



A systematic literature review of preference-based health-related quality-of-life measures applied and validated for use in childhood and adolescent populations in sub-Saharan Africa

Journal:	<i>Value in Health Regional Issues</i>
Manuscript ID	VIHRI-CEEWAA-2020-0085.R1
Article Type:	Systematic Literature Review
Health Areas List:	AIDS / HIV < Health Areas, Other health conditions < Health Areas, Other infectious disease < Health Areas
Methods of Interest List:	PRO/QoL: psychometrics < Methods of Interest, Systematic Literature Reviews < Methods of Interest, Preferences: methods < Methods of Interest
Keywords Enter Your Own:	childhood, preference-based measures, PBMs, health-related quality of life, HRQoL, validation, psychometrics, sub-Saharan Africa, SSA

SCHOLARONE™
Manuscripts

1
2
3 Title: A systematic literature review of preference-based health-related quality-of-life
4 measures applied and validated for use in childhood and adolescent populations in sub-Saharan
5 Africa
6
7
8
9

10
11
12 **Running title:** Review of childhood PB HRQoL measures in sSA
13
14

15
16 **Precis:** Application of childhood preference-based measures (PBMs) in sub-Saharan Africa
17 has been limited. There is cross-cultural validation need for PBMs in this region to measure
18 HRQoL.
19

20
21
22
23 **Word count:**253 (abstract)

24 **Word count:** 4,135 (excluding Abstract, Reference, Figures, Tables and Appendices)

25 **Number of Pages:** 24

26 **Number of Figures:** 2

27 **Number of Tables:** 6 (excluding Additional files)

28 **Number of References:**35
29

30
31 **Appendix:**

32 Pages: 15

33 Figures:0

34 Tables: 4
35

36
37
38 **Key words:** childhood, PBM, HRQoL, validation, sSA
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Review of childhood PB HRQoL measures in sSA

ABSTRACT

Objectives

Consideration of health status in children and adolescents now includes broader concepts such as health-related quality-of-life (HRQoL). Globally there is a need for relevant preference-based HRQoL measures (PBMs) for use in children and adolescents, yet measurement of HRQoL in these groups presents particular challenges. This paper systematically reviews the available generic childhood PBMs, and their application and cross-cultural validation in sub-Saharan African (sSA).

Methods

A systematic review of published literature from 01/01/1990 to 02/08/2017 was conducted using MEDLINE (through OVID Sp), EMBASE (OVID Sp), EconLit (EBSCOhost), PsycINFO, Web of Science and PubMed.

Results

220 full-text articles were included in a qualitative synthesis. Ten generic childhood PBMs were identified, of which nine were adapted from adult versions and only one developed specifically for children. None of the measures were originally developed in sSA or other resource-constrained settings. The HUI3 and the EQ-5D-Y-3L were the only measures that had been applied in sSA settings. Further, the HUI3 and the EQ-5D-Y-3L were the only generic childhood PBM that attempted to establish cross-cultural validation in sSA. Five of the six of these validation studies were conducted using the EQ-5D-Y-3L in a single country, South Africa.

Conclusions

Review of childhood PB HRQoL measures in sSA

1
2
3 The findings show that application of generic childhood PBMs in sSA settings has hitherto been
4
5 limited to the HUI3 and EQ-5D-Y-3L. Most adaptations of existing measures take an absolutist
6
7 approach, which assumes that measures can be used across cultures. However, there is also need
8
9 to ensure linguistic and conceptual equivalence and undertake validation across a range of sSA
10
11 cultural contexts.
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Review of childhood PB HRQoL measures in sSA

Highlights:

i. What is already known about the topic?

Since the 1990s, there has been growing interest in developing and using child and adolescent (childhood) preference-based health-related quality of life measures (PBMs). The increased prominence of childhood PBMs has been driven by their use in measuring health-related quality-of-life (HRQoL) as part of economic evaluations, thereby informing health care policies.

ii. What does the paper add to existing knowledge?

While this paper examines existing childhood PBMs and summarises their psychometric properties, it focuses on the validation and application of these measures in the sub-Saharan African (sSA) region. Development of existing childhood PBMs have been restricted to high-income countries, and none has been developed in low-income countries. Validation and application of childhood PBMs in the sSA region has been limited to the HUI3 and EQ-5D-Y-3L in two countries.

iii. What insights does the paper provide for informing health care-related decision making?

Findings from health technology assessments (HTA) and economic evaluations are increasingly used to inform health policy in sSA. These studies should incorporate local preferences for health outcomes. Preference-based values for health outcomes are commonly generated using tools that are adapted to local contexts. If the process of adapting these measures in sSA contexts is not robust, the findings from subsequent HTAs and economic evaluations may not be applicable.

1 Review of childhood PB HRQoL measures in sSA
2

3 **Introduction:**
4

5 In sub-Saharan Africa (sSA) i.e. the region below the Sahara desert, there has been a remarkable
6 decline in mortality rates in infants and under 5-year olds since 2000,¹ partly driven by policy
7 responses to the Millennium Development Goals (MDGs).² In 2015, the United Nations set 17
8 sustainable development goals as a replacement for MDGs for nations to achieve by 2030. Goal
9 number three focuses on ‘good health and wellbeing’, specifically targeting child and maternal
10 mortality and chronic disease reduction.³ As childhood mortality continues to fall, and increasing
11 numbers of children survive into adolescence and beyond, future interventions will increasingly
12 focus on improving wider outcomes including children’s health-related quality of life (HRQoL).⁴
13
14
15
16
17
18
19
20
21
22
23
24
25

26 Preference-based HRQoL measures (PBMs) are designed to “assess patient preferences across
27 broad areas including symptoms, physical functioning, work and social activities, and mental well-
28 being”.⁵ PBMs have increased in prominence over recent years because of their use in economic
29 evaluations, which often inform policy.⁶ This can only be meaningful if PBMs are relevant, valid
30 and robust. However, most generic PBMs have been developed in the English language⁷ and, as
31 the development of new measures is time and resource intensive, use of PBMs in other contexts
32 has mostly relied on translating and/or adapting existing measures, with an implicit assumption
33 that adaptation produces an equivalent measure.⁸ Adaptation of an existing measure requires an
34 iterative process of translation and back translation with a qualitative review to establish both
35 linguistic and conceptual equivalence.⁽⁷⁾ In addition, the adapted measure needs to be re-validated
36 using standard psychometric methods to re-establish validity and reliability for the new form of the
37 measure in the new context.
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1 Review of childhood PB HRQoL measures in sSA

2
3 Several reviews have assessed the development and application of childhood PBMs and non-PBMs
4 (restricted to ages 6-17 years),⁹⁻¹¹ their cross-cultural validation,^{12,13} and availability of measures
5 for young children (<8 years) in alternative settings.¹⁴ However, to our knowledge, no review has
6 specifically studied the use or validation of childhood PBMs across sSA.
7
8
9
10

11
12
13
14 We carried out a systematic review of generic childhood PBMs, particularly focussing on their use
15 and cross-cultural validation in sSA settings. The objectives of the review were threefold: (1) to
16 identify existing generic childhood PBMs and summarise their psychometric quality; (2) to explore
17 which of these generic childhood PBMs have been applied in sSA; and (3) to assess the extent of
18 cross-cultural validation of generic childhood PBMs for application in sSA.
19
20
21
22
23
24
25
26
27

28 **Methods**

29
30 We conducted the literature searches using a two-stage process compatible with the preferred
31 reporting items for systematic reviews and meta-analyses (PRISMA) guidelines.¹⁵ For the first
32 stage, we searched the literature to identify existing childhood PBMs (objective 1). In the second
33 stage (to address objectives two and three), we refined our search strategy to identify all published
34 research studies that applied or cross-culturally validated any one of these PBMs in the sSA region.
35
36 Cross-cultural validation in our search criteria was defined as i) linguistic equivalence (achieved
37 through translation), ii) conceptual equivalence (achieved through focus group discussion or
38 cognitive interviews) and iii) psychometric validation. The peer reviewed literature published
39 between 1st January 1990 and 2nd August 2017 was examined for both stages of the review.
40
41
42
43
44
45
46
47
48
49
50

51 ***Search strategy***

1 Review of childhood PB HRQoL measures in sSA

2
3 In stage one, we carried out systematic searches using an intersection of four sub-groups of search
4 terms: 1) HRQoL terms; 2) childhood terms; 3) instrument terms; and 4) psychometric terms. The
5 psychometric terms were included in order to identify validated measures only. A full list of search
6 terms is provided in Appendix 1.1. The search was performed in six databases via Ovid: PubMed,
7 EMBASE, Web of Science, EconLit, PsycINFO and MEDLINE, as per COSMIN
8 recommendations.¹⁶

9
10
11
12
13
14
15
16
17
18
19 In stage two, we searched through the same databases using search terms for the identified generic
20 childhood PBMs and combined them with sSA terms to identify their application and/or cross-
21 cultural validation in sSA settings. The full list of identified measures and sSA search terms used
22 to address the second and third study objectives are provided in Appendix 1.2.

23 24 25 26 27 28 29 30 31 ***Inclusion and exclusion criteria***

32 To address the first study objective, studies reporting development of measures were included if:
33 (1) they reported the development of at least one childhood PBM (this was limited to the original
34 work by the developers only); 2) the target age(s) for completion of the measure included children
35 between 6-17 years; and (3) were published in the English language. Studies were excluded if they
36 reported measures designed for adults (≥ 18 years) or young children (< 6 years) as the review
37 focussed on measures with the potential for self-completion by children.¹⁷ Additionally, all studies
38 that reported on psychometric properties of the identified childhood PBMs were restricted to work
39 by the primary developers.

1 Review of childhood PB HRQoL measures in sSA
2

3 To address the second and third study objectives, studies were included if they had applied and/or
4 cross-culturally validated any measure identified at stage one of the review in a sSA setting. Studies
5 using generic childhood PBMs that were not carried out either in part or whole in a sSA setting
6 were excluded.
7
8
9
10
11
12
13

14 ***Procedures***

15
16 Three reviewers (LGN, KK and LS) independently assessed the titles, and subsequent abstracts, to
17 select potential full-text articles for screening. Study selection was agreed upon at each stage (title,
18 title and abstract, full article) by the three reviewers with disagreements resolved via consensus.
19
20
21
22
23
24
25

26 ***Data extraction and analysis***

27 To meet **objective 1**, for each identified generic childhood PBM the following was extracted:
28
29
30
31
32

33 *(i) General description of the measure*

34 The descriptive system consisting of the classification system; where, how and who developed the
35 classification system; target population; number of dimensions; response levels; and time taken to
36 respond.
37
38
39
40
41
42
43

44 *(ii) Underpinning utility theory*

45 This included the number of health of states generated by the classification system, the valuation
46 technique used, model used to extrapolate the utilities to all possible health states, range of utilities
47 for each measure and when and how these were generated.
48
49
50
51
52
53
54

55 *(iii) Psychometric validation*

Review of childhood PB HRQoL measures in sSA

The quality of each identified childhood PBM was appraised against standard psychometric criteria¹⁸⁻²¹ as outlined below.

Conceptual framework including a detailed description of the conceptual basis of the construct measured, is usually based on qualitative work with the target population.²⁰

Cross-cultural language and conceptual equivalence/adaptation is the extent to which a scale shows linguistic and conceptual equivalence, usually assessed through qualitative methodology, including translation, cognitive interviews or focus group discussions with the target population.⁸

Reliability, the extent to which the instrument is free from random error, was assessed in terms of *internal consistency*, defined using Cronbach's $\alpha \geq 0.70$ for group level comparisons and ≥ 0.90 for individual comparisons;¹⁹ and *test re-test* assessed by an Intraclass-Correlation Coefficient (ICC) ≥ 0.70 or Kappa $\kappa \geq 0.70$ for nominal scales.^{18,19,21}

Validity, the extent to which a scale measures what it purports to measure, was assessed using: *Content validity* which assesses the scale's conceptual breadth, including all aspects of the construct being measured;^{18,20} *Criterion validity*, which assesses how the scale correlates with a gold standard measure of the same construct administered at the same time (*concurrent*) or at some point in the future (*predictive*), defined using correlation ≥ 0.70 ;²¹ *Convergent validity*, the extent to which the scale is positively related to other scales that purport to measure the same thing; *Discriminant validity*, the extent to which the scale is not related to other scales that purport to measure different things, assessed using correlation coefficients; *Known groups validity*, which assesses the extent to which scores on the scale are different between two groups that are known to be different on another clinically defined variable;²¹ and *Responsiveness*, the ability of the scale to detect clinically important change over time, which was assessed using effect size statistics, score changes or standardized response means.¹⁸

1 Review of childhood PB HRQoL measures in sSA

2
3 *Interpretability of scores* refers to the degree to which one can easily assign meanings to an
4 instrument's quantitative scores, assessed in several ways including normative data.¹⁸⁻²¹

5
6
7 *Practicality* is the burden on those completing and/or those administering the measure and includes
8 completion time, and completion and response rates.²¹

9
10
11
12
13
14 In our review, all psychometric properties described above were classified as (+) if they met the
15 criteria; (-) if not; and (NR) if not reported. We assessed the psychometric properties as a whole
16 rather than using a single statistic as evidence of psychometric quality.

17
18
19
20
21
22
23 To achieve **objective 2**, all studies included in this second stage of the review were assessed for
24 their quality using the Critical Appraisal Skills Programme (CASP) toolkit for case-control
25 studies.²² The CASP tool consists of ten questions that address issues such as study aim and design,
26 data collection and analysis, and ethical issues raised by the published study.

27
28
29
30
31
32
33
34 In order to achieve **objective 3**, each identified generic childhood PBM was assessed according to
35 whether there had been: i) cross-cultural translation into any native language spoken in sSA; ii)
36 any conceptual equivalence through focus group discussion or cognitive interviews on word
37 meanings and changes to concepts to fit what would be locally equivalent into a sSA setting and
38 iii) psychometric validation. Cross-cultural validation was classified as (+) if it was appraised and
39 performed; (-) if appraised but not performed; and (NR) if not reported.

40
41
42
43
44
45
46
47
48
49
50
51 This review was not part of any clinical trial and was therefore not registered as would have been
52 required otherwise.

Review of childhood PB HRQoL measures in sSA

Results

Identifying childhood preference-based HRQoL measures (objective 1)

The first stage of our search identified 59,342 unique articles (Figure 1). 58,366 articles were excluded after screening the title and a further 734 articles excluded after screening the abstract. In addition to the remaining 242 articles, we identified 91 articles through snowball searching. We therefore reviewed the full texts of 333 articles for eligibility. 113 of these articles were excluded as they did not meet eligibility criteria; a large number (n=51) of these articles were excluded as they did not report use of a generic childhood PBM. In total, we identified 220 articles that reported development of at least one generic childhood PBM, from which we identified ten unique measures (Figure 1).

