PharmacoEconomics
https://doi.org/10.1007/s40273-024-01357-z

SYSTEMATIC REVIEW

Exploring the Use of Pictorial Approaches in the Development of Paediatric Patient-Reported Outcome Instruments: A Systematic Review

Christine Mpundu-Kaambwa[1] · Norma B. Bulamu[2] · Lauren Lines[1] · Gang Chen[3] · David G. T. Whitehurst[4] · Kim Dalziel[5,6] · Nancy Devlin[6] · Julie Ratcliffe[1] on behalf of the Quality of Life in Kids: Key Evidence for Decision Makers in Australia (QUOKKA) project team

Accepted: 30 October 2023

Abstract

Introduction Children may find self-reporting health-related quality of life (HRQoL) using patient-reported outcome measures (PROMs) presented in text-based formats difficult, particularly younger children and children with developmental delays or chronic illness. In such cases, pictorial PROMs (where pictorial representations are used alongside or to replace text) may offer a valid alternative.

Aim This systematic literature review focused on identifying and describing paediatric PROMs that incorporate pictorial approaches, providing children with more effective means to express their HRQoL.

Methods Preferred Reporting Items for Systematic Reviews and Meta-Analyses guidelines were followed. Seven electronic databases were searched from inception to 1 March 2022. There were no country restrictions applied to the search; all English-language studies were considered for inclusion in the review. Characteristics and development methods of the identified pictorial PROMs were evaluated against context-specific good practice guidelines published by The Professional Society for Health Economics and Outcomes Research (ISPOR).

Results A total of 22 paediatric pictorial PROMs, comprising 28 unique versions, were identified. These PROMs were predominantly developed in the USA and the UK, targeting children aged 3–18 years. Likert scales with pictorial anchors, particularly happy–sad faces, were commonly used for response options, appearing in 15 (54%) of the PROMs. Various graphic methods, such as happy–sad faces, cartoons, and thermometers, were adapted to specific content domains. These PROMs covered a wide range of domains, including physical and emotional health and social functioning. Emphasis was placed on content validity, including active child participation in developing pictorial elements. Notably, children's participation was sought during the development of the pictorial elements for 13 (46%) of the PROMs. Various development methods were employed, with 43% of paediatric PROMs using literature reviews, 43% using focus groups, and 32% involving expert consultation. Interviews emerged as the primary method, being employed in 61% of the studies. Additionally, three measures specifically addressed cross-cultural considerations.

Conclusion Paediatric pictorial PROMs offer child-friendly tools for assessing HRQoL for application with children who find reading and understanding text-based PROMs challenging. There is some evidence that pictorial PROMs facilitate self-report in this population and improve measurement properties compared to text-only PROMs. Further research is needed to develop, validate, and test paediatric pictorial PROMs, with an emphasis on including children from the inception in the co-design process.

1 Introduction

Patient-reported outcome measures (PROMs) are increasingly being incorporated into clinical practice and health research to inform economic evaluation and quality assessment [1–3]. PROMs report an individual's subjective assessment of their health-related quality of life (HRQoL) at a particular time point or repeated time intervals over an extended period. PROMs may vary from simple assessments to complex multi-dimensional instruments [4]. PROMs can be categorised into those that are preference weighted and those with scoring systems that are not based

Extended author information available on the last page of the article

Published online: 13 February 2024

Key Points for Decision Makers

Pictorial patient-reported outcome measures (PROMs) improve paediatric health-related quality of life (HRQoL) assessment.

Pictorial HRQoL PROMs are well-suited for children facing difficulties in understanding textual content.

Incorporating children in the development of pictorial HRQoL PROMs improves content validity.

When choosing pictorial HRQoL PROMs, it is essential to factor in their cross-cultural applicability.

on preferences; the former are suitable for application in economic evaluation.

Preference-weighted PROMs incorporate scoring algorithms that are typically based on the preferences assigned by general population samples, generated using valuation methods such as discrete choice modelling, standard gamble, or time trade-off, and are typically anchored on the 0 (dead) to 1 (full health) utility scale [4].

PROMs can be further differentiated into condition-specific (e.g. Paediatric Asthma Health Outcome Measure [PAHOM]) [5] and generic (e.g. EQ-5D-Y) [6] PROMs. Generic PROMs are designed for application in all subgroups and allow comparisons between them. Condition-specific PROMs are intended to apply to a single group (e.g. individuals with a specific health condition) and, compared with generic PROMs, generally have a higher sensitivity for detecting changes resulting from the condition or associated factors [7, 8].

