5L value sets. In: Devlin N, Roudijk B, Ludwig K, editors. Value sets for EQ-5D-5L. Chapter 6. New York: Springer; 2020.
27. Brazier J, Ratcliffe J, Saloman J, Tsuchiya A. Measuring and valuing health benefits for economic evaluation. Oxford: Oxford University Press; 2017.
28. Buckingham K, Devlin N. A theoretical framework for TTO valuations of health. Health Econ. 2006;15(10):1149–54. https://doi.org/10.1002/hec.1122.
29. Shah KK, Ramos-Goñi JM, Kreimeier S, Devlin NJ. An exploration of methods for obtaining 0 = dead anchors for latent scale EQ-5D-Y values. Eur J Health Econ. 2020;21(7):1091–103. https://doi.org/10.1007/s10198-020-01205-9.
30. Devlin N, Lovett R, Rowen D. Challenges in measuring and valuing childen's health related quality of life. Value and Outcomes Spotlight. 2021.
31. Stevens K. Valuation of the Child Health Utility 9D Index. Pharmacoeconomics. 2012;30:729–47.
32. Rowen D, Mulhern B, Steves K, Vermaire JH. Estimating a Dutch value set for the pediatric preference-based CHU9D using a discrete choice experiment with duration. Value Health. 2012;21(10):124–42.
33. Furlong W, Rae C, Feeny D, Ghotra S, Breakey VR, Carter T, et al. Generic health-related quality of life utility measure for preschool children (health utilities preschool): design, development, and properties. Value Health. 2022. https://doi.org/10.1016/j.jval.2022.07.015.
34. Rowen D, Rivero-Arias O, Devlin N, et al. Review of valuation methods of preference-based measures of health for economic evaluation in child and adolescent populations: where are we now and where are we going? Pharmacoeconomics. 2020;38:325–40. https://doi.org/10.1007/s40273-019-00873-7.
35. Ratcliffe J, et al. Valuing the Child Health Utility 9D: using profile case best worst scaling methods to develop a new adolescent specific scoring algorithm. Soc Sci Med. 2016;157:48–59.
36. Prevolnik Rupel V, Ramos-Goñi JM, Ogorevc M, Kreimeier S, Ludwig K, Greiner W. Comparison of adult and adolescent preferences toward EQ-5D-Y-3L health states. Value Health. 2021;24(9):1350–9. https://doi.org/10.1016/j.jval.2021.03.019.
37. Dalziel K, Catchpool M, García-Lorenzo B, Gorostiza I, Norman R, Rivero-Arias O. Feasibility, validity and differences in adolescent and adult EQ-5D-Y health state valuation in Australia and Spain: an application of best-worst scaling. Pharmacoeconomics. 2020;38(5):499–513. https://doi.org/10.1007/s40273-020-00884-9.
38. Petrou S. Methodological challenges surrounding QALY estimation for paediatric economic evaluation. Cost Eff Resour Alloc. 2022;20(1):10. https://doi.org/10.1186/s12962-022-00345-4.
39. Devlin NJ. Valuing child health isn't child's play. Value Health. 2022;25(7):1087–9. https://doi.org/10.1016/j.jval.2022.05.009.
40. Brouwer WBF, Culyer AJ, van Exel NJA, Rutten FFH. Welfarism vs. extra-welfarism. J Health Econ. 2008;27(2):325–38.