Identified childhood PBMs

i) General description

Our review identified ten existing childhood PBMs [in Table 1](#): 16-Dimensional (16D), 17-Dimensional (17D), Adolescent Health Utility Measurement (AHUM), Assessment of Quality of Life-6Dimension (AQoL-6D), Child Health Utility 9Dimension (CHU-9D), EQ-5D-Youth (EQ-5D-Y), Health Utilities Index Mark 2 (HUI2), Health Utilities Index Mark 3 (HUI3), Quality of Well Being-Self Administered (QWB-SA) and, Child Health-6 Dimension (CH-6D). The CH-6D was a childhood PBM recently developed in South Korea for children aged 7-12 years.²³ All ten measures were developed in high income countries, adapted predominantly from existing adult versions with little or no involvement of children at the development stage²⁴ with the exception of the CHU-9D.²⁵ [Table 1 provides an overview of these ten identified measures.](#)

Review of childhood PB HRQoL measures in sSA

ii) *Underpinning utility theory of identified measures (Table 2):*

Utility values were generated using [time trade-off \(TTO\)](#) for the AQoL-6D, AHUM and CH-6D; combining TTO with a visual analogue scale (VAS) for the EQ-5D-Y; a rating scale (RS) for 16D, 17D, and QWB-SA; and combining a RS and the standard gamble (SG) approach for the HUI2 and HUI3. The utilities were elicited using adult preferences - whether own or proxy – for the HUI2, HUI3, QWB-SA, 17D, CHU-9D, EQ-D-Y, AHUM and CH-6D. The AQoL-6D and 16D were the only two measures that elicited the preferences of adolescents. Through either statistical additive or multiplicative models, the utilities generated for a sample health states were then extrapolated for all health states of the measures. For the 16D, 17D, AHUM, QWB-SA, EQ-5D-Y, CHU-9D this was through an additive model and through a multiplicative model for the HUI2, HUI3 and AQoL-6D. [Table 2 presents an overview of utility valuations for each of these measures.](#)

iii) *Psychometric validation:*

Table 3 presents the psychometric appraisal of the 10 identified measures. The psychometric evidence was mixed. Development of a conceptual framework, content validity, and convergent validity were most frequently evaluated. Only six measures (HUI2, HUI3, 16D, 17D, CHU-9D and EQ-5D-Y) had at least one piece of adequate evidence for both reliability and validity. A detailed description of the psychometric properties can be found in Appendix S1.

Application and cross-cultural validation of generic childhood PBMs in sub-Saharan Africa

The second stage of our search aimed to identify the application and cross-cultural validation of generic childhood PBMs in sSA. We identified 40 unique articles; 30 articles were excluded after

1 Review of childhood PB HRQoL measures in sSA

2
3 screening the title leaving ten articles for full-text assessment. Four of these articles were excluded
4
5 because they pertained to adults. Of the remaining six articles, two articles reported on the
6
7 application (objective 2) of two childhood PBMs in sSA,^{26,27} but also provided evidence on aspects
8
9 of validation (Figure 2). The included studies were of good quality, meeting most of the criteria as
10
11 defined by the CASP guidelines (see Table 5).
12
13
14
15
16

17 ***Application of generic childhood PBMs in sSA (objective 2)***

18
19
20 The two studies reporting the use of childhood PBMs in sSA were undertaken in Uganda and South
21
22 Africa.^{26,27} The quality of both studies was good according to the criteria defined by the CASP
23
24 checklist with clearly focussed topics, methodology, analysis and interpretation of the results
25
26 (Table 4). These two studies used the HUI3 and the EQ-5D-Y, respectively (Table 5).
27
28
29
30
31
32

33 Sims-Williams and colleagues²⁶ used the HUI3 to evaluate the HRQoL of children with spina bifida
34
35 and that of their caregiver adult proxies (self-reported from their perspective). There was little
36
37 difference in HRQoL perception between children and their proxies (0.575 and 0.549, respectively,
38
39 Pearson correlation=0.848). This study also found that male sex, urinary incontinence, large family
40
41 size and hydrocephalus were factors that were associated with children's HRQoL.
42
43
44
45
46
47

48 Jelsma and Ramma²⁷ used the EQ-5D-Y to compare HRQoL amongst children (aged 7-12 years)
49
50 attending open schools (healthy, n=567) and those attending special schools (with functional
51
52 limitations, n=41); and to compare the responses to the measure provided by the children and their
53
54 parents in South Africa. The parents (proxies), without consulting their children, completed the
55
56
57
58
59
60

1 Review of childhood PB HRQoL measures in sSA
2

3 EQ-5D-Y proxy version whilst the children completed the self-report version of the EQ-5D-Y. The
4
5 parents' responses to the EQ-5D-Y proxy version required them to rate their child's HRQoL from
6
7 the perspective of the child. The study found no differences in perception of HRQoL between
8
9 children in open and special schools in this population. However, differences were observed
10
11 between HRQoL reported by children and their female parent proxies in open schools but not in
12
13 special schools.
14
15
16
17
18
19

20 ***Cross-cultural validation of generic childhood PBMs in sub-Saharan Africa (objective 3)*** 21

22
23 The HUI3 and the EQ-5D-Y were the two generic childhood PBM measures that attempted cross-
24
25 cultural validation in six studies in sSA settings (Table 6 summary and Table S2 detailed).²⁸ Five
26
27 of the six studies were for the validation of the EQ-5D-Y and this was in one country alone, South
28
29 Africa.
30
31
32
33
34
35

36 The HUI3
37

38 *i) Linguistic equivalence* 39

40
41
42 The HUI3 language translation involved forward translation only from the original English
43
44 language into several local languages spoken in Uganda.²⁶ Members of the study team translated
45
46 HUI3 into Luganda, Lugisu, Lugwere, Lunyole, Lusoga, Kiswahili, Ateso, Dhapadhadola, Samia
47
48 and Kubsabiny as they went along interviewing families that they saw. However, the authors didn't
49
50 specify whether the study members were qualified translators or that a backward translation was
51
52 done prior to using the translated questionnaire.
53
54
55
56
57
58
59
60

1 Review of childhood PB HRQoL measures in sSA
2
3
4
5

6 *ii) Conceptual equivalence*
7

8 Cognitive interviews were not conducted prior to using the translated HUI3 among children and
9 parents conversant in the translated language in Uganda. The study members translated the HUI3
10 into the local language as it was being administered to children with spina bifida and their proxies.
11
12
13
14
15

16
17
18
19 *iii) psychometric validation*
20

21 The validation involved comparison of the performance of the HUI3 tool and the VAS marked 0
22 to 10, and the scores of children and their proxies (completed from the child's perspective). There
23 was moderate correlation (0.488) between the HUI3 and VAS scores of children in open and special
24 schools, and even lower correlation (0.380) between those of children and their proxies.
25
26
27
28
29
30
31

32
33
34
35 The EQ-5D-Y
36

37
38 *i) Linguistic equivalence*
39

40 The EQ-5D-Y was not translated into any native South African language as the EQ-5D-Y English
41 UK version was used.
42
43
44
45

46
47
48
49 *ii) Conceptual equivalence*
50

51 Cognitive interviews to establish conceptual equivalence were done for the EQ-5D-Y in four
52 countries (Germany, Italy, Spain and Sweden), but not in South Africa during the development
53
54
55
56
57
58
59
60

1 Review of childhood PB HRQoL measures in sSA

2
3 stage.²⁸ The adaptation involved revising wording of the adult EQ-5D dimensions into a child-
4 friendly version followed by psychometric validation in five countries including South Africa.
5
6
7
8
9

10
11 *iii) Psychometric validation*

12
13
14 The psychometric validation study demonstrated that the EQ-5D-Y was feasible in South Africa
15 since only 2% of the children had inappropriate or missing responses.²⁹ Further, there was
16 convergent validity between the EQ-5D-Y and both the KIDSCREEN-10 and KIDSCREEN-27 on
17 similar dimensions (correlation coefficient $r = -0.17$ to -0.52); the EQ-5D-Y and self-reported
18 general health ($r = 0.25-0.27$); and with the EQ-5D-Y VAS (Pearson, $r = -0.56$ to 0.55). The
19 negative correlation between the EQ-5D-Y and KIDSCREEN is because these are scored in
20 opposite directions (higher score on EQ-5D-Y indicates better HRQoL health status whereas for
21 the KIDSCREEN higher scores mean worse healthHRQoL). Known-group validity was evident in
22 that the EQ-5D-Y was able to differentiate reported problems between groups with different health
23 conditions ($p < 0.05$).²⁹ The p-value refers to the statistical significance between those reporting
24 'fair' or 'poor' self-reported health and more problems on EQ-5D-Y dimensions.
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42

43 Scott et al.³⁰ found that the EQ-5D-Y demonstrated test-retest reliability across both the descriptive
44 component and the VAS scores (kappa, $k = 0.365$ to 0.653 and $ICC = 0.77$); convergent validity with
45 the PedsQL amongst the acutely ill children in the study ($p < 0.001$); and known-group validity
46 amongst groups with different health states both by composite score (derived from a model in order
47 to generate a single utility score for the descriptive component) and VAS scores (all differences
48 $p < 0.001$), but not on composite scores.
49
50
51
52
53
54
55
56
57
58
59
60

Review of childhood PB HRQoL measures in sSA

In another South African psychometric validation study of the EQ-5D-Y, Jelsma and Ramma²⁷ demonstrated no inter-rater reliability between children and their adult proxies on all dimensions ($r < 0.33$), and no known-group validity between children with and without disabilities on ‘doing usual activities’, ‘having pain or discomfort’, and ‘feeling worried, sad or unhappy’ dimensions and on the VAS scores, mean 88.4 and 87.9, respectively. However, there was evidence of known-group validity on the ‘mobility’ and ‘looking after myself’ dimensions ($p < 0.001$). The last of the EQ-5D-Y validation studies, by Jelsma,³¹ demonstrated that the EQ-5D-Y and the EQ-5D shouldn’t be used interchangeably since the former performs better in children and adolescent populations in this setting. The EQ-5D-Y had statistically significant fewer missing responses on both dimensions ($\chi^2 = 9.404$; d.f.=1; $p = 0.002$) and the VAS scale ($\chi^2 = 26.159$; d.f.=1; $p < 0.001$), respectively, than the adult EQ-5D.

Summary assessment of results

Our review found that in previous studies both the HUI and the EQ-5D-Y did meet some but not all of the cross-cultural validation criteria. The HUI was translated into a native sub-Saharan language but the translation did not involve backward translation nor cognitive interviews to establish conceptual equivalence. Additionally, some but not all of the psychometric properties were evaluated for the HUI3.

On the other hand, previous studies of the EQ-5D-Y have established conceptual equivalence, but have not translated the measure into a native language spoken in sSA. Similar to the HUI3, we did not find any single study that had evaluated all the psychometric properties of the EQ-5D-Y. When

1 Review of childhood PB HRQoL measures in sSA
2

3 all the findings from these individual EQ-5D-Y studies are considered together, it is evident that
4
5 the psychometric properties have been comprehensively evaluated for the EQ-5D-Y except for
6
7 internal consistency, which was not reported in any of the studies. In summary, we found that the
8
9 majority of psychometric properties were not reported by individual studies. However, of those
10
11 that did report, the majority did meet the psychometric criteria.
12
13
14
15
16
17
18

19 **Discussion**

20
21 This review aimed to identify the available childhood PBMs and to determine the extent to which
22
23 they have been used and/or cross-culturally validated in the sSA region. Ten PBM measures were
24
25 identified that were all originally developed in high-income countries and mostly adapted from
26
27 adult versions of the measures. The HUI3 and the EQ-5D-Y were the only two measures used in
28
29 child and adolescent populations in sSA. Our review also highlighted the lack of involvement of
30
31 children and adolescents themselves in the process of adapting these measures for use in sSA.
32
33
34
35
36
37
38

39 ***Existing childhood PBMs***

40
41 There is a need for greater involvement of children and adolescents themselves in the development
42
43 of conceptual frameworks underlying childhood PBMs.²⁴ The CHU-9D is the only measure
44
45 identified by our review whose dimensions were generated directly from children through
46
47 qualitative interviews and analysis (although we have no data for the recently developed CH-6D).²⁴
48
49 Even though it has been demonstrated that adult domains/dimensions can relate to children and
50
51 adolescents,²⁸ it is also well known that children have different perspectives and goals, implying
52
53 that dimension structures should ideally be developed directly from children themselves within
54
55
56
57
58
59
60

1 Review of childhood PB HRQoL measures in sSA

2
3 specific contexts.³² This is consistent with previous authors¹⁴ who have emphasised that
4
5 interpretation of HRQoL may differ across cultures.^{8,33}
6
7
8
9

10 ***PBMs in sub-Saharan Africa***

11
12
13 The use of generic childhood PBMs in sSA settings is sparse. In part, this may reflect a lack of
14
15 policy relevance in sSA settings; although funders have set standards for childhood outcomes as
16
17 evidenced by the [Sustainable Development Goals \(SDGs\)](#), these have primarily focussed on
18
19 reducing child mortality.² While this is important, there should be some effort to measure the
20
21 quality of lifeHRQoL gained through the SDG interventions using the existing preference-based
22
23 measures. It may also reflect the fact that the measurement of HRQoL in children has lagged
24
25 behind that of their adult counterparts due to the challenges of eliciting robust self-reported
26
27 information across the developmental stages of childhood. To be able to use PBMs effectively in
28
29 sSA requires more research that develops methods of eliciting reliable information from children
30
31 themselves via questionnaires and other modes of administration.
32
33
34
35
36
37
38
39

40 Our findings suggest that there is potential for the EQ-5D-Y to be used in sSA, but more work is
41
42 needed to investigate its conceptual equivalence in this setting. There has been no cross-cultural
43
44 validation in sSA of the ten identified childhood PBMs with the exception of the HUI3 and the
45
46 EQ-5D-Y. The cross-cultural validation of the HUI3 involved translation and cognitive interviews
47
48 with children and parents, and psychometric evaluation.²⁶ Three theoretical underpinnings for
49
50 adapting preference-based HRQoL measures have previously been described, depending on the
51
52 impact of culture: absolutist, universalist and relativism.³³ Some adaptations of existing measures
53
54
55
56
57
58
59
60

Review of childhood PB HRQoL measures in sSA

take an absolutist approach, which assumes culture does not have an impact. This means that ~~measures are only directly translated in the intended culture~~ without seeking to establishing a conceptual equivalence, measures are exported, after crude translation, into the local cultural milieu.^{7,14} However, it is well known that local culture has an impact on health and health-related concepts.⁸ It is important therefore that conceptual equivalence, which is part of the adaptation process, be established before using PBMs cross-culturally. Developers of the EQ-5D-Y themselves suggest that future studies should consider other aspects of HRQoL that are important to children and adolescents for inclusion in the measure.³⁴ Moreover, the measures have only been applied in a small selection of sSA countries, raising questions about their applicability across sSA as a whole. Apart from Uganda, where the HUI3 was applied, the use of the childhood PBMs was limited to South Africa alone, which is not representative of most economies in sSA as it is ranked as an upper middle-income country according to the World Bank.^{28,29} The use of the EQ-5D-Y therefore needs to be explored in other countries outside South Africa in order to assess generalisability across sSA settings.