Most instrument developers for PROMs for measuring HRQoL in children require for self-completion that children are 8 years old or older [9–11]; for younger children and those with developmental delays and/or intellectual impairments, proxy-reported (by a parent/guardian and/or health care professional) or interviewer-administered modes of administration are recommended [4]. The Professional Society for Health Economics and Outcomes Research (ISPOR) Patient Reported Outcomes (PRO) Good Research Practices for assessing children and adolescents Task Force Report found insufficient evidence to establish whether self-reporting of HRQoL by children younger than 8 years of age is reliable and valid [12]. Self-reporting of HRQoL is generally considered preferable to proxy assessment wherever possible [4].

Prior research in HRQoL reporting in children has reported discrepancies between child self-assessments and proxy assessments [13–16], and the need for psychometrically sound measures for young children [17]. Three previous systematic reviews on the accuracy of proxy-reported versus child-self-reported assessments in deriving childhood utility values for economic evaluation found disparities between the two perspectives [13–15].

To support decision-making in research and clinical practice, the ISPOR PRO Good Research Practices report makes five recommendations for assessing childhood PROMs: (1) consider developmental differences and determine age-based criteria for PROM administration; (2) establish content validity of paediatric PROM instruments; (3) determine whether an informant-reported outcome instrument is necessary; (4) ensure that the instrument is designed and formatted appropriately for the target age group; and (5) consider cross-cultural issues [12].

It is well documented that younger children [18, 19] and children with severe chronic conditions [20] face challenges with self-reporting health-related outcomes using PROMs presented in traditional text-based formats (written words). Though proxy reporting is an option, finding ways to support and extend self-report in the aforementioned groups is essential to ensure that HRQoL data represents the child's views. Measures that use pictorial representations such as pictures, cartoons, and smiley faces (happy–sad faces) to convey information may facilitate reliable self-assessment of HRQoL and ensure the inclusion of children who find reading text challenging.

Previous literature reviews have shown that incorporating pictorial formats alongside text-based information may facilitate a better understanding of PROMs to assess HRQoL among children [21, 22]. Images can also facilitate word memory, increased attention, and conceptual processing, and they demand less cognitive effort to understand than words [21, 22]. Other advantages of using pictures to support written text include engaging young children's interest and sustaining attention, which may contribute to more meaningful responses and improve completion rates [23].

Four systematic reviews have assessed HRQoL measures used in child populations, which included several pictorial PROMs [19, 24–26]. However, these reviews have highlighted certain limitations that are relevant to the objectives of the present systematic literature review. The previous reviews by Cremeens et al. [25] (limited to ages 3–8 years), Arsiwala et al. [26] (primarily focused on generic PROMs with Likert scales and with pictorial elements), and Coombes et al. [24] (focused on 'faces' scales for pain reporting, limited diversity, age group specificity, incomplete coverage of recent developments) all had certain limitations that are relevant to the objectives of this systematic literature review. These limitations include narrow scope, age group specificity, limited diversity in pictorial scales, and potential gaps in coverage of recent developments. Additionally, the study by Solans et al. [19] was undertaken some time ago

(1980–2000 timeframe) and primarily focused on generic and disease-specific PROMs without specific attention to pictorial approaches. These limitations indicate the need for a more comprehensive and up-to-date review, encompassing a wider age range, diverse pictorial approaches, and recent developments in paediatric PROMs that incorporate pictorial elements. Therefore, the aims of this review were twofold, to (1) identify PROMs designed for children that use pictorial approaches (where pictorial representations are used either alongside or replace text) to assess HRQoL and (2) provide a comprehensive description of the development and key characteristics of the instruments.

The assessment of these PROMs focussed on ISPOR PRO Good Research Practices report recommendations 1, 2, 4, and 5. The scope of the review did not include proxy-reported outcomes, which typically involve reports from caregivers or parents. Therefore, we did not address ISPOR PRO Good Research Practices recommendation 3, which pertains to informant-reported outcome instruments, as it did not align with the objectives and scope of our study.

In our review, the ISPOR recommendations for the evaluation of pictorial paediatric PROMs were incorporated as follows:

1. Developmental differences were considered by assessing the appropriateness of pictorial elements for various age groups and addressing age-related factors.
2. Content validity was assessed by examining children's involvement in creating pictorial elements and their alignment with specific content domains.
3. While not explicitly addressed, our review centred on self-reported aspects of pictorial paediatric PROMs, which are typically child centred.
4. Suitable design and formatting were emphasised by evaluating how graphical content aligned with cognitive abilities and readability, highlighting features like large fonts and accessibility options.
5. Cultural considerations in the development of pictorial PROMs were discussed, acknowledging the need for further exploration of cross-cultural applicability.

Although no exclusive frameworks exist for pictorial PROMs, we adapted ISPOR guidelines originally designed for PRO instruments, enabling us to effectively assess the development and content validity of these instruments. This adaptation clarified our rationale for including relevant information and enhanced the alignment of our review with a focus on the presentation of graphical techniques in pictorial paediatric PROMs.