Strengths and limitations

This is the first study to our knowledge that assessed the application and cross-cultural adaption of childhood PBMs in sSA. Our search generated a high volume of data, and a large volume of articles were excluded during the title screening of our review, which could have led to some relevant articles being excluded. Nevertheless, the findings related to our first study objective are largely consistent with the most recent review of existing generic childhood PBMs.¹⁰ Although we appraised the measures for their psychometric properties, this was limited to the original work by the developers. Despite the fact that psychometric evaluations for widely used measures, such as

1 Review of childhood PB HRQoL measures in sSA

2
3 the HUI2 and HUI3, are ongoing, for practical reasons we only included the original psychometric
4 evaluations of each measure.³⁵ Our searches were also limited to peer reviewed material, which
5 could have overlooked important evidence in the grey literature. In relation to this, we excluded all
6 studies other than those in the English language, which could have had the effect of excluding
7 potentialsome local studies. Finally, the quality appraisal of the studies applied in sSA settings was
8 performed using the CASP case-control checklist due to the lack of toolkits designed specifically
9 for cross-sectional studies.
10
11
12
13
14
15
16
17
18
19
20
21
22
23

24 **Conclusion**

25
26 The findings of this review show that application of generic childhood PBMs in sSA settings has
27 hitherto been limited to the HUI3 and EQ-5D-Y-3L. Most adaptations of existing measures take
28 an absolutist approach, which assumes that measures can be used across cultures. However, there
29 is a need to undertake cross-cultural linguistic and conceptual equivalence, testing and
30 psychometric validation across a range of sSA cultural contexts. Unless cross-cultural validation
31 of PBMs is robust when applied in the sSA region, the application and interpretation of economic
32 evaluations reliant on preference-based outcome measures will fail to reflect local relevance in this
33 part of the world.
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1 Review of childhood PB HRQoL measures in sSA
2

3 **Additional files:**
4

5 Table S1: Detailed psychometric properties as evaluated by developers of the identified generic
6 childhood PBMs
7
8
9

10
11
12 Table S2: Detailed cross-cultural validation of generic childhood preference-based HRQoL
13 measures in sSA (n=6)
14
15
16
17
18

19 **Declarations**
20

21 **Competing interests**
22

23
24 The authors declare that they have no competing interests.
25
26
27

28 **Authors' contributions**
29

30 LGN, LN, SP and HM conceived the concept; LGN, KK and LS independently assessed the titles
31 and abstracts; LGN, SS, HM, SP and LN drafted the original manuscript; LGN, LNM, HM, SS, SP
32 and LN reviewed the manuscript. All authors read and approved the final manuscript
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Review of childhood PB HRQoL measures in sSA

REFERENCES

1. WHO G, Switzerland. UNESCO SCIENCE REPORT. last accessed on 11th October 2017.
2. WHO G, Switzerland. Millenium Development Goals and Beyond 2015. Accessed 23rd November 2017, 2017.
3. WHO G, Switzerland. Sustainable development goals. <http://www.un.org/sustainabledevelopment/sustainable-development-goals/>. last assessed 15th September 2017.
4. WHOQOL Group. Development of the World Health Organization WHOQOL-BREF quality of life assessment. *Psychological Medicine*. 1998;28(3):551-558.
5. Brazier J, Deverill M, Green C. A review of the use of health status measures in economic evaluation. *Journal of Health Services Research and Policy*. 1999;4(3):174-184.
6. Wallander J, Schmitt M, Koot H. Quality of Life Measurement in Children and Adolescents: Issues, Instruments, and Applications. *Journal of Clinical Psychology*. 2001;57(4):571-585.
7. Bowden A, Fox-Rushby JA. A systematic and critical review of the process of translation and adaptation of generic health-related quality of life measures in Africa, Asia, Eastern Europe, the Middle East, South America. *Soc Sci Med*. 2003;57(7):1289-1306.
8. Epstein J, Santo RM, Guillemin F. A review of guidelines for cross-cultural adaptation of questionnaires could not bring out a consensus. *J Clin Epidemiol*. 2015;68(4):435-441.
9. Janssens A, Thompson Coon J, Rogers M, et al. A systematic review of generic multidimensional patient-reported outcome measures for children, part I: descriptive characteristics. *Value Health*. 2015;18(2):315-333.

Review of childhood PB HRQoL measures in sSA

10. Chen G, Ratcliffe J. A Review of the Development and Application of Generic Multi-Attribute Utility Instruments for Paediatric Populations. *Pharmacoeconomics*. 2015;33(10):1013-1028.
11. Thorrington D, Eames K. Measuring Health Utilities in Children and Adolescents: A Systematic Review of the Literature. *PLoS One*. 2015;10(8):e0135672.
12. Colbourn T, Masache G, Skordis-Worrall J. Development, reliability and validity of the Chichewa WHOQOL-BREF in adults in Lilongwe, Malawi. *BMC Res Notes*. 2012;5:346.
13. Westmoreland K, Reeve BB, Amuquandoh A, et al. Translation, psychometric validation, and baseline results of the Patient-Reported Outcomes Measurement Information System (PROMIS) pediatric measures to assess health-related quality of life of patients with pediatric lymphoma in Malawi. *Pediatr Blood Cancer*. 2018:e27353.
14. Paltzer J, Barker E, Witt WP. Measuring the health-related quality of life (HRQoL) of young children in resource-limited settings: a review of existing measures. *Qual Life Res*. 2013;22(6):1177-1187.
15. Moher D, Liberati A, Tetzlaff J, Altman DG, Group P. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *J Clin Epidemiol*. 2009;62(10):1006-1012.
16. Mokkink LB, Terwee CB, Patrick DL, et al. The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. *Qual Life Res*. 2010;19(4):539-549.

Review of childhood PB HRQoL measures in sSA

17. Canaway AG, Frew EJ. Measuring preference-based quality of life in children aged 6-7 years: a comparison of the performance of the CHU-9D and EQ-5D-Y--the WAVES pilot study. *Qual Life Res.* 2013;22(1):173-183.
18. (FDA) UDoHaHSFaDA. Guidance for industry: patient-reported outcome measures: use in medical product development to support labeling claims. Rockville, MD: FDA; 2009.
19. Trust SACotMO. Assessing health status and quality-of-life instruments: attributes and review criteria. *Qual Life Res.* 2002;11(3):193-205.
20. Reeve BB, Wyrwich KW, Wu AW, et al. ISOQOL recommends minimum standards for patient-reported outcome measures used in patient-centered outcomes and comparative effectiveness research. *Qual Life Res.* 2013;22(8):1889-1905.
21. Terwee CB, Bot SD, de Boer MR, et al. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol.* 2007;60(1):34-42.
22. Programme CAS. CASP Systematic Review Checklist. 2018; <https://casp-uk.net/casp-tools-checklists/>. Accessed 14th January, 2019.
23. Kang E. Validity of Child Health-6 Dimension(CH-6D) for Adolescents. *Value in Health.* 2016;19(7).
24. Stevens KJ. How Well Do the Generic Multi-attribute Utility Instruments Incorporate Patient and Public Views Into Their Descriptive Systems? *Patient.* 2016;9(1):5-13.
25. Stevens K. Developing a descriptive system for a new preference-based measure of health-related quality of life for children. *Qual Life Res.* 2009;18(8):1105-1113.
26. Sims-Williams HJ, Sims-Williams HP, Mbabazi Kabachelor E, Warf BC. Quality of life among children with spina bifida in Uganda. *Arch Dis Child.* 2017;102(11):1057-1061.

Review of childhood PB HRQoL measures in sSA

27. Jelsma J, Ramma L. How do children at special schools and their parents perceive their HRQoL compared to children at open schools? *Health and Qual Life Outcomes*. 2010;8(1):72.
28. Wille N, Badia X, Bonsel G, et al. Development of the EQ-5D-Y: a child-friendly version of the EQ-5D. *Qual Life Res*. 2010;19(6):875-886.
29. Ravens-Sieberer U, Wille N, Badia X, et al. Feasibility, reliability, and validity of the EQ-5D-Y: results from a multinational study. *Qual Life Res*. 2010;19(6):887-897.
30. Scott D, Ferguson GD, Jelsma J. The use of the EQ-5D-Y health related quality of life outcome measure in children in the Western Cape, South Africa: psychometric properties, feasibility and usefulness - a longitudinal, analytical study. *Health Qual Life Outcomes*. 2017;15(1):12.
31. Jelsma J. A comparison of the performance of the EQ-5D and the EQ-5D-Y health-related quality of life instruments in South African children. *Int J Rehabil Res*. 2010;33(2):172-177.
32. 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 Econ*. 2016;25(4):486-496.
33. Herdman M, Fox-Rushby J, Badia X. 'Equivalence' and the translation and adaptation of health-related quality of life questionnaires. *Quality of Life Research*. 1997;6:237-247.
34. Connolly MA, Johnson JA. Measuring quality of life in paediatric patients. *Pharmacoeconomics*. 1999;16(6):605-625.
35. Feeny D. The Health Utilities Index: A tool for assessing health benefits. *PRO Newsletter*. 2005;34.

1 Review of childhood PB HRQoL measures in sSA

2
3 ABSTRACT

4
5 *Objectives*

6
7 Consideration of health status in children and adolescents now includes broader concepts such as
8
9 health-related quality-of-life (HRQoL). Globally there is a need for relevant preference-based
10
11 HRQoL measures (PBMs) for use in children and adolescents, yet measurement of HRQoL in these
12
13 groups presents particular challenges. This paper systematically reviews the available generic
14
15 childhood PBMs, and their application and cross-cultural validation in sub-Saharan African (sSA).
16
17
18
19

20
21 *Methods*

22
23 A systematic review of published literature from 01/01/1990 to 02/08/2017 was conducted using
24
25 MEDLINE (through OVID Sp), EMBASE (OVID Sp), EconLit (EBSCOhost), PsycINFO, Web
26
27 of Science and PubMed.
28
29
30
31

32
33 *Results*

34
35 220 full-text articles were included in a qualitative synthesis. Ten generic childhood PBMs were
36
37 identified, of which nine were adapted from adult versions and only one developed specifically for
38
39 children. None of the measures were originally developed in sSA or other resource-constrained
40
41 settings. The HUI3 and the EQ-5D-Y-3L were the only measures that had been applied in sSA
42
43 settings. Further, the HUI3 and the EQ-5D-Y-3L were the only generic childhood PBM that
44
45 attempted to establish cross-cultural validation in sSA. Five of the six of these validation studies
46
47 were conducted using the EQ-5D-Y-3L in a single country, South Africa.
48
49
50
51

52
53 *Conclusions*

1
2
3 Review of childhood PB HRQoL measures in sSA
4

5 The findings show that application of generic childhood PBMs in sSA settings has hitherto been
6 limited to the HUI3 and EQ-5D-Y-3L. Most adaptations of existing measures take an absolutist
7 approach, which assumes that measures can be used across cultures. However, there is also need
8 to ensure linguistic and conceptual equivalence and undertake validation across a range of sSA
9 cultural contexts.
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Review of childhood PB HRQoL measures in sSA

Highlights:

i. What is already known about the topic?

Since the 1990s, there has been growing interest in developing and using child and adolescent (childhood) preference-based health-related quality of life measures (PBMs). The increased prominence of childhood PBMs has been driven by their use in measuring health-related quality-of-life (HRQoL) as part of economic evaluations, thereby informing health care policies.

ii. What does the paper add to existing knowledge?

While this paper examines existing childhood PBMs and summarises their psychometric properties, it focuses on the validation and application of these measures in the sub-Saharan African (sSA) region. Development of existing childhood PBMs have been restricted to high-income countries, and none has been developed in low-income countries. Validation and application of childhood PBMs in the sSA region has been limited to the HUI3 and EQ-5D-Y-3L in two countries.

iii. What insights does the paper provide for informing health care-related decision making?

Findings from health technology assessments (HTA) and economic evaluations are increasingly used to inform health policy in sSA. These studies should incorporate local preferences for health outcomes. Preference-based values for health outcomes are commonly generated using tools that are adapted to local contexts. If the process of adapting these measures in sSA contexts is not robust, the findings from subsequent HTAs and economic evaluations may not be applicable.

1 Review of childhood PB HRQoL measures in sSA
2

3 **Introduction:**
4

5 In sub-Saharan Africa (sSA) i.e. the region below the Sahara desert, there has been a remarkable
6 decline in mortality rates in infants and under 5-year olds since 2000,¹ partly driven by policy
7 responses to the Millennium Development Goals (MDGs).² In 2015, the United Nations set 17
8 sustainable development goals as a replacement for MDGs for nations to achieve by 2030. Goal
9 number three focuses on ‘good health and wellbeing’, specifically targeting child and maternal
10 mortality and chronic disease reduction.³ As childhood mortality continues to fall, and increasing
11 numbers of children survive into adolescence and beyond, future interventions will increasingly
12 focus on improving wider outcomes including children’s health-related quality of life (HRQoL).⁴
13
14
15
16
17
18
19
20
21
22
23
24
25

26 Preference-based HRQoL measures (PBMs) are designed to “assess patient preferences across
27 broad areas including symptoms, physical functioning, work and social activities, and mental well-
28 being”.⁵ PBMs have increased in prominence over recent years because of their use in economic
29 evaluations, which often inform policy.⁶ This can only be meaningful if PBMs are relevant, valid
30 and robust. However, most generic PBMs have been developed in the English language⁷ and, as
31 the development of new measures is time and resource intensive, use of PBMs in other contexts
32 has mostly relied on translating and/or adapting existing measures, with an implicit assumption
33 that adaptation produces an equivalent measure.⁸ Adaptation of an existing measure requires an
34 iterative process of translation and back translation with a qualitative review to establish both
35 linguistic and conceptual equivalence.⁽⁷⁾ In addition, the adapted measure needs to be re-validated
36 using standard psychometric methods to re-establish validity and reliability for the new form of the
37 measure in the new context.
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1 Review of childhood PB HRQoL measures in sSA

2
3 Several reviews have assessed the development and application of childhood PBMs and non-PBMs
4 (restricted to ages 6-17 years),⁹⁻¹¹ their cross-cultural validation,^{12,13} and availability of measures
5 for young children (<8 years) in alternative settings.¹⁴ However, to our knowledge, no review has
6 specifically studied the use or validation of childhood PBMs across sSA.
7
8
9
10

11
12 We carried out a systematic review of generic childhood PBMs, particularly focussing on their use
13 and cross-cultural validation in sSA settings. The objectives of the review were threefold: (1) to
14 identify existing generic childhood PBMs and summarise their psychometric quality; (2) to explore
15 which of these generic childhood PBMs have been applied in sSA; and (3) to assess the extent of
16 cross-cultural validation of generic childhood PBMs for application in sSA.
17
18
19
20
21
22
23
24
25
26
27