2 Methods

2.1 Protocol and Registration

The protocol for this review was registered with the International Prospective Register of Systematic Reviews (PROSPERO), registration number CRD42021222771. This review was conducted following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) 2020 guidelines for systematic reviews [27, 28]. The PRISMA checklist is presented in Table B in the electronic supplementary material (ESM).

2.2 Data Sources and Search Strategy

The systematic search covered seven electronic academic research databases: Medline (via Ovid), Scopus, Web of Science Core Collection (via ISI Web of Science), CINAHL (via EBSCOHost), Emcare, and Embase and PsycINFO (both via Ovid). A combination of Medical Subject Heading (MeSH) terms and keyword searches were used and adapted to enhance search sensitivity and specificity across databases. The initial search encompassed content from the inception of each database up to 1 March 2022. This process was then updated for each individual database, spanning from its inception to 31 January 2023.

Four themes were used to categorise the search terms. These were (1) ‘health-related quality of life’; (2) communication methods, such as ‘pictogram’, ‘pictorial’, and ‘animation’; (3) the population (e.g. ‘child’ and ‘adolescent’); and (4) the instrument (e.g. ‘tool’, ‘measure’, and ‘questionnaire’). The full search strategy is presented in Table A in the ESM. The search strategy was first developed by the first author (CMK) and refined with feedback from a research librarian and the authorship team. The research librarian translated the search strategy for each database.

2.3 Selection Criteria

2.3.1 Inclusion Criteria

This systematic review included peer-reviewed studies of children < 18 years of age that focus on the development of HRQoL PROMs that use pictorial representations, such as animations, pictures, pictograms, and easy-to-read techniques all designed to enhance children's comprehension of information. Our criteria encompassed both generic and condition-specific HRQoL PROMs intended for children that were either self-report or interviewer assisted. We also considered both preference-weighted and non-preference-weighted HRQoL PROMs. We also extended our criteria beyond published articles to include discussion papers,

reports, and published theses for a comprehensive evaluation. This comprehensive approach allowed us to thoroughly assess studies within the scope of our systematic review.

Our search criteria did not impose any country restrictions; we considered all English-language studies for inclusion in the review. For studies identified in languages other than English, we used Google Translate [29] to ensure their inclusion in the review.

2.3.2 Exclusion Criteria

Studies that included participants 18 years of age or over were excluded. Studies examining participants across both child and adult age ranges were only included if the majority of the sample were children under 18 years and if results for the children were reported separately. Reviews, reports, conference papers, book chapters, and opinion pieces were also excluded.

2.4 Article Screening

Identified articles were imported into Endnote X9.3 (2020) reference management software and transferred to Covidence (www.covidence.org), online screening and data extraction software for systematic reviews. The title and abstract of each identified paper were reviewed independently by two reviewers from a team of three (CMK, LL, NB) to determine eligibility. Disagreements were resolved via discussion among all three reviewers. Full texts of publications were further screened independently by each of the three reviewers against the inclusion criteria. All full texts were forward and backward citation checked.

Overall agreement among the reviewers was assessed by Cohen's kappa, with categories defined as follows: values $\leq$ 0 indicated no agreement, 0.01–0.20 none to slight (poor), 0.21–0.40 fair, 0.41–0.60 moderate, 0.61–0.80 substantial (good), 0.81–0.99 almost perfect (very good), and 1.00 perfect agreement [30].

2.5 Data Extraction

A standardised data extraction form was developed to ensure consistency and completeness when extracting data from the included articles. The unit of analysis was a PROM and not a research article. Data for each PROM were extracted independently by two reviewers. Discrepancies throughout this process were resolved through discussion. The data extraction form included the following:

1. Characteristics of the PROM: name and acronym of the PROM, name of the author(s)/developer(s), year developed, country/countries where the measure was developed, target age range, type of respondent (e.g. child, proxy), mode of administration (e.g. self-report, interviewer administered), recall period, number of dimensions and items, response options, scoring methods, preference-based or non-preference-based measures, length of PROM, pictorial representations, and electronic data collection (ePRO).
2. Content and item generation process or methodology (e.g. literature review, focus groups, cognitive interviews) and whether children were involved in item generation, pilot testing, and content validation.

2.6 Data Synthesis

An overview of the PROMs identified was presented, with the rest of the data synthesised according to ISPOR PRO Good Research Practices report recommendations 1, 2, 4, and 5 (each described below). A description of measurement properties of the PROMs, if reported, was also provided. Note that a full appraisal of measurement properties for the identified PROMs was beyond the scope of the review. Given the nature of our research question and the focus on the description of these instruments, we did not conduct a risk of bias assessment for the included articles.

In the overview of PROMs identified, the type of PROM (generic and condition- and domain-specific PROMs), health conditions focussed on for condition-specific PROMs, the scoring method used (preference or non-preference based), year of development, country of development, target population, setting in which the PROM was developed, and versions of PROMs available were reported.

To align with ISPOR PRO Good Research Practices report recommendation 1 (determine developmental differences and age-based criteria for PROM administration), the following age groups were reported: < 5 years or early childhood (including infants, toddlers, or pre-schoolers); 5–7 years (younger children); 8–11 years (pre-adolescents); and 12–18 years (adolescents) [12]. PROMs covering more than one age group were classified into a fifth group, 'multi-age'.

To address ISPOR PRO Good Research Practices report recommendation 2 (establish the content validity of paediatric PROMs), we assessed the methodology and processes used to develop these PROMs. Each PROM was summarised according to (1) the methods used for concept elicitation and development and domain and item generation and (2) whether individuals from the target population were involved in the development, pilot testing, and/or validation of the measure.

To investigate ISPOR PRO Good Research Practices report recommendation 4 (ensure that the instrument is designed and formatted appropriately for the target age group), the response scales were categorised according to the scale format used (Likert or visual analogue) and presentation style, i.e. written, pictorial, verbal/audio, animation,

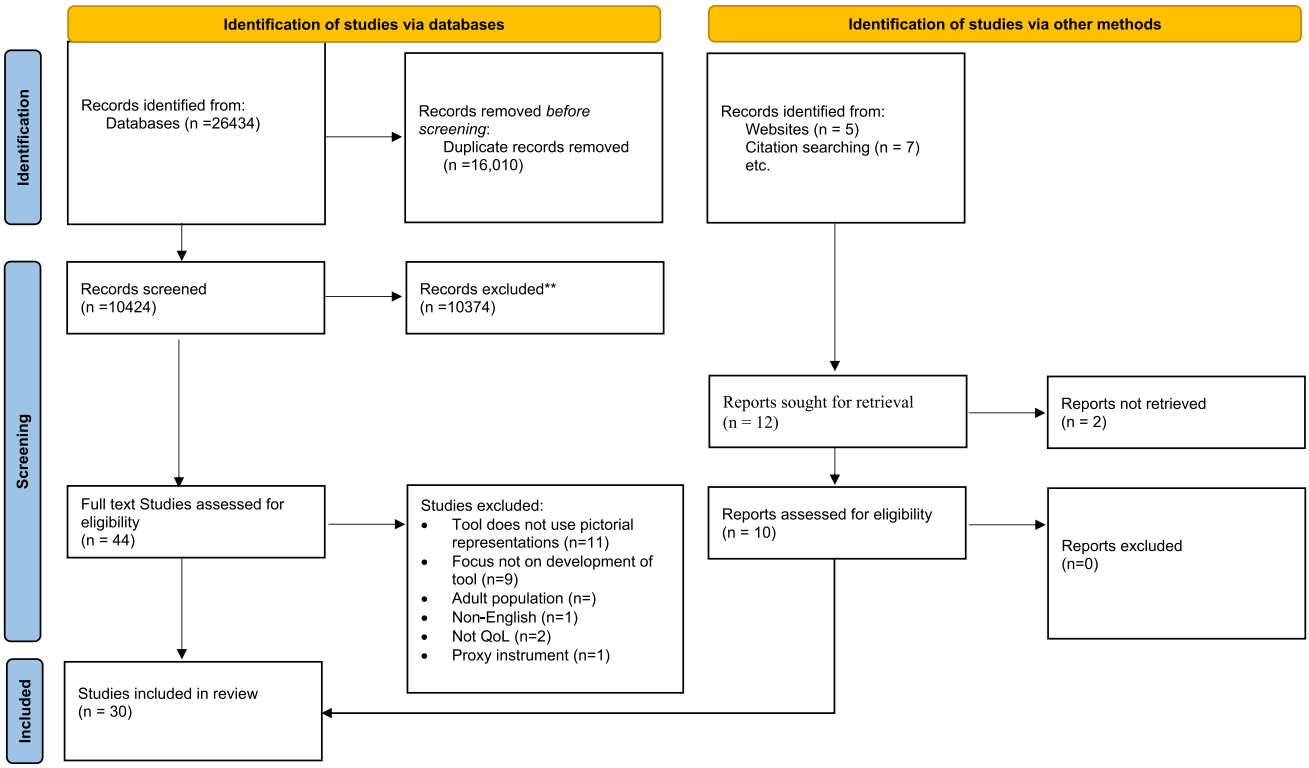

Fig. 1 Literature search flow diagram using the PRISMA checklist. From: Page MJ, McKenzie JE, Bossuyt PM, et al., The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021;372: n71. *PRISMA* Preferred Reporting Items for Systematic Reviews and Meta-Analyses, *QoL* quality of life

or props. Likert scaling typically measures a positive or negative response to a statement with two or more response options. The visual analogue scale (VAS) is a rating scale in a continuous graphical format, usually presented as a horizontal or vertical straight line extending from one end of the scale (e.g. 'strongly agree') to the other (e.g. 'strongly disagree') and generally ranges between 0 and 100. We also examined other key characteristics of childhood PROMs as recommended by ISPOR, including the recall period, length of the PROM (number of items), type of pictorial representations applied (e.g. pictorial, verbal/audio, animation, or props) administration approaches, and availability of electronic PROMs.