28 **Methods**

29
30 We conducted the literature searches using a two-stage process compatible with the preferred
31 reporting items for systematic reviews and meta-analyses (PRISMA) guidelines.¹⁵ For the first
32 stage, we searched the literature to identify existing childhood PBMs (objective 1). In the second
33 stage (to address objectives two and three), we refined our search strategy to identify all published
34 research studies that applied or cross-culturally validated any one of these PBMs in the sSA region.
35
36 Cross-cultural validation in our search criteria was defined as i) linguistic equivalence (achieved
37 through translation), ii) conceptual equivalence (achieved through focus group discussion or
38 cognitive interviews) and iii) psychometric validation. The peer reviewed literature published
39 between 1st January 1990 and 2nd August 2017 was examined for both stages of the review.
40
41
42
43
44
45
46
47
48
49
50
51
52

53 ***Search strategy***

1 Review of childhood PB HRQoL measures in sSA
2

3 In stage one, we carried out systematic searches using an intersection of four sub-groups of search
4 terms: 1) HRQoL terms; 2) childhood terms; 3) instrument terms; and 4) psychometric terms. The
5 psychometric terms were included in order to identify validated measures only. A full list of search
6 terms is provided in Appendix 1.1. The search was performed in six databases via Ovid: PubMed,
7 EMBASE, Web of Science, EconLit, PsycINFO and MEDLINE, as per COSMIN
8 recommendations.¹⁶
9
10
11
12
13
14
15
16
17
18

19 In stage two, we searched through the same databases using search terms for the identified generic
20 childhood PBMs and combined them with sSA terms to identify their application and/or cross-
21 cultural validation in sSA settings. The full list of identified measures and sSA search terms used
22 to address the second and third study objectives are provided in Appendix 1.2.
23
24
25
26
27
28
29

30 ***Inclusion and exclusion criteria***

31 To address the first study objective, studies reporting development of measures were included if:
32
33 (1) they reported the development of at least one childhood PBM (this was limited to the original
34 work by the developers only); 2) the target age(s) for completion of the measure included children
35 between 6-17 years; and (3) were published in the English language. Studies were excluded if they
36 reported measures designed for adults (≥ 18 years) or young children (< 6 years) as the review
37 focussed on measures with the potential for self-completion by children.¹⁷ Additionally, all studies
38 that reported on psychometric properties of the identified childhood PBMs were restricted to work
39 by the primary developers.
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1 Review of childhood PB HRQoL measures in sSA

2
3 To address the second and third study objectives, studies were included if they had applied and/or
4 cross-culturally validated any measure identified at stage one of the review in a sSA setting. Studies
5 using generic childhood PBMs that were not carried out either in part or whole in a sSA setting
6 were excluded.
7
8
9
10
11
12
13

14 ***Procedures***

15
16 Three reviewers (LGN, KK and LS) independently assessed the titles, and subsequent abstracts, to
17 select potential full-text articles for screening. Study selection was agreed upon at each stage (title,
18 title and abstract, full article) by the three reviewers with disagreements resolved via consensus.
19
20
21
22
23
24
25

26 ***Data extraction and analysis***

27 To meet **objective 1**, for each identified generic childhood PBM the following was extracted:
28
29
30
31
32

33 *(i) General description of the measure*

34 The descriptive system consisting of the classification system; where, how and who developed the
35 classification system; target population; number of dimensions; response levels; and time taken to
36 respond.
37
38
39
40
41
42
43

44 *(ii) Underpinning utility theory*

45 This included the number of health of states generated by the classification system, the valuation
46 technique used, model used to extrapolate the utilities to all possible health states, range of utilities
47 for each measure and when and how these were generated.
48
49
50
51
52
53
54
55

56 *(iii) Psychometric validation*

Review of childhood PB HRQoL measures in sSA

The quality of each identified childhood PBM was appraised against standard psychometric criteria¹⁸⁻²¹ as outlined below.

Conceptual framework including a detailed description of the conceptual basis of the construct measured, is usually based on qualitative work with the target population.²⁰

Cross-cultural language and conceptual equivalence/adaptation is the extent to which a scale shows linguistic and conceptual equivalence, usually assessed through qualitative methodology, including translation, cognitive interviews or focus group discussions with the target population.⁸

Reliability, the extent to which the instrument is free from random error, was assessed in terms of *internal consistency*, defined using Cronbach's $\alpha \geq 0.70$ for group level comparisons and ≥ 0.90 for individual comparisons;¹⁹ and *test re-test* assessed by an Intraclass-Correlation Coefficient (ICC) ≥ 0.70 or Kappa $\kappa \geq 0.70$ for nominal scales.^{18,19,21}

Validity, the extent to which a scale measures what it purports to measure, was assessed using: *Content validity* which assesses the scale's conceptual breadth, including all aspects of the construct being measured;^{18,20} *Criterion validity*, which assesses how the scale correlates with a gold standard measure of the same construct administered at the same time (*concurrent*) or at some point in the future (*predictive*), defined using correlation ≥ 0.70 ;²¹ *Convergent validity*, the extent to which the scale is positively related to other scales that purport to measure the same thing; *Discriminant validity*, the extent to which the scale is not related to other scales that purport to measure different things, assessed using correlation coefficients; *Known groups validity*, which assesses the extent to which scores on the scale are different between two groups that are known to be different on another clinically defined variable;²¹ and *Responsiveness*, the ability of the scale to detect clinically important change over time, which was assessed using effect size statistics, score changes or standardized response means.¹⁸

1 Review of childhood PB HRQoL measures in sSA

2
3 *Interpretability of scores* refers to the degree to which one can easily assign meanings to an
4 instrument's quantitative scores, assessed in several ways including normative data.¹⁸⁻²¹
5

6
7 *Practicality* is the burden on those completing and/or those administering the measure and includes
8 completion time, and completion and response rates.²¹
9

10
11
12
13
14 In our review, all psychometric properties described above were classified as (+) if they met the
15 criteria; (-) if not; and (NR) if not reported. We assessed the psychometric properties as a whole
16 rather than using a single statistic as evidence of psychometric quality.
17
18
19

20
21
22
23
24 To achieve **objective 2**, all studies included in this second stage of the review were assessed for
25 their quality using the Critical Appraisal Skills Programme (CASP) toolkit for case-control
26 studies.²² The CASP tool consists of ten questions that address issues such as study aim and design,
27 data collection and analysis, and ethical issues raised by the published study.
28
29
30

31
32
33
34
35 In order to achieve **objective 3**, each identified generic childhood PBM was assessed according to
36 whether there had been: i) cross-cultural translation into any native language spoken in sSA; ii)
37 any conceptual equivalence through focus group discussion or cognitive interviews on word
38 meanings and changes to concepts to fit what would be locally equivalent into a sSA setting and
39 iii) psychometric validation. Cross-cultural validation was classified as (+) if it was appraised and
40 performed; (-) if appraised but not performed; and (NR) if not reported.
41
42
43
44
45
46
47
48
49
50

51 This review was not part of any clinical trial and was therefore not registered as would have been
52 required otherwise.
53
54
55
56
57
58
59
60

Review of childhood PB HRQoL measures in sSA

Results

Identifying childhood preference-based HRQoL measures (objective 1)

The first stage of our search identified 59,342 unique articles (Figure 1). 58,366 articles were excluded after screening the title and a further 734 articles excluded after screening the abstract. In addition to the remaining 242 articles, we identified 91 articles through snowball searching. We therefore reviewed the full texts of 333 articles for eligibility. 113 of these articles were excluded as they did not meet eligibility criteria; a large number (n=51) of these articles were excluded as they did not report use of a generic childhood PBM. In total, we identified 220 articles that reported development of at least one generic childhood PBM, from which we identified ten unique measures (Figure 1).

Identified childhood PBMs

i) General description

Our review identified ten existing childhood PBMs in Table 1: 16-Dimensional (16D), 17-Dimensional (17D), Adolescent Health Utility Measurement (AHUM), Assessment of Quality of Life-6Dimension (AQoL-6D), Child Health Utility 9Dimension (CHU-9D), EQ-5D-Youth (EQ-5D-Y), Health Utilities Index Mark 2 (HUI2), Health Utilities Index Mark 3 (HUI3), Quality of Well Being-Self Administered (QWB-SA) and, Child Health-6 Dimension (CH-6D). The CH-6D was a childhood PBM recently developed in South Korea for children aged 7-12 years.²³ All ten measures were developed in high income countries, adapted predominantly from existing adult versions with little or no involvement of children at the development stage²⁴ with the exception of the CHU-9D.²⁵

1 Review of childhood PB HRQoL measures in sSA

2
3 *ii) Underpinning utility theory of identified measures (Table 2):*

4
5
6 Utility values were generated using time trade-off (TTO) for the AQoL-6D, AHUM and CH-6D;
7
8 combining TTO with a visual analogue scale (VAS) for the EQ-5D-Y; a rating scale (RS) for 16D,
9
10 17D, and QWB-SA; and combining a RS and the standard gamble (SG) approach for the HUI2 and
11
12 HUI3. The utilities were elicited using adult preferences - whether own or proxy – for the HUI2,
13
14 HUI3, QWB-SA, 17D, CHU-9D, EQ-D-Y, AHUM and CH-6D. The AQoL-6D and 16D were the
15
16 only two measures that elicited the preferences of adolescents. Through either statistical additive
17
18 or multiplicative models, the utilities generated for a sample health states were then extrapolated
19
20 for all health states of the measures. For the 16D, 17D, AHUM, QWB-SA, EQ-5D-Y, CHU-9D
21
22 this was through an additive model and through a multiplicative model for the HUI2, HUI3 and
23
24 AQoL-6D.
25
26
27
28
29
30
31

32 *iii) Psychometric validation:*

33
34
35 Table 3 presents the psychometric appraisal of the 10 identified measures. The psychometric
36
37 evidence was mixed. Development of a conceptual framework, content validity, and convergent
38
39 validity were most frequently evaluated. Only six measures (HUI2, HUI3, 16D, 17D, CHU-9D and
40
41 EQ-5D-Y) had at least one piece of adequate evidence for both reliability and validity. A detailed
42
43 description of the psychometric properties can be found in Appendix S1.
44
45
46
47
48
49

50 ***Application and cross-cultural validation of generic childhood PBMs in sub-Saharan Africa***

51
52 The second stage of our search aimed to identify the application and cross-cultural validation of
53
54 generic childhood PBMs in sSA. We identified 40 unique articles; 30 articles were excluded after
55
56
57
58
59
60

1 Review of childhood PB HRQoL measures in sSA

2
3 screening the title leaving ten articles for full-text assessment. Four of these articles were excluded
4
5 because they pertained to adults. Of the remaining six articles, two articles reported on the
6
7 application (objective 2) of two childhood PBMs in sSA,^{26,27} but also provided evidence on aspects
8
9 of validation (Figure 2). The included studies were of good quality, meeting most of the criteria as
10
11 defined by the CASP guidelines (see Table 5).
12
13
14
15
16

17 ***Application of generic childhood PBMs in sSA (objective 2)***

18
19
20 The two studies reporting the use of childhood PBMs in sSA were undertaken in Uganda and South
21
22 Africa.^{26,27} The quality of both studies was good according to the criteria defined by the CASP
23
24 checklist with clearly focussed topics, methodology, analysis and interpretation of the results
25
26 (Table 4). These two studies used the HUI3 and the EQ-5D-Y, respectively (Table 5).
27
28
29
30
31
32

33 Sims-Williams and colleagues²⁶ used the HUI3 to evaluate the HRQoL of children with spina bifida
34
35 and that of their caregiver adult proxies (self-reported from their perspective). There was little
36
37 difference in HRQoL perception between children and their proxies (0.575 and 0.549, respectively,
38
39 Pearson correlation=0.848). This study also found that male sex, urinary incontinence, large family
40
41 size and hydrocephalus were factors that were associated with children's HRQoL.
42
43
44
45
46
47

48 Jelsma and Ramma²⁷ used the EQ-5D-Y to compare HRQoL amongst children (aged 7-12 years)
49
50 attending open schools (healthy, n=567) and those attending special schools (with functional
51
52 limitations, n=41); and to compare the responses to the measure provided by the children and their
53
54 parents in South Africa. The parents (proxies), without consulting their children, completed the
55
56
57
58
59
60

1 Review of childhood PB HRQoL measures in sSA

2
3 EQ-5D-Y proxy version whilst the children completed the self-report version of the EQ-5D-Y. The
4
5 parents' responses to the EQ-5D-Y proxy version required them to rate their child's HRQoL from
6
7 the perspective of the child. The study found no differences in perception of HRQoL between
8
9 children in open and special schools in this population. However, differences were observed
10
11 between HRQoL reported by children and their female parent proxies in open schools but not in
12
13 special schools.
14
15
16
17
18
19

20 ***Cross-cultural validation of generic childhood PBMs in sub-Saharan Africa (objective 3)***

21
22
23 The HUI3 and the EQ-5D-Y were the two generic childhood PBM measures that attempted cross-
24
25 cultural validation in six studies in sSA settings (Table 6 summary and Table S2 detailed).²⁸ Five
26
27 of the six studies were for the validation of the EQ-5D-Y and this was in one country alone, South
28
29 Africa.
30
31
32
33
34
35

36 The HUI3

37 38 39 *i) Linguistic equivalence*

40
41
42 The HUI3 language translation involved forward translation only from the original English
43
44 language into several local languages spoken in Uganda.²⁶ Members of the study team translated
45
46 HUI3 into Luganda, Lugisu, Lugwere, Lunyole, Lusoga, Kiswahili, Ateso, Dhapadhadola, Samia
47
48 and Kubsabiny as they went along interviewing families that they saw. However, the authors didn't
49
50 specify whether the study members were qualified translators or that a backward translation was
51
52 done prior to using the translated questionnaire.
53
54
55
56
57
58
59
60

1 Review of childhood PB HRQoL measures in sSA
2
3
4
5

6 *ii) Conceptual equivalence*
7

8 Cognitive interviews were not conducted prior to using the translated HUI3 among children and
9 parents conversant in the translated language in Uganda. The study members translated the HUI3
10 into the local language as it was being administered to children with spina bifida and their proxies.
11
12
13
14
15

16
17
18
19 *iii) psychometric validation*
20

21 The validation involved comparison of the performance of the HUI3 tool and the VAS marked 0
22 to 10, and the scores of children and their proxies (completed from the child's perspective). There
23 was moderate correlation (0.488) between the HUI3 and VAS scores of children in open and special
24 schools, and even lower correlation (0.380) between those of children and their proxies.
25
26
27
28
29
30
31

32
33
34
35 The EQ-5D-Y
36

37
38 *i) Linguistic equivalence*
39

40 The EQ-5D-Y was not translated into any native South African language as the EQ-5D-Y English
41 UK version was used.
42
43
44
45

46
47
48
49 *ii) Conceptual equivalence*
50

51 Cognitive interviews to establish conceptual equivalence were done for the EQ-5D-Y in four
52 countries (Germany, Italy, Spain and Sweden), but not in South Africa during the development
53
54
55
56
57
58
59
60

1 Review of childhood PB HRQoL measures in sSA

2
3 stage.²⁸ The adaptation involved revising wording of the adult EQ-5D dimensions into a child-
4 friendly version followed by psychometric validation in five countries including South Africa.
5
6
7
8
9