For ISPOR PRO Good Research Practices report recommendation 5 (consider cross-cultural issues), countries or jurisdictions of PROM origins, cross-cultural influences, and language translations were considered.

3 Results

3.1 Search Results

Figure 1, the PRISMA flowchart, presents the systematic review process and search results.

A total of 26,434 studies were identified. Of these, 16,010 duplicate articles were removed, resulting in 10,424 studies. A further 12 were identified through web searching via Google Scholar. Titles and abstracts were screened, and 10,374 studies did not meet the inclusion criteria and were removed. The reasons for exclusion were (1) adult population and (2) childhood HRQoL PROMs that do not use pictorial response scales or items. Full texts for 44 studies were assessed independently by the three reviewers. A total of 22 PROMs, including 28 unique versions, were identified in the review. The reviewers' overall agreement (inter-rater reliability) was calculated using Cohen's kappa and yielded a good agreement ($k = 0.72$).

3.2 Overview of Identified PROMs

3.2.1 Origin, Population, and Setting

Tables 1 and 2 provide a detailed summary of the identified generic and condition-specific PROMs, respectively. More measures were developed in the USA ($n = 9$, [32%]) [31–38] and the UK ($n = 8$, [29%]) [39–44] than elsewhere. Four PROMs were from the Netherlands ($n = 4$, [14%]), three from France ($n = 3$, [11%]) [45–47], and one each from Finland [48] and Australia [49]. Two PROMs, the DISABKIDS

Smileys TAKE 6 [50] and the Hvidoere Smiley Faces International diabetes quality of life assessment tool for young children (Hvidoere Smiley Faces) questionnaire [51], were developed as part of multi-country collaborative projects.

All the PROMs were developed for children < 18 years, except for one (Dartmouth Primary Care Cooperative Information Project Charts [COOP Charts]) [33], designed for children 12–18. Study participants were recruited from schools, clinics, hospitals, or the general public.

3.2.2 Description of PROMs

A detailed description of the identified generic and condition-specific PROMs is provided in Tables 1 and 2. We identified 22 PROMs, including 28 unique versions. Two PROMs were versions of the Auto Questionnaire Enfant Image/Child Pictured Self Report (AUQUEI), i.e. AUQUEI Nursery and AUQUEI Primary [46]. Two were different versions of the Pediatric Quality of Life Inventory (PedsQL): PedsQL Generic Core Scales—Young Child (PedsQL GCS—Young Child) [34] and PedsQL Short Form 15—Young Child (PedsQL SF15—Young Child) [35]. Three were different versions of the Dutch Children's AZL/TNO Quality of Life (DUX): the Short Form (DUX-25) [52], the celiac disease version of the DUX-25 (CDDUX) [53], and the Bone Tumour DUX-25 (Bt-DUX) [54]. Another three were different versions of the Childhood Asthma Questionnaire (CAQ): Childhood Asthma Questionnaire-A (CAQ-A), Childhood Asthma Questionnaire-B (CAQ-B), and Childhood Asthma Questionnaire-C (CAQ-C) [42]. A further three PROMs were versions of the Child Health Rating Inventories (CHRIS): CHRIS—general health scales [32], CHRIS—preoperative anxiety, and CHRIS—postoperative pain management [36]. The remaining 15 were single-version measures, presented in Box 1.

Box 1 Single-version pictorial patient-reported outcome measures (PROMs)

#	PROM name and acronym
1	Children's Health States Preferences Learnt from Animation (CHILDSPLA) [39]
2	Seventeen-dimensional measure of health-related quality of life (17D) [48]
3	Child Health and Illness Profile—Child Edition Child Report Form (CHIP-CE CRF) [31]
4	Dartmouth Primary Care Cooperative Information Project Charts (COOP Charts) [33]
5	Exeter Quality of Life (ExqoL) [55]
6	KidIQoL [47]
7	Paediatric Measure Yourself Medical Outcomes Profile (P-MYMOP) [49]
8	Quality of Life Scale for Children (QoL-C) [56]
9	TedQL [44]
10	Animated computer program to assess quality of life in children with inflammatory bowel disease [57]
11	Children's Dermatology Life Quality Index: Cartoon version (CDLQI Cartoon version) [43]
12	DISABKIDS Smileys TAKE 6 [50]
13	Hvidoere Smiley Faces: International diabetes quality of life assessment tool for young children (Hvidoere Smiley Faces) [51]
14	Pediatric Asthma Quality of Life Questionnaire—Pictorial version (PAQLQ—Pictorial version) [37]
15	Cochlear Implant Quality of Life (Cochlear Implants—QOL) [38]

Of the 28 unique PROM versions, two (7%) are generic PROMs accompanied by preference-based value sets (i.e. Child Health Utility 9 Dimension [CHU9D] animation/Children's Health States Preferences Learnt from Animation [CHILDSPLA] and 17-dimensional measure of health-related quality of life [17D]), and the remaining 26 (93%) are not preference-weighted measures.

All the PROMs were developed between 1993 and 2021, with the majority developed after 2000 (n = 20, [71%]). The earliest measures to be developed were the CAQ-A, CAQ-B, and CAQ-C in 1993. The most recently developed are the CHRIS—preoperative anxiety and CHRIS—postoperative pain management (2021).

Fifteen (54%) of the PROMs were generic, 11 (39%) were condition specific, and two (7%) were domain specific. The conditions targeted include asthma (n = 4), celiac disease (n = 1), chronic conditions (n = 1), cochlear implants (n = 1), bone tumour (n = 1), diabetes (n = 1), inflammatory bowel disease (IBD) (n = 1), and skin conditions (n = 1). The domains targeted were preoperative anxiety (n = 1) and postoperative pain (n = 1).

3.3 Criteria for Pictorial PROMs Administration

3.3.1 Target Age Groups

The youngest age for self-reported childhood PROMs with pictorial illustrations available is 3 years (TedQL), while the oldest is 18 years (COOP Charts). Of the 28 PROM versions examined, one measure, AUQUEI Nursery, targets children younger than 5. The PedsQL GCS—Young

Table 1 Summary of characteristics of included generic preference-based and non-preference-based childhood patient-reported outcome measures

#	Acronym: name	References (first author); country	Target age (years)	Respondent type	Generic/ disease-specific	Administration mode	Recall period	Completion time (min)	Domain/dimensions (number, *n*)	Items (number)	Response options and presentation [scoring method]
	Generic preference-based measures										
1.	CHILDSPLA: Children's Health States Preferences Learnt from Animation (computer animation program)	Abrines Jaume [39]; UK	4–14	Child	Generic	Self-report: via a touch screen device /computer	Today/last night	3–5	Sad, annoyed, worried, pain, tired, schoolwork, daily routine, able to join in activities ($n = 9$)	9	5-level pictorial Likert scale [value set available]
2.	17D: Seventeen-dimensional measure of health-related quality of life	Apajasalo [48]; Finland	8–11	Child; observer	Generic	Self-report (proxy administration can be used for younger children)	Today	20–30	Mobility, vision, hearing, breathing, sleeping; eating, speech, excretion, school and hobbies, mental function, discomfort and symptoms, depression, distress, vitality, appearance, friends, concentration ($n = 17$)	17	5-level Likert scale illustrated with pictures representing each domain [value set available]
	Generic non-preference-based measures										
3.	AUQUEI Nursery: Auto Questionnaire Enfant Image/ Child Pictured Self Report—Nursery	Manificat [45, 46]; France	3–5	Child	Generic	Self-report aided by parents	Not stated	10–15	Family life; social life; children's activities [school and leisure]; health ($n = 4$)	26	4-level pictorial happy–sad faces (26 items); open-ended (2 items) [item average]
4.	AUQUEI Primary: Auto Questionnaire Enfant Image/ Child Pictured Self Report—Primary	Manificat [45, 46]; France	6–11	Child	Generic	Self-report	Not stated	10	Family life; social life; activities at school and leisure; health ($n = 4$)	31	4-level pictorial happy–sad faces (31 items); open-ended (2 items) [item average]

Table 1 (continued)