10
11 *iii) Psychometric validation*
12

13
14 The psychometric validation study demonstrated that the EQ-5D-Y was feasible in South Africa
15 since only 2% of the children had inappropriate or missing responses.²⁹ Further, there was
16 convergent validity between the EQ-5D-Y and both the KIDSCREEN-10 and KIDSCREEN-27 on
17 similar dimensions (correlation coefficient $r = -0.17$ to -0.52); the EQ-5D-Y and self-reported
18 general health ($r = 0.25-0.27$); and with the EQ-5D-Y VAS (Pearson, $r = -0.56$ to 0.55). The
19 negative correlation between the EQ-5D-Y and KIDSCREEN is because these are scored in
20 opposite directions (higher score on EQ-5D-Y indicates better HRQoL whereas for the
21 KIDSCREEN higher scores mean worse HRQoL). Known-group validity was evident in that the
22 EQ-5D-Y was able to differen reported problems between with ($p < 0.05$).²⁹ The p-value refers to
23 the statistical significance between those reporting 'fair' or 'poor' self-reported health and more
24 problems on EQ-5D-Y dimensions.
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42

43 Scott et al.³⁰ found that the EQ-5D-Y demonstrated test-retest reliability across both the descriptive
44 component and the VAS scores (kappa, $k = 0.365$ to 0.653 and ICC = 0.77); convergent validity with
45 the PedsQL amongst the acutely ill children in the study ($p < 0.001$); and known-group validity
46 amongst groups with different health states both by composite score (derived from a model in order
47 to generate a single utility score for the descriptive component) and VAS scores (all differences
48 $p < 0.001$), but not on composite scores.
49
50
51
52
53
54
55
56
57
58
59
60

Review of childhood PB HRQoL measures in sSA

In another South African psychometric validation study of the EQ-5D-Y, Jelsma and Ramma²⁷ demonstrated no inter-rater reliability between children and their adult proxies on all dimensions ($r < 0.33$), and no known-group validity between children with and without disabilities on ‘doing usual activities’, ‘having pain or discomfort’, and ‘feeling worried, sad or unhappy’ dimensions and on the VAS scores, mean 88.4 and 87.9, respectively. However, there was evidence of known-group validity on the ‘mobility’ and ‘looking after myself’ dimensions ($p < 0.001$). The last of the EQ-5D-Y validation studies, by Jelsma,³¹ demonstrated that the EQ-5D-Y and the EQ-5D shouldn’t be used interchangeably since the former performs better in children and adolescent populations in this setting. The EQ-5D-Y had statistically significant fewer missing responses on both dimensions ($\chi^2 = 9.404$; d.f.=1; $p = 0.002$) and the VAS scale ($\chi^2 = 26.159$; d.f.=1; $p < 0.001$), respectively, than the adult EQ-5D.

Summary assessment of results

Our review found that in previous studies both the HUI and the EQ-5D-Y did meet some but not all of the cross-cultural validation criteria. The HUI was translated into a native sub-Saharan language but the translation did not involve backward translation nor cognitive interviews to establish conceptual equivalence. Additionally, some but not all of the psychometric properties were evaluated for the HUI3.

On the other hand, previous studies of the EQ-5D-Y have established conceptual equivalence, but have not translated the measure into a native language spoken in sSA. Similar to the HUI3, we did not find any single study that had evaluated all the psychometric properties of the EQ-5D-Y. When

1 Review of childhood PB HRQoL measures in sSA
2

3 all the findings from these individual EQ-5D-Y studies are considered together, it is evident that
4
5 the psychometric properties have been comprehensively evaluated for the EQ-5D-Y except for
6
7 internal consistency, which was not reported in any of the studies. In summary, we found that the
8
9 majority of psychometric properties were not reported by individual studies. However, of those
10
11 that did report, the majority did meet the psychometric criteria.
12
13
14
15
16
17
18

19 **Discussion**

20
21 This review aimed to identify the available childhood PBMs and to determine the extent to which
22
23 they have been used and/or cross-culturally validated in the sSA region. Ten PBM measures were
24
25 identified that were all originally developed in high-income countries and mostly adapted from
26
27 adult versions of the measures. The HUI3 and the EQ-5D-Y were the only two measures used in
28
29 child and adolescent populations in sSA. Our review also highlighted the lack of involvement of
30
31 children and adolescents themselves in the process of adapting these measures for use in sSA.
32
33
34
35
36
37

38 ***Existing childhood PBMs***

39
40
41 There is a need for greater involvement of children and adolescents themselves in the development
42
43 of conceptual frameworks underlying childhood PBMs.²⁴ The CHU-9D is the only measure
44
45 identified by our review whose dimensions were generated directly from children through
46
47 qualitative interviews and analysis (although we have no data for the recently developed CH-6D).²⁴
48
49 Even though it has been demonstrated that adult domains/dimensions can relate to children and
50
51 adolescents,²⁸ it is also well known that children have different perspectives and goals, implying
52
53 that dimension structures should ideally be developed directly from children themselves within
54
55
56
57
58
59
60

1 Review of childhood PB HRQoL measures in sSA

2
3 specific contexts.³² This is consistent with previous authors¹⁴ who have emphasised that
4
5 interpretation of HRQoL may differ across cultures.^{8,33}
6
7
8
9

10 ***PBMs in sub-Saharan Africa***

11
12
13 The use of generic childhood PBMs in sSA settings is sparse. In part, this may reflect a lack of
14
15 policy relevance in sSA settings; although funders have set standards for childhood outcomes as
16
17 evidenced by the Sustainable Development Goals (SDGs), these have primarily focussed on
18
19 reducing child mortality.² While this is important, there should be some effort to measure the
20
21 HRQoL gained through the SDG interventions using the existing preference-based measures. It
22
23 may also reflect the fact that the measurement of HRQoL in children has lagged behind that of their
24
25 adult counterparts due to the challenges of eliciting robust self-reported information across the
26
27 developmental stages of childhood. To be able to use PBMs effectively in sSA requires more
28
29 research that develops methods of eliciting reliable information from children themselves via
30
31 questionnaires and other modes of administration.
32
33
34
35
36
37
38
39

40 Our findings suggest that there is potential for the EQ-5D-Y to be used in sSA, but more work is
41
42 needed to investigate its conceptual equivalence in this setting. There has been no cross-cultural
43
44 validation in sSA of the ten identified childhood PBMs with the exception of the HUI3 and the
45
46 EQ-5D-Y. The cross-cultural validation of the HUI3 involved translation and cognitive interviews
47
48 with children and parents, and psychometric evaluation.²⁶ Three theoretical underpinnings for
49
50 adapting preference-based HRQoL measures have previously been described, depending on the
51
52 impact of culture: absolutist, universalist and relativism.³³ Some adaptations of existing measures
53
54
55
56
57
58
59
60

Review of childhood PB HRQoL measures in sSA

1
2
3 take an absolutist approach, which assumes culture does not have an impact. This means that,
4
5 without seeking to establish a conceptual equivalence, measures are exported, after crude
6
7 translation, into the local cultural milieu.^{7,14} However, it is well known that local culture has an
8
9 impact on health and health-related concepts.⁸ It is important therefore that conceptual equivalence,
10
11 which is part of the adaptation process, be established before using PBMs cross-culturally.
12
13 Developers of the EQ-5D-Y themselves suggest that future studies should consider other aspects
14
15 of HRQoL that are important to children and adolescents for inclusion in the measure.³⁴ Moreover,
16
17 the measures have only been applied in a small selection of sSA countries, raising questions about
18
19 their applicability across sSA as a whole. Apart from Uganda, where the HUI3 was applied, the
20
21 use of the childhood PBMs was limited to South Africa alone, which is not representative of most
22
23 economies in sSA as it is ranked as an upper middle-income country according to the World
24
25 Bank.^{28,29} The use of the EQ-5D-Y therefore needs to be explored in other countries outside South
26
27 Africa in order to assess generalisability across sSA settings.
28
29
30
31
32
33
34
35

Strengths and limitations

36
37
38 This is the first study to our knowledge that assessed the application and cross-cultural adaption of
39
40 childhood PBMs in sSA. Our search generated a high volume of data, and a large volume of articles
41
42 were excluded during the title screening of our review, which could have led to some relevant
43
44 articles being excluded. Nevertheless, the findings related to our first study objective are largely
45
46 consistent with the most recent review of existing generic childhood PBMs.¹⁰ Although we
47
48 appraised the measures for their psychometric properties, this was limited to the original work by
49
50 the developers. Despite the fact that psychometric evaluations for widely used measures, such as
51
52 the HUI2 and HUI3, are ongoing, for practical reasons we only included the original psychometric
53
54
55
56
57
58
59
60

1 Review of childhood PB HRQoL measures in sSA

2
3 evaluations of each measure.³⁵ Our searches were also limited to peer reviewed material, which
4
5 could have overlooked important evidence in the grey literature. In relation to this, we excluded all
6
7 studies other than those in the English language, which could have had the effect of excluding some
8
9 local studies. Finally, the quality appraisal of the studies applied in sSA settings was performed
10
11 using the CASP case-control checklist due to the lack of toolkits designed specifically for cross-
12
13 sectional studies.
14
15
16
17
18
19
20

21 **Conclusion**

22
23 The findings of this review show that application of generic childhood PBMs in sSA settings has
24
25 hitherto been limited to the HUI3 and EQ-5D-Y-3L. Most adaptations of existing measures take
26
27 an absolutist approach, which assumes that measures can be used across cultures. However, there
28
29 is a need to undertake cross-cultural linguistic and conceptual equivalence, testing and
30
31 psychometric validation across a range of sSA cultural contexts. Unless cross-cultural validation
32
33 of PBMs is robust when applied in the sSA region, the application and interpretation of economic
34
35 evaluations reliant on preference-based outcome measures will fail to reflect local relevance in this
36
37 part of the world.
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1 Review of childhood PB HRQoL measures in sSA
2

3 **Additional files:**
4

5 Table S1: Detailed psychometric properties as evaluated by developers of the identified generic
6 childhood PBMs
7
8
9

10
11
12 Table S2: Detailed cross-cultural validation of generic childhood preference-based HRQoL
13 measures in sSA (n=6)
14
15
16
17
18

19 **Declarations**
20

21 **Competing interests**
22

23
24 The authors declare that they have no competing interests.
25
26
27

28 **Authors' contributions**
29

30 LGN, LN, SP and HM conceived the concept; LGN, KK and LS independently assessed the titles
31 and abstracts; LGN, SS, HM, SP and LN drafted the original manuscript; LGN, LNM, HM, SS, SP
32 and LN reviewed the manuscript. All authors read and approved the final manuscript
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Review of childhood PB HRQoL measures in sSA

REFERENCES

1. WHO G, Switzerland. UNESCO SCIENCE REPORT. last accessed on 11th October 2017.
2. WHO G, Switzerland. Millenium Development Goals and Beyond 2015. Accessed 23rd November 2017, 2017.
3. WHO G, Switzerland. Sustainable development goals. <http://www.un.org/sustainabledevelopment/sustainable-development-goals/>. last assessed 15th September 2017.
4. WHOQOL Group. Development of the World Health Organization WHOQOL-BREF quality of life assessment. *Psychological Medicine*. 1998;28(3):551-558.
5. Brazier J, Deverill M, Green C. A review of the use of health status measures in economic evaluation. *Journal of Health Services Research and Policy*. 1999;4(3):174-184.
6. Wallander J, Schmitt M, Koot H. Quality of Life Measurement in Children and Adolescents: Issues, Instruments, and Applications. *Journal of Clinical Psychology*. 2001;57(4):571-585.
7. Bowden A, Fox-Rushby JA. A systematic and critical review of the process of translation and adaptation of generic health-related quality of life measures in Africa, Asia, Eastern Europe, the Middle East, South America. *Soc Sci Med*. 2003;57(7):1289-1306.
8. Epstein J, Santo RM, Guillemin F. A review of guidelines for cross-cultural adaptation of questionnaires could not bring out a consensus. *J Clin Epidemiol*. 2015;68(4):435-441.
9. Janssens A, Thompson Coon J, Rogers M, et al. A systematic review of generic multidimensional patient-reported outcome measures for children, part I: descriptive characteristics. *Value Health*. 2015;18(2):315-333.

Review of childhood PB HRQoL measures in sSA

10. Chen G, Ratcliffe J. A Review of the Development and Application of Generic Multi-Attribute Utility Instruments for Paediatric Populations. *Pharmacoeconomics*. 2015;33(10):1013-1028.
11. Thorrington D, Eames K. Measuring Health Utilities in Children and Adolescents: A Systematic Review of the Literature. *PLoS One*. 2015;10(8):e0135672.
12. Colbourn T, Masache G, Skordis-Worrall J. Development, reliability and validity of the Chichewa WHOQOL-BREF in adults in Lilongwe, Malawi. *BMC Res Notes*. 2012;5:346.
13. Westmoreland K, Reeve BB, Amuquandoh A, et al. Translation, psychometric validation, and baseline results of the Patient-Reported Outcomes Measurement Information System (PROMIS) pediatric measures to assess health-related quality of life of patients with pediatric lymphoma in Malawi. *Pediatr Blood Cancer*. 2018:e27353.
14. Paltzer J, Barker E, Witt WP. Measuring the health-related quality of life (HRQoL) of young children in resource-limited settings: a review of existing measures. *Qual Life Res*. 2013;22(6):1177-1187.
15. Moher D, Liberati A, Tetzlaff J, Altman DG, Group P. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *J Clin Epidemiol*. 2009;62(10):1006-1012.
16. Mokkink LB, Terwee CB, Patrick DL, et al. The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. *Qual Life Res*. 2010;19(4):539-549.

Review of childhood PB HRQoL measures in sSA

17. Canaway AG, Frew EJ. Measuring preference-based quality of life in children aged 6-7 years: a comparison of the performance of the CHU-9D and EQ-5D-Y--the WAVES pilot study. *Qual Life Res.* 2013;22(1):173-183.
18. (FDA) UDoHaHSFaDA. Guidance for industry: patient-reported outcome measures: use in medical product development to support labeling claims. Rockville, MD: FDA; 2009.
19. Trust SACotMO. Assessing health status and quality-of-life instruments: attributes and review criteria. *Qual Life Res.* 2002;11(3):193-205.
20. Reeve BB, Wyrwich KW, Wu AW, et al. ISOQOL recommends minimum standards for patient-reported outcome measures used in patient-centered outcomes and comparative effectiveness research. *Qual Life Res.* 2013;22(8):1889-1905.
21. Terwee CB, Bot SD, de Boer MR, et al. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol.* 2007;60(1):34-42.
22. Programme CAS. CASP Systematic Review Checklist. 2018; <https://casp-uk.net/casp-tools-checklists/>. Accessed 14th January, 2019.
23. Kang E. Validity of Child Health-6 Dimension(CH-6D) for Adolescents. *Value in Health.* 2016;19(7).
24. Stevens KJ. How Well Do the Generic Multi-attribute Utility Instruments Incorporate Patient and Public Views Into Their Descriptive Systems? *Patient.* 2016;9(1):5-13.
25. Stevens K. Developing a descriptive system for a new preference-based measure of health-related quality of life for children. *Qual Life Res.* 2009;18(8):1105-1113.
26. Sims-Williams HJ, Sims-Williams HP, Mbabazi Kabachelor E, Warf BC. Quality of life among children with spina bifida in Uganda. *Arch Dis Child.* 2017;102(11):1057-1061.