#	Acronym: name	References (first author); country	Target age (years)	Respondent type	Generic/ disease-specific	Administration mode	Recall period	Completion time (min)	Domain/dimensions (number, *n*)	Items (number)	Response options and presentation [scoring method]
5.	CHIP-CE CRF: Child Health and Illness Profile—Child Edition Child Report Form	Riley [31]; USA	6–11	Child	Generic	Self-report	Past 4 weeks	20	Satisfaction; comfort; resilience; risk avoidance; achievement ($n = 5$)	45	5-level response graduated circles; cartons at beginning and end [item average]
6.	CHRIS: Child Health Rating Inventories—general health module	Kaplan [36]; USA	5–12	Child	Generic	Self-report and self-administered via computer		7	Physical health, mental health ($n = 2$)	20	5-point scale anchored with cartoons [item average]
7.	COOP Charts: Dartmouth Primary Care Cooperative Information Project Charts	Wasson [33]; USA	12–21	Child	Generic	self-report or clinician administered	Past 4 weeks	5	Physical and social functioning; symptoms and emotional feelings; perceptions; and social support ($n = 4$)	9	5-point scale with cartoons and written response options (picture and word chart) [item average]
8.	DUX-25: Dutch Children's AZL/TNO Quality of Life—Short Form	Koopman [52]; Netherlands	5–16	Child	Generic	Self-report	Current health status	-	Body, emotion, social, home ($n = 4$)	25	5-point Likert (smiling to crying) faces scale with expressions from happy to sad [item average]
9.	ExqoL: Exeter Quality of Life	Eiser [55]; UK	6–12	Child	Generic	Interviewer assisted; computer based	Not stated	20	Symptoms; social well-being; school achievements; physical activity; worry; family relationships ($n = 6$)	12	12 pictures, each rated twice using VAS—(1) first in terms of actual self ('like me') (2) in terms of ideal self 'I would like to be' [item average]

Table 1 (continued)

#	Acronym: name	References (first author); country	Target age (years)	Respondent type	Generic/ disease-specific	Administration mode	Recall period	Completion time (min)	Domain/dimensions (number, *n*)	Items (number)	Response options and presentation [scoring method]
10.	KidIQoL	Gayral-Taminh [47]; France	6–12	Child	Generic	Self-reported. Administered via computer		20	Physical and psychological health, family life, school life and social and physical environment ($n = 4$)	44	5-level scale with pictures [item average]
11	PedsQL GCS—— Young Child: Pediatric Quality of Life Inventory, Generic Core Scales—Young Child	Varni [34]; USA	5–7	Child; observer	Generic	Self-report or interviewer-administered (telephone)	Past month (past week for acute version)	–	Functioning: physical; emotional; social; school ($n = 4$)	23	3-level pictorial faces, Informant: 5-point scale [item average per domain]
12	PedsQL SF15—Young Child: Pediatric Quality of Life Inventory, Short Form 15—Young Child	Chan [35]; USA	5–7	Child: observer	Generic	Self-report or interviewer administered (telephone)	Past month (past week of acute version)	–	Functioning: physical; emotional; social; school ($n = 4$)	15	3-level pictorial faces, informant: 5-point scale [item average per domain]
13.	P-MYMOP: Paediatric Measure Yourself Medical Outcomes Profile	Ishaque [49]; Australia	7–11	Child, aided by the parent	Generic	Self-report	Past week	–	Symptom 1; symptom 2; activity impairment; general well-being (child specifies symptoms and activity impairment affecting her most) ($n = 4$)	4	7-point pictorial faces [item average]

Table 1 (continued)

#	Acronym: name	References (first author); country	Target age (years)	Respondent type	Generic/ disease-specific	Administration mode	Recall period	Completion time (min)	Domain/dimensions (number, *n*)	Items (number)	Response options and presentation [scoring method]
14.	QoL-C: Quality of Life Scale for Children	Thompson [41]; UK	4–9	Child	Generic	Age 4–7: interviewer assisted. Age 8–9: self-report, informant: by self; mail	Today	7	Moving; looking after myself; doing usual activities; having pain; feeling worried, sad, or unhappy; global health VAS ($n = 5 +$ VAS)	5	3-level happy–sad face scale and a horizontal health meter from ranging from 0 to 10 [item sum]
15.	TedQL	Lawford [44]; UK	3–8	Child	Generic	Interviewer assisted	Present health		Physical competence; peer acceptance; maternal acceptance; psychological functioning; cognitive functioning ($n = 5$)	23	Two 4-level scales: colour circles and linear scale ranging from really bad/dislike to really good/like. This was with the aid of props (2 teddy bears presented by the interviewer)

VAS visual analogue scale

Table 2 Summary of characteristics of included condition-specific non-preference-based childhood patient-reported outcome measures