Review of childhood PB HRQoL measures in sSA

27. Jelsma J, Ramma L. How do children at special schools and their parents perceive their HRQoL compared to children at open schools? *Health and Qual Life Outcomes*. 2010;8(1):72.
28. Wille N, Badia X, Bonsel G, et al. Development of the EQ-5D-Y: a child-friendly version of the EQ-5D. *Qual Life Res*. 2010;19(6):875-886.
29. Ravens-Sieberer U, Wille N, Badia X, et al. Feasibility, reliability, and validity of the EQ-5D-Y: results from a multinational study. *Qual Life Res*. 2010;19(6):887-897.
30. Scott D, Ferguson GD, Jelsma J. The use of the EQ-5D-Y health related quality of life outcome measure in children in the Western Cape, South Africa: psychometric properties, feasibility and usefulness - a longitudinal, analytical study. *Health Qual Life Outcomes*. 2017;15(1):12.
31. Jelsma J. A comparison of the performance of the EQ-5D and the EQ-5D-Y health-related quality of life instruments in South African children. *Int J Rehabil Res*. 2010;33(2):172-177.
32. 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 Econ*. 2016;25(4):486-496.
33. Herdman M, Fox-Rushby J, Badia X. 'Equivalence' and the translation and adaptation of health-related quality of life questionnaires. *Quality of Life Research*. 1997;6:237-247.
34. Connolly MA, Johnson JA. Measuring quality of life in paediatric patients. *Pharmacoeconomics*. 1999;16(6):605-625.
35. Feeny D. The Health Utilities Index: A tool for assessing health benefits. *PRO Newsletter*. 2005;34.

Fig 1: Flow-diagram for systematic review of generic childhood preference-based HRQoL measures addressing objective 1

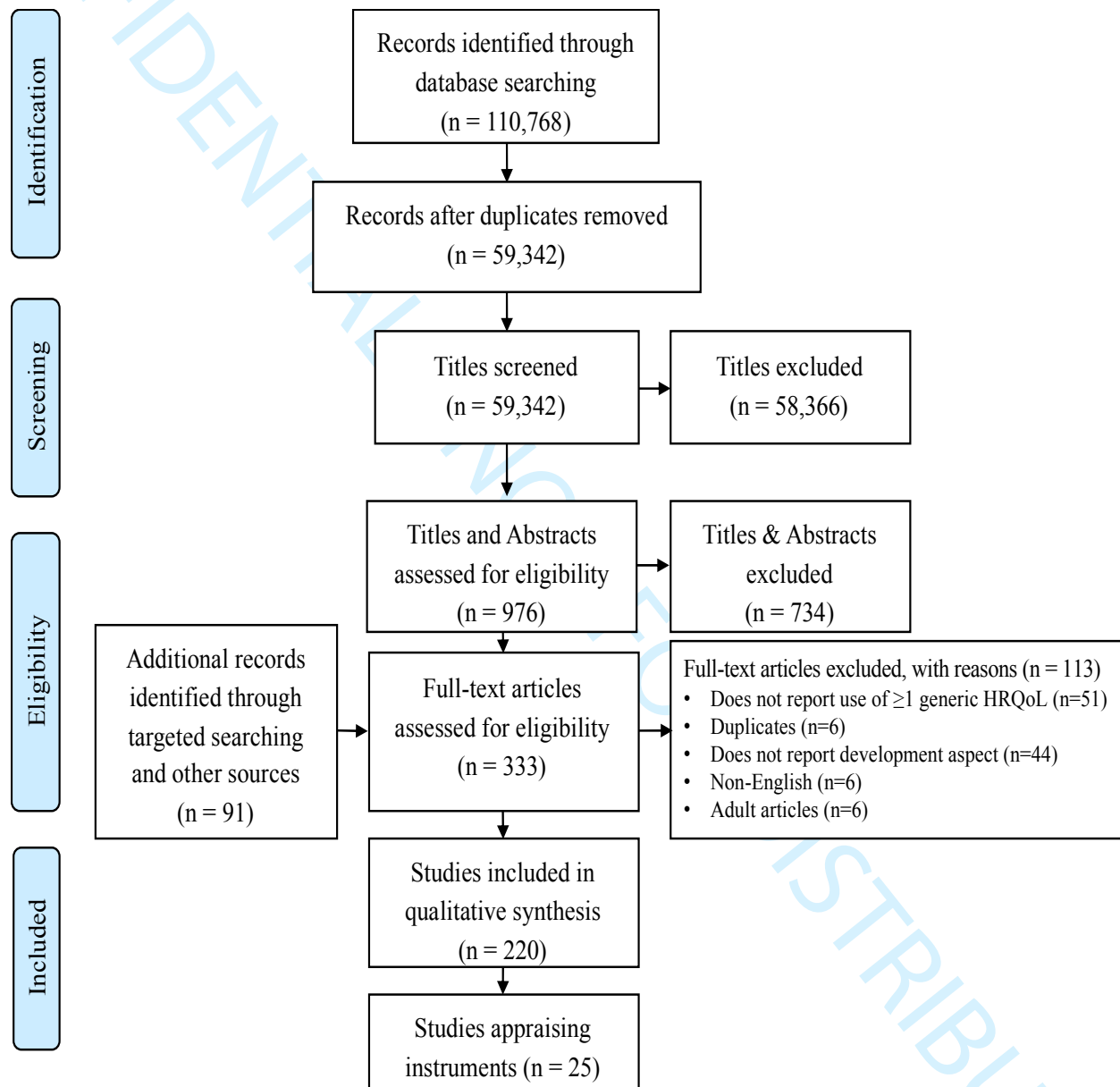
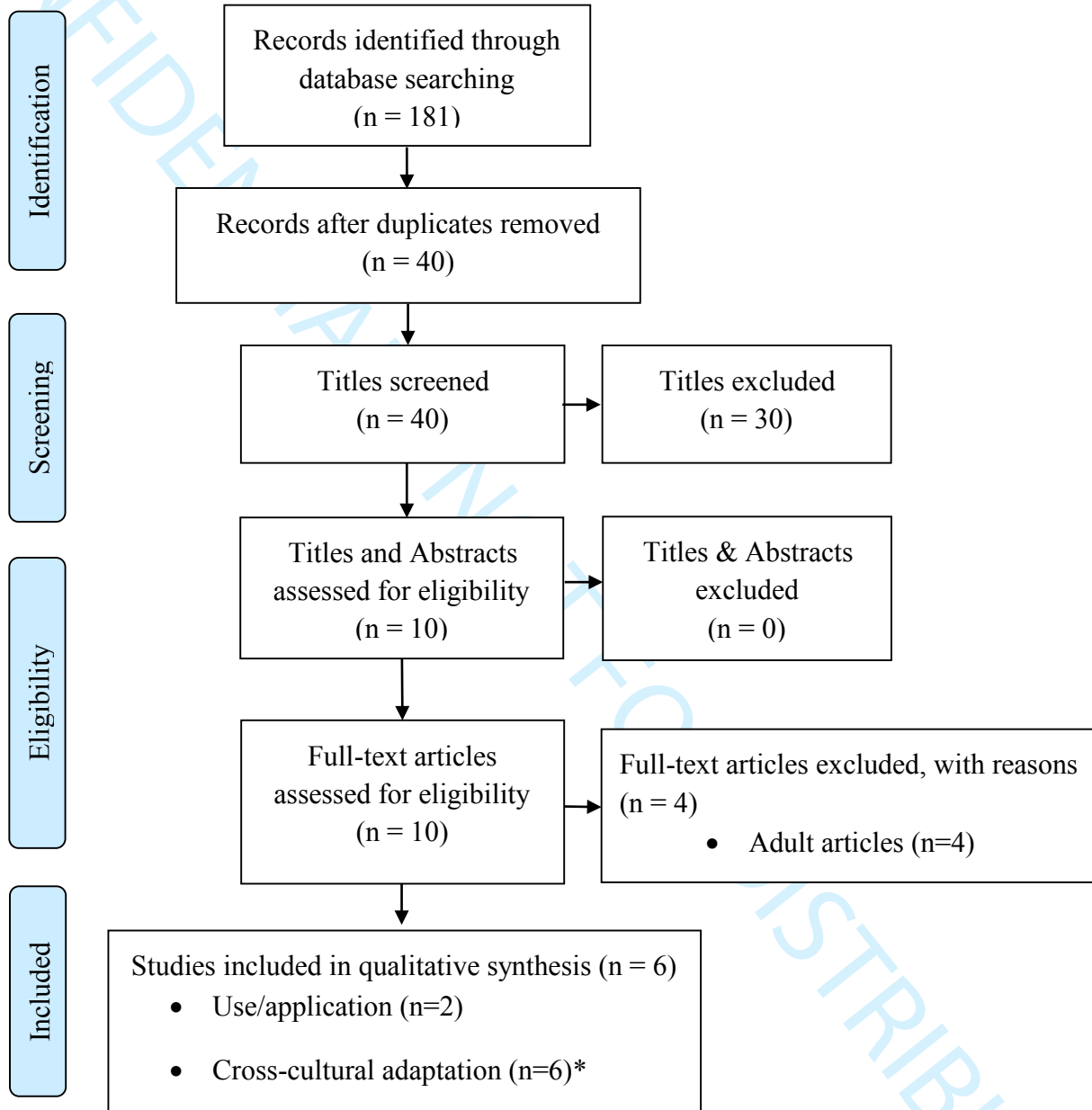


Fig 2: Flow-diagram for generic childhood preference-based HRQoL measures applications and cross-cultural validations in sSA



* application/use of childhood PBMs (n=2) but also discussed aspects of validation

Review of childhood PB HRQoL measures in sSA

Table 1: Generic child and adolescent preference-based HRQoL instruments arranged in chronological order of development

Measure (Acronym)	Age range (yrs)	Author, year (Ref)	Domain development	Domain refinement	Instrument Testing/ pilot	Filled by	Domain s/ Items	Resp -onse level s	Comple tion time (min)	Recall period	Country developed (languages available)
Health Utilities Index Mark 3 (HUI3)	6-16	Feeny et al 1995(27)	Experts	No	No	Self (proxy for 5- 8yrs)	8	5-6 level s	8-10 (min)	Past 4/2/ 1week	Canada, France (35)
Health Utilities Index Mark 2 (HUI2)	6-16	Feeny et al 1992(26)	Experts (surveys and literature review)	Y. 84x Child & parent pairs (same gender + living in same household) rated items	No. Experts decided attributes to keep	Self (proxy for 5- 8yrs)	7	3-5 level s	8-10 (min)	Past 4/2/ 1week	Canada (35)

Review of childhood PB HRQoL measures in sSA

Quality Well-Being Self-Administered (QWB-SA)	12-18	Kaplan et al 1996(31)	Experts (QWB)			Self	3	2-4	NA	3 days	USA (9)
17 Dimensional (17D)	8-11	Apajasalo et al 1996(35)	Experts (16D, literature review)	Pilot study with 79 healthy children + interview	Final dimension decision by experts	Self	17	5	5-10	Today	Finland (4)
16 Dimensional (16D)	12-15	Apajasalo et al 1996(34)	Experts (review of health policy)	Yes (Patient/public involved)	Yes (Patient/public involved)	Self	16	5	5-10	Today	Finland (5)
Assessment of Quality of Life (AQoL-6D)	≥15	Richardson et al 2012(36)	Experts (AQoL)	No (as was developed from AQoL)	Yes, semantics and language were tested	Self	6	4-6	5-10	Past week	Australia (5)

Review of childhood PB HRQoL measures in sSA

Child Health Utility -9D (CHU-9D)	7-17	Stevens 2009(40)	Children (focus group discussion with children)	Yes, Yes (Patient/public involved)	Yes, patients and public	Self	9	5	4-5	Today / last night	UK (7)
Euroqol-5D-Youth (EQ-5D-Y)	8-11	Wille et al 2010(45)	Experts (EQ-5D)	Yes (Patient/public involved)	Yes (Patient/public involved)	Self and/ or proxy	5	3	3	Today	Europe/UK (26)
Adolescent Health Utility Measure (AHUM)	12-18	Beusterien et al 2012(48)	Experts (EQ-5D, SF-6D, and literature review)	Yes. 11x children (11-18yrs) with Hunter syndrome and 27 parents	No	Self	6	4-7	N/A	N/A	UK (N/A)
Child Health-6 Dimensional (CH-6D)	7-12	Kang 2016(49)	N/A	N/A	N/A	Self	6	3-4	N/A	N/A	South Korea (N/A)

NA Not Available

Review of childhood PB HRQoL measures in sSA

CH-6D Child Health 6 Dimension, *AHUM* Adolescent of Health Utility Measure, *AQoL-6D* Assessment of Quality of Life -6 dimension, *EQ-5D-Y* EQ-5D Youth, *CHU-9D* Child Health Utility 9 Dimension, *16D* Sixteen dimensional measure of health related quality of life, *17D* Seventeen dimensional measure of health related quality of life, *QWB-SA* Quality Well Being- Scale, *HUI3* Health Utilities Index Mark 3; *HUI2* Health Utilities Index Mark 2,

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47

CONFIDENTIAL - NOT FOR DISTRIBUTION

Review of childhood PB HRQoL measures in sSA

Table 2 Descriptive system development and utility valuation technique of generic child and adolescent preference-based HRQoL instruments

Measure (Ref)	Health states	Valuation technique	Source preference/ population (Age years)	of HSUV Model (Age years)	Instrument boundaries (min, max)	Year of preference made available
HUI 3 (25, 29, 30, 44)	972,000	SG/ RS	Adult (>=16 yrs)	Multiplicative	0.036, 1.00	2002
HUI2 (25, 28, 30, 44)	24,000	SG/ RS (VAS used then converted to a SG using power function)(26)	Adult proxy	Multiplicative	-0.03, 1.00	1996
QWB-SA (adults and children)(25, 31-33)	945	RS	Adults (>=18yrs)	Statistical additive except for symptom 1 problem	0.08, 1.00	1976

Review of childhood PB HRQoL measures in sSA

17D(10, 35)	7.6x 10 ¹¹	RS	Adult proxy	Additive	0, 1.00	1996
16D(10, 34)	1.5x 10 ¹¹	RS	Adolescents (12-15 yrs)	Additive	0, 1.00	1996
AQoL-6D(25, 37-39)	7.8 x10 ¹³	TTO	Students (adolescents)	Multiplicative/ exponential	0.053, 1.00 min/country: Fiji 0.094; Australia 0.072; Tonga 0.068; NZ 0.053	2004
CHU-9D(25, 41-44)	1,943,125	SG	Adults (mean age= 49yrs)	Additive	0.34, 1.00	2012
EQ-5D-Y(25, 46, 47)	243	Adult utilities TTO, VAS	Adults	Statistical additive	-0.59, 1.00	1993
AHUM (48)	16,800	TTO	Adults (≥18yrs)	Additive	0.42, 0.99	2012
CH-6D (49)	2,304	TTO	General population	N/A	N/A	N/A