#	Acronym: name	References (first author); country	Target age (years)	Respondent type	Generic/ disease-specific	Administration mode	Recall period	Completion time (min)	Domains (number, *n*)	Items (number)	Response options [scoring method]
	Disease specific										
1.	Animated computer program to assess quality of life in children with inflammatory bowel disease	Buller [57]; Netherlands	5–12	Child	Inflammatory bowel disease	Animated computer program (with touch screen or mouse)	Not stated	Not stated	Not stated	35	4-level pictorial scale supported by a bear, clown, faces and hand illustrations
2.	CAQ-A: Childhood Asthma Questionnaire—A	Christie [42]; UK	4–7	Child	Asthma	Self-report aided by parents	Present health	Not stated	Quality of life, distress ($n = 2$)	14	4-level smiley face scale [item average]
3.	CAQ-B: Childhood Asthma Questionnaire—B	Christie [42]; UK	8–11	Child	Asthma	Self-report aided by parents	Present health	Not stated	Active quality of life, passive quality of life, distress, severity ($n = 4$)	23	5-level smiley faces scale [item average]
4.	CAQ-C: Childhood Asthma Questionnaire—C	Christie [42]; UK	12–16	Child	Asthma	Self-report aided by parents	Present health	Not stated	Active quality of life, teenage quality of life, distress, severity, reactivity ($n = 5$)	41	5-level smiley face scale [item average]
5.	CDLQI: Children's Dermatology Life Quality Index: *Cartoon version*	Holme [43]; UK	4–16	Child	Skin conditions	Self-report	Past Week	Not stated	Symptoms, embarrassment, friendships, clothes, playing, sports, school, bullying, sleep, and impact of treatment ($n = 10$)	10	4-level response scale (checkboxes) to each question illustrated by a cartoon dog in colour [summed total score]

Table 2 (continued)

#	Acronym: name	References (first author); country	Target age (years)	Respondent type	Generic/ disease-specific	Administration mode	Recall period	Completion time (min)	Domains (number, *n*)	Items (number)	Response options [scoring method]
6.	CDDUX: Celiac Disease Dutch Children's AZL/TNO Quality of Life—Short Form	van Doorn [53]; Netherlands	8–18	Child	Celiac disease	Self-report	General	Not stated	Communication, diet, having celiac disease ($n = 3$)	12	5-point Likert (smiling to crying) faces scale with expressions from happy to sad [item average]
7.	CHRIS: Child Health Rating Inventories—*preoperative anxiety*	Kaplan [36]; USA	4–12	Child	Preoperative anxiety	Self-report	–	Not stated	Preoperative anxiety ($n = 1$)	5	5-point scale anchored with cartoons [item average]
8.	CHRIS: Child Health Rating Inventories—*postoperative pain management*	Kaplan [36]; USA	4–12	Child	postoperative pain management	Self-report	–	Not stated	Postoperative pain management ($n = 1$)	3	5-point scale anchored with cartoons [item average]
9.	DISABKIDS Smileys TAKE 6	Chaplin [50, 58]; multiple countries ($n = 6$)	4–7	Child	Chronic conditions	Interviewer-assisted report	Not stated	Not stated	Mental, social, physical ($n = 3$)	6	5-level smiley face scale ranging from very sad to very happy.
10.	Hvidoere Smiley Faces: Hvidoere Smiley Faces International diabetes quality of life assessment tool for young children	Hoey [51]; multiple countries ($n = 17$)	4–11	Child	Diabetes	Interviewer-assisted report	Not stated		Disease state and physical symptoms, functional status, psychological and social functioning ($n = 4$)	10	5-level happy–sad face response scale ranging from very happy to very unhappy
11.	PAQLQ: Pediatric Asthma Quality of Life Questionnaire—*Pictorial version*	Everhart [37]; USA	5–7	Child	Asthma	Interviewer-assisted	Past week		Physical, emotional, and social impairment due to asthma ($n = 4$)	16	3-level thermometer response scale [item average]

Table 2 (continued)

#	Acronym: name	References (first author); country	Target age (years)	Respondent type	Generic/ disease-specific	Administration mode	Recall period	Completion time (min)	Domains (number, *n*)	Items (number)	Response options [scoring method]
12.	Cochlear Implants—QOL: Cochlear Implant Quality of Life	Hoffman [38]; USA	6–12	Child	Cochlear implants	Self-report. Administered electronically via tablet device	One week	10	Noisy environments, academic functioning, child acceptance, oral communication, social functioning, fatigue, emotional functioning, and device management ($n = 8$)	33	4-point scale; items and response choices are presented in three modalities: written, auditory and pictorial forms

Child, PedsQL SF15—Young Child, and the Pediatric Asthma Quality of Life Questionnaire (PAQLQ)—Pictorial version target children aged 5–7 years, while three measures (17D, Paediatric Measure Yourself Medical Outcomes Profile [P-MYMOP], and CAQ-B) are intended for children 8–11 years. Measures explicitly designed for adolescents (12- to 18-year-olds) include the COOP Charts and the CAQ-C. Seventeen PROMs covered multiple age groups: the AUQUEI Primary, TedQL, Child Health and Illness Profile—Child Edition Child Report Form (CHIP-CE CRF), CHRIS, DUX-25, Exeter Quality of Life (ExqoL), KidIQoL, Quality of Life Scale for Children (QoL-C), the animated computer program for children with IBD, Bt-DUX, CAQ-A, Children's Dermatology Life Quality Index (CDLQI), CHRIS—preoperative anxiety, CHRIS—postoperative pain management, DISABKIDS Smileys