CH-6D Child Health 6 Dimension, *AHUM* Adolescent of Health Utility Measure, *AQoL-6D* Assessment of Quality of Life -6 dimension, *EQ-5D-Y* EQ-5D Youth, *CHU-9D* Child Health Utility 9 Dimension, *16D* Sixteen dimensional measure of health related quality of life, *17D* Seventeen dimensional measure of health related quality of life, *QWB-SA* Quality Well Being- Scale, *HUI3*

Review of childhood PB HRQoL measures in sSA

Health Utilities Index Mark 3; *HUI2* Health Utilities Index Mark 2, *N/A* Not Available, *TTO* time trade-off, *VAS* Visual Analogue Scale, *SG* Standard Gamble

Psychometric property*/ Measure	Conceptual framework ^{a,b,c,d}	Cross-cultural & language adaptation ^{a,b,c,d}	Reliability (dependability/ consistency of an instrument) ^{a,b,c,d}			Validity (accuracy/ represent exactly what is on the ground) ^{a, b,c,d}							Interpretation of scores ^{a,b,c,d}	Practicality / Burden ^{a,b,c,d}	Validated in sSA		
			Internal consistency ^{a,b,c,d}	test re-test ^{a,b,c,d}	inter-rater ^{a,b,d}	Content ^{a,b,c,d}	Criterion ^{a,b,c,d}	concurrent	predictive	Construct ^{a,b,c,d}	Convergent ^a	Discriminant ^a	Knowledge group ^a				
HUI2 & HUI3 (30, 44)	NR	+	-	+	NR	NR	NR	NR	NR	NR	NR	NR	+	+	+	NR	Yes
QWB-SA(32, 33)	NR	NR	NR	NR	NR	+	NR	NR	NR	+	NR	NR	NR	NR	NR	+	No
17D (35)	+	NR	+	+	-	+	NR	NR	NR	NR	NR	NR	+	NR	NR	+	No
16D(34)	+	NR	++	+	NR	+	NR	NR	NR	NR	NR	NR	+	NR	NR	+	No
AQoL-6D(38, 39)	+	+	NR	NR	NR	+	NR	NR	NR	+	NR	NR	NR	NR	NR	NR	No
CHU-9D(40, 44)	+	NR	+	NR	NR	+	NR	NR	NR	+	NR	NR	+	NR	NR	+	No

Review of childhood PB HRQoL measures in sSA

EQ-5D-Y(45)	+	+	NR	+	+	-	NR	NR	-	NR	+	NR	+	+	Yes
AHUM(44, 48)	NR	NR	NR	NR	NR	+	NR	NR	NR	NR	NR	NR	NR	NR	No
CH-6D(49)	NR	NR	NR	NR	NR	NR	NR	NR	+	NR	+	NR	NR	NR	No

Table 3 Summary of psychometric properties as evaluated by developers of the identified generic childhood PBMs[#]

[#] Detailed description of psychometric definition, evaluation criteria and reference to specific studies reporting reporting on those properties can be provided on request (Additional files 3).

^a US Department of Health and Human Services Food and Drug Administration (FDA). Guidance for industry: patient-reported outcome measures: use in medical product development to support labeling claims. Rockville, MD: FDA; 2009;

^b Scientific Advisory Committee of the Medical Outcomes Trust. Assessing health status and quality-of-life instruments: attributes and review criteria. Qual Life Res. 2002;11(3):193-205;

^c Reeve BB, Wyrwich KW, Wu AW, Velikova G, Terwee CB, Snyder CF, et al. ISOQOL recommends minimum standards for patient-reported outcome measures used in patient-centered outcomes and comparative effectiveness research. Qual Life Res. 2013;22(8):1889-905;

^d Terwee CB, Bot SD, de Boer MR, van der Windt DA, Knol DL, Dekker J, et al. Quality criteria were proposed for measurement properties of health status questionnaires. J Clin Epidemiol. 2007;60(1):34-42.

*Psychometric evaluation for both measures was extracted from one article provided by the developer; +: indicate meeting minimum review criteria; -: indicate not meeting minimum review criteria; NR: indicate not reported;

HUI2 Health Utility Mark Index 2; HUI3 Health Utility Mark Index 3; QWB-SA Quality of Well-Being Self-Administered; 16D 16Dimension; 17D 17Dimension; AQoL-6D Assessment of Quality of Life 6Dimension adolescent; CHU-9D Child Health Utility 9 Dimension; EQ-5D-Y EuroQoL 5Dimension Youth; AHUM Assessment of Health Utility Measurement; CH-6D Child Health 6Dimension;

1 Review of childhood PB HRQoL measures in sSA
2
3

4 ICC- Intra-class correlation coefficient; MID-meaningfully important difference; +: indicate meeting minimum review criteria; -:
5 indicate not meeting minimum review criteria; NR: indicate not reported
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47

Review of childhood PB HRQoL measures in sSA

Table 4: Quality appraisal of studies identified for application or adaptation of generic childhood PBMs in sSA (n=3)

		Sims-Williams et al (50)	Jelsma & Ramma (51)	Wille et al (45)
	CASP case-control study checklist*			
1	Did the study address a clearly focused issue?	Y	Y	Y
2	Did the authors use an appropriate method to answer their question?	Y	Y	Y
3	Were the cases recruited in an acceptable way?	Y	Y	Y
4	Were the controls recruited in an acceptable way?	Y	Y	N/ A
5	Was the exposure accurately measured to minimise bias?	N/A	N/A	N/ A
6a	Aside from the experimental intervention, were the groups treated equally?	Y	Y	Y**
6b	Have the authors taken account of the potential confounding factors in the design and/or in their analysis?	Y	N#	Y
7	How large was the treatment effect?	Mean VAS score=0.61	Mean VAS score=0.5	N/ A
8	How precise was the estimate of the treatment effect?	N/A	N/A	N/ A
9	Do you believe the results?	Y	Y	Y
10	Can the results be applied to the local population?	Y	Y	Y
11	Do the results in this study fit with other available evidence?	Y	Y	Y

Review of childhood PB HRQoL measures in sSA

*The CASP case-control checklist was used because there is no CASP checklist for appraising cross-sectional studies which was the case with all the three studies; Y=Yes; N/A= Not Applicable; #Was not part of the objectives of the study; **Was a multi-national study across seven countries

Review of childhood PB HRQoL measures in sSA

Table 5: Summary of application of generic childhood preference-based HRQoL measures in sSA (n=2)

Measure	Country	Author, year	Age range (yrs)	Study objective	Study design	study sample size	Methods of recruitment/ selection	Results	Study evaluation
HUI3	Uganda	Sims-Williams HJ, Sims-Williams HP, Mbabazi Kabachelor E, Warf BC. Arch Dis Child. 2017;102(11):1057-61.(50)	10-14	ascertain the QoL of surviving children with spina bifida	Cross-sectional	68	Surgical cases Both children and parent (proxies) completed HUI3 and VAS. The parent (proxies) reported HRQoL from	There was high completion rate by both children and adult proxies, 94% and 100% respectively. The HUI3 HRQoL scores for children and adult proxies was correlated. However, there was poor correlation between QoL VAS scores of children and caregivers.	The HUI3 was found to be acceptable and able to distinguish between children and adult proxy scores in this population. However, there was no

Review of childhood PB HRQoL measures in sSA

							their own perspective		agreement between QoL VAS scores of children and that of their adult proxies.
EQ-5D-Y	South Africa	Jelsma J, Ramma L. Health and Quality Life Outcomes. 2010;8(1):72.(51)	7-12yrs	to compare the HRQoL between healthy children and those with disabilities; and between parents proxies and	Cross sectional	628	All children in eligible grades	There was low completion rate (28.2%-28.4%). The EQ-5D-Y demonstrated practicality (few missing responses). Children with disabilities (mean VAS score=87.9) didn't perceive their HRQoL worse compared	The EQ-5D-Y seems was feasible in this population judging by the few missing responses. However, the low response rate by

Review of childhood PB HRQoL measures in sSA

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47

children's
own health

to those without children could
disabilities (mean VAS not be
score= 88.4). ascertained
Low correlation (r=0.33, whether it was
p<0.001) between due to lack of
children and female comprehension
parent proxies' VAS of the measure
scores in open schools by the children
and no correlation in this setting.
between children and
parent proxies in special
schools (r=0.16, p<0.22).

HUI3 Health Utilities Index Mark 3, EQ-5D-Y EQ-5D Youth

Review of childhood PB HRQoL measures in sSA

Table 6: Summary cross-cultural validation of generic childhood preference-based HRQoL measures in sSA (n=6)

Psychometric property/ Measure (author/ country)#	Linguistic equivalence into any native sSA language		Conceptual equivalence in sSA	Psychometric validation							
				Reliability (dependability/ consistency of an instrument) in sSA		Validity (accuracy/ represent exactly what is on the ground) in sSA				Practicality/ Burden in sSA	
	Forward Translation	Backward Translation	Cognitive interviews	test re-test	inter-rater	Content	Criterion	Construct			Responsiveness
							Concurrent/ predictive	Convergent/ Discriminant	Known group		
HUI3 (Sims-Williams et al/ Uganda)	+	-	-	NR	- +	NR	NR	-	NR	+	+
EQ-5D-Y (Scott et al 2017/ SA)	-	-	-	+	NR	+	-	+	+ - +	++	+
EQ-5D-Y (Jelsma & Ramma 2010/ SA)	-	-	-	NR	-	NR	NR	NR	-	NR	+
EQ-5D-Y EQ5D (Jelsma 2010/ SA)	-	-	-	NR	NR	NR	+	NR	NR	NR	+
EQ-5D-Y (Wille et al 2010/ SA)	-	-	-	NR	NR	NR	NR	NR	NR	NR	+++
EQ-5D-Y (Ravensieberer et al 2010/ SA)	-	-	-	NR	NR	NR	NR	+	+	NR	+

*Internal consistency not reported in any of the studies

#SA South Africa

HUI3 Health Utilities Index Mark 3, EQ-5D-Y EQ-5D Youth

Review of childhood PB HRQoL measures in sSA

Additional files

[Additional file 1 Search terms for identifying available generic childhood preference-based HRQoL measures \(objective 1\)](#)

Search categories		Search terms
Question 1 apply date filter	#30	29 AND (01/01/1990-02/08/2016)
Question 1	#29	8 AND 16 AND 22 AND 28
Psychometric terms	#28	OR (23 to 27)
	#27	psychometrics
	#26	psychometrics[MeSH Terms]
	#25	assess*
	#24	reliab*
	#23	valid*
measure terms	#22	OR (17 to 21)
	#21	measures
	#20	index
	#19	scale
	#18	measurement
	#17	measure
child terms	#16	OR (9 to 15)
	#15	pediatri*
	#14	paediatri*
	#13	child* OR adolesc*
	#12	adolescent
	#11	adolescent[MeSH Terms]
	#10	child
	#9	child[MeSH Terms]
Qol terms	#8	OR (1 to 7)
	#7	health status
	#6	functional status
	#5	Health state utilities
	#4	QALY
	#3	quality adjusted life years
	#2	Quality of life
	#1	Quality of life[MeSH Terms]

Review of childhood PB HRQoL measures in sSA

[Additional file 2 Search terms for identifying applications and cross-cultural validations of generic childhood preference-based HRQoL measures \(objectives 2&3\)](#)

Tool/ Measure		Search terms
	#10	8 AND 9
	#9	sub-saharan africa
	#8	(((((EQ-5D or "EQ 5D" or EQ5D or Euroqol or "Euro qol" or EQ-5D-Y or "EQ 5D Y")) OR ((16D Health-Related Quality of Life or 16D HRQoL or 17D Health-Related Quality of Life or 17D HRQoL))) OR ((AQoL-6D or Assessment of Quality of Life-6D))) OR ((("Child Health Utility 9 Dimension" or CHU9D or CHU-9D or "CHU 9D")) OR ((("quality of well being" or "quality of well-being" or QWB))) OR "health utility index"
16D & 17D	#7	16D Health-Related Quality of Life or 16D HRQoL or 17D Health-Related Quality of Life or 17D HRQoL
AHUM	#6	Adolescent Health Utility Measure or AHUM
AQoL-6D	#5	AQoL-6D or Assessment of Quality of Life-6D
CHU-9D	#4	"Child Health Utility 9 Dimension" or CHU9D or CHU-9D or "CHU 9D"
EQ 5D-Y	#3	EQ-5D or "EQ 5D" or EQ5D or Euroqol or "Euro qol" or EQ-5D-Y or "EQ 5D Y"
HUI2 & HUI3	#2	"health utilities index"
QWB	#1	"quality of well being" or "quality of well-being" or QWB

Review of childhood PB HRQoL measures in sSA

Additional file 3: Detailed psychometric properties as evaluated by developers of the identified generic childhood PBMs*

Measure	Conceptual framework ^{a,b,c,d}	Reliability (dependability/ consistency of an instrument) ^{a,b,c,d}			Validity* (accuracy/ represent exactly what is on the ground) ^{a,b,c,d}			Interpretation of scores ^{a,b,c,d}	Practicality/ Burden ^{a,b,c,d}	Cross-cultural & language adaptation ^{a,b,c,d}	
		Internal consistency ^{a,b,c,d}	test re-test (reproducibility) ^{a,b,c,d}	inter-rater ^{a,b,d}	Content ^{a,b,c,d}	Construct ^{a,b,c,d}	Responsiveness ^{a,b,c,d}				
						Convergent ^a	Discriminant ^a	Known group ^a			
HUI2 & HUI3*		Low correlation among attributes compared to other measures (1)	ICC≥0.70 (1)					MID≥0.03 for clinically different groups (1)	Demonstrated responsiveness (1)	Few ceiling & floor effect in clinical studies but in health surveys(1)	
QWB-SA					<i>QWB has content validity similar to SF-36 (2)</i>			<i>QWB-SA and QWB were found to highly correlated (3)</i>		<i>Found to be practical (2)</i>	
16D		Repeatability coefficient by individual level	Good. Overall reliability coefficient r=0.91 (lying 2		Developed by a team of children and adolescent experts and domains			Able to differentiate between sick and healthy children		Easy to complete taking between five to ten minutes	

Review of childhood PB HRQoL measures in sSA

		dimension s ranged from 86- 100%, and majority (11 out of 16 dimension s) had repeatabilit y coefficient $\geq 97\%$ (<i>only sleeping, and discomfort and symptoms <94%</i>). (4); Cronbach' s $\alpha =$ 0.831(5)	standard deviation (SD) from mean difference) and between 0.86-1.00 by dimension being over and above the accepted standards (4)		were mostly modified from an existing 15D instrument (4)			HRQoL states- (4)		(4) <i>quite quick to complete among psychiatri c symptoms (5)</i>
17D	A clear methodo logy of question naire develop ment. Question naire revised with input from	11 out of 17 dimension s had reliability coefficient $r = 0.97$ with only two (sleeping and breathing) $r < 0.93$. (6)	Reliability coefficient overall score, $r =$ 0.95 with cases lying within 2 SD of the mean difference and by dimensions $r = 0.91-1.0$ (6)	No evidence of inter- rater correlati on between children and proxies especial ly on	Developed from 16D by a team of experts in paediatrics and adolescent (6)			Children with more health problems not only reported more problems but also scored lower (0.895 \pm 0.072) than		Children did not just complete filling the form but also found it enjoyable . (6)

Review of childhood PB HRQoL measures in sSA

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47

	children (6)			speech (6)				those without health problem (0.937+/-0.053) on particular dimensions . (6)				
AQoL-6D	Four broad steps involving interviews with target population. (7)				Five of the six dimensions had a correlation coefficient of ≥ 0.73 (lambda loading gamma coefficients are equivalent of correlation coefficient) except for vision (=0.51) (8)	Five of the six dimensions had a correlation coefficient of ≥ 0.73 (lambda loading gamma coefficients are equivalent of correlation coefficient) except for vision (=0.51) (8)						Demonstrated good cross-cultural validity when adapted for adolescents in four pacific countries (7, 9)
CHU-9D	<i>Excellent. Qualitative work, cognitive interviews then constructed</i>	Cronbach's $\alpha = 0.781(12)$	NR	NR	Excellent Only child MAU instrument with domains constructed directly	Child-rated level of health correlated with descriptive system demonstrati	NR	CHU9D able to discriminate between clinically known different groups	NR	NR	Response rate, completion rate for each question and mean completio	NR

Review of childhood PB HRQoL measures in sSA

	<i>ted descriptive system (10, 11)</i>				from children (7-11 yrs) using qualitative research (10, 13) Good content validity as only 11% of children felt something missing in content (12)	ng construct validity (12)		based on indicators of long-standing illness and disability and general health as self-reported (14-18)			n time that ranged from 64% - 98%, 98.7%-100% and 3.8 - 5.3 minutes respectively, many finding it quick and easy to administer.(12)	
EQ-5D-Y	The modification process included revision of domain definitions, revision of wording and layout, translation, cognitive interviews,	NR=0	For dimensions, agreement ranged from 69.8 to 99.7% in the EQ-5D-Y; Kappa coefficients up to 0.67 (19)	<i>Good. Self-report and proxy version agreement ranged from 72.9% to 97.1% (19)</i>	<i>Poor, as it was developed from adult EQ-5D version, domains not generated specifically for children. There was also specifically lack of in-depth examination of domain conceptualization (20)</i>	Kappa coefficients were up to 67%, but no correlation with other self-rated health measures, correlation coefficient ($r=-0.56$) (19)	NR=0	Detect differences between known clinically ill groups (19)	NR=0	<i>Between 91 to 100% provided valid scores (19)</i>	<i>Found to be feasible in a number of countries with either nothing missing or very few inappropriate responses both for DS and VAS ranging from 0-2%, and</i>	There is also evidence of cross-cultural validity from multiple trials showing EQ-5D-Y is well understood and practical across several countries (19)

Review of childhood PB HRQoL measures in sSA

	integrati on and decision- making on harmoni zation, and compari ng the two versions. <i>Cognitiv e intervie ws involved target populati on (19)</i>									<i>time taken to complete <5 minutes (19)</i>
AHUM					<i>Contents derived from EQ- 5D and SF- 6D (21)</i>					
CH-6D						Spearman correlation coefficient with HUI2 was 0.658 (22)		Able to differentiat e between those with known different status (22)		

*Criterion validity (concurrent and/or predicted) not reported by any study

^a US Department of Health and Human Services Food and Drug Administration (FDA). Guidance for industry: patient-reported outcome measures: use in medical product development to support labeling claims. Rockville, MD: FDA; 2009;

Review of childhood PB HRQoL measures in sSA

^b Scientific Advisory Committee of the Medical Outcomes Trust. Assessing health status and quality-of-life instruments: attributes and review criteria. *Qual Life Res.* 2002;11(3):193-205;

^c Reeve BB, Wyrwich KW, Wu AW, Velikova G, Terwee CB, Snyder CF, et al. ISOQOL recommends minimum standards for patient-reported outcome measures used in patient-centered outcomes and comparative effectiveness research. *Qual Life Res.* 2013;22(8):1889-905;

^d Terwee CB, Bot SD, de Boer MR, van der Windt DA, Knol DL, Dekker J, et al. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol.* 2007;60(1):34-42.

*Psychometric evaluation for both measures was extracted from one article provided by the developer;

+: indicate meeting minimum review criteria; -: indicate not meeting minimum review criteria; NR: indicate not reported;

HUI2 Health Utility Mark Index 2; HUI3 Health Utility Mark Index 3; QWB-SA Quality of Well-Being Self-Administered; 16D 16Dimension; 17D 17Dimension; AQoL-6D Assessment of Quality of Life 6Dimension adolescent; CHU-9D Child Health Utility 9 Dimension; EQ-5D-Y EuroQol 5Dimension Youth; AHUM Assessment of Health Utility Measurement; CH-6D Child Health 6Dimension;

ICC- Intra-class correlation coefficient; MID-meaningfully important difference; +: indicate meeting minimum review criteria; -: indicate not meeting minimum review criteria; NR: indicate not reported

Review of childhood PB HRQoL measures in sSA

Additional file 4: Detail cross-cultural validation of generic childhood preference-based HRQoL measures in sSA (n=6)

Psycho metric property / Measure (author/ country)	Conceptual framework	Reliability* (dependability/ consistency of an instrument) in sSA		Validity (accuracy/ represent exactly what is on the ground) in sSA						Practicality / Burden in sSA	Cross-cultural & language adaptation in sSA	
		test re-test	inter-rater	Content	Criterion	Construct			Respon siveness			
					concu rent	predi ctive	Converg ent	Discri minant	Known group			
HUI3 (Sims-William s et al/ Uganda)	+ Developed from HUI2	NR	- No correlation between children and proxies using VAS scores and HUI3 and VAS scores, pearson correlation=0.133 and 0.380 respectively; +High correlation (0.848) between children and proxies using	NR	NR	- Not highly correlated (pearson correlation=0.488) between HUI3 and VAS scores.			NR	+ Large treatment effect (Wilcoxon signed-rank test for VAS=0.43) observed in acutely ill children.	+Feasible: only 4 of 68 children unable to complete the HUI3 due to deafness and/or learning disability.	+ (adaptation involved translation of the HUI3); + (cognitive interviews with parents and children)

Review of childhood PB HRQoL measures in sSA

			HUI3							
EQ-5D-Y (Scott et al 2017/ South Africa)	+ Developed from adult EQ-5D version	+All dimensions (Kappa 0.365-0.653) except usual activities (kappa=0.199); + VAS scores ICC=0.77	NR	+ Nine clinical therapists found it to have right content	- No evidence of correlation between VAS and composite scores of different groups except the acutely ill	+ Only evident in acutely ill between EQ-5D-Y and PedsQL	+On dimension: Able to discriminate across groups with different health states (acutely ill v healthy children (p<0.001)); -On composite scores: unable to diccriminate +On VAS scores: able to discriminate between groups (p<0.001)	++ able to detect medium effect size for both the chronic ill and acute ill children for both the composite and VAS scores	+ All children completed the EQ-5D-Y and was found to be quick and easy to apply by six of nine therapists who took part in the study	- (no translation was required as the English version was used); + (the original EQ-5D-Y established this during development which included South Africa)
EQ-5D-Y (Jelsma & Ramma 2010 /	+ Developed from adult EQ-5D version	NR	- No agreement between Children and their proxies on all	NR	NR	NR	-Children with disabilities didn't perceive their HRQoL	NR	+ Of 62 responents, there were only 5 missing responses	- (none required was in English); + (the original EQ-5D-Y

Review of childhood PB HRQoL measures in sSA

South Africa)			domains (Kappa= 0.01-0.60)				worse compared to those without			established this during development)
EQ-5D-Y EQ5D (Jelsma 2010 /South Africa)	+ Developed from adult EQ-5D version	NR	NR	NR	+ EQ-5D-Y performed better than EQ-5D	NR	NR	NR	++ More missing responses on the EQ-5D compared to EQ-5D-Y both for dimension and VAS, Chi-squared=9.404; d.f.=1; (P=0.002a nd Chi-squared=26.159; d.f.=1; (P<0.001) respectively	- (none required was in English); + (the original EQ-5D-Y established this during development which included South Africa)
EQ-5D-Y (Wille et al 2010/ South	+ Developed from adult EQ-5D version	NR	NR	NR	NR	NR	NR	NR	++ Children and adolescents	- (none required was in English); + (the original EQ-

Review of childhood PB HRQoL measures in sSA

Africa)									satisfactorily understood EQ-5D-Y than EQ-5D. Also, better accepted and proved more feasible than the EQ-5D	5D-Y established this during development which included South Africa)
EQ-5D-Y (Ravensieberer et al 2010/ South Africa)	+ Developed from adult EQ-5D version	NR	NR	NR	NR	+ There was evidence of convergent validity between the EQ-5D-Y and three other measures: KIDSCREEN-10 HRQoL Index, KIDSCREEN-27 and the Life Satisfaction Ladder.	+ There was evidence of known-group validity between groups with different health conditions	NR	+ Missing or inappropriate responses on the EQ-5D-Y was 2%	- (none required was in English); + (the original EQ-5D-Y established this during development which included South Africa)

* Internal consistency not reported by any study

*Psychometric evaluation for both measures was extracted from one article provided by the developer;

Review of childhood PB HRQoL measures in sSA

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

+: indicate meeting minimum review criteria; -: indicate not meeting minimum review criteria;
NR: indicate not reported;

HUI2 Health Utility Mark Index 2; HUI3 Health Utility Mark Index 3; QWB-SA Quality of Well-Being Self-Administered; 16D 16Dimension; 17D 17Dimension; AQoL-6D Assessment of Quality of Life 6Dimension adolescent; CHU-9D Child Health Utility 9 Dimension; EQ-5D-Y EuroQol 5Dimension Youth; AHUM Assessment of Health Utility Measurement; CH-6D Child Health 6Dimension;

ICC- Intra-class correlation coefficient; MID-meaningfully important difference; +: indicate meeting minimum review criteria; -: indicate not meeting minimum review criteria; NR: indicate not reported

Review of childhood PB HRQoL measures in sSA

REFERENCES

1. Feeny D. The Health Utilities Index: A tool for assessing health benefits. *PRO Newsletter*. 2005;34.
2. Kaplan R, Ganiats T, Sieber W, Anderson J. The Quality of Well-Being Scale: critical similarities and differences with SF-36. *International Journal for Quality of Health Care*. 1998;10(6):509-20.
3. Kaplan R, Sieber W, Ganiats T. The quality of well-being scale: comparison of the interviewer-administered version with a self-administered questionnaire. *Psychol Health*. 1997;12(6):783-91.
4. Apajasalo M, Sintonen H, Holmberg C, Sinkkonen J, Aalberg V, Pihko H, et al. Quality of life in early adolescence: A sixteen-dimensional health-related measure (16D). *Quality of Life Research*. 1996;5:205-11.
5. Grano N, Kiesepa T, Karjalainen M, Roine M. Exploratory factor analysis of a 16D Health-Related Quality of Life instrument with adolescents seeking help for early psychiatric symptoms. *Nord J Psychiatry*. 2016;70(2):81-7.
6. Apajasalo M, Rautonen J, Holmberg C, Sinkkonen J, Aalberg V, Pihko H, et al. Quality of life in pre-adolescence: a 17-dimensional health-related measure (17D). *Qual Life Res*. 1996;5(6):532-8.
7. Richardson J, Peacock S, Iezzi A, Day N. Construction and Validation of the Assessment of Quality of Life (AQoL) Mark II Instrument. Melbourne: Monash University; 2007.
8. Richardson JR, Peacock SJ, Hawthorne G, Iezzi A, Elsworth G, Day NA. Construction of the descriptive system for the Assessment of Quality of Life AQoL-6D utility instrument. *Health Qual Life Outcomes*. 2012;10:38.
9. Moodie M, Richardson J, Rankin B, Iezzi A, Sinha K. Predicting time trade-off health state valuations of adolescents in four Pacific countries using the Assessment of Quality-of-Life (AQoL-6D) instrument. *Value Health*. 2010;13(8):1014-27.
10. Stevens K. Developing a descriptive system for a new preference-based measure of health-related quality of life for children. *Qual Life Res*. 2009;18(8):1105-13.
11. Stevens KJ. How Well Do the Generic Multi-attribute Utility Instruments Incorporate Patient and Public Views Into Their Descriptive Systems? *Patient*. 2016;9(1):5-13.
12. Stevens K. 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. *Appl Health Econ Health Policy*. 2011;9(3):157-69.
13. Stevens KJ. Working with children to develop dimensions for a preference-based, generic, pediatric, health-related quality-of-life measure. *Qual Health Res*. 2010;20(3):340-51.
14. Canaway AG, Frew EJ. Measuring preference-based quality of life in children aged 6-7 years: a comparison of the performance of the CHU-9D and EQ-5D-Y--the WAVES pilot study. *Qual Life Res*. 2013;22(1):173-83.
15. Furber G, Segal L. The validity of the Child Health Utility instrument (CHU9D) as a routine outcome measure for use in child and adolescent mental health services. *Health Qual Life Outcomes*. 2015;13:22.
16. 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. *Qual Life Res*. 2012;21(4):717-25.
17. Stevens K, Ratcliffe J. 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. *Value Health*. 2012;15(8):1092-9.
18. Chen G, Flynn T, Stevens K, Brazier J, Huynh E, Sawyer M, et al. Assessing the Health-Related Quality of Life of Australian Adolescents: An Empirical Comparison of the Child Health Utility 9D and EQ-5D-Y Instruments. *Value Health*. 2015;18(4):432-8.

Review of childhood PB HRQoL measures in sSA

19. Ravens-Sieberer U, Wille N, Badia X, Bonsel G, Burstrom K, Cavrini G, et al. Feasibility, reliability, and validity of the EQ-5D-Y: results from a multinational study. *Qual Life Res.* 2010;19(6):887-97.
20. Wille N, Badia X, Bonsel G, Burstrom K, Cavrini G, Devlin N, et al. Development of the EQ-5D-Y: a child-friendly version of the EQ-5D. *Qual Life Res.* 2010;19(6):875-86.
21. Beusterien KM, Yeung JE, Pang F, Brazier J. Development of the multi-attribute Adolescent Health Utility Measure (AHUM). *Health Qual Life Outcomes.* 2012;10:102.
22. Kang E. Validity of Child Health-6 Dimension(CH-6D) for Adolescents. *Value in Health.* 2016;19(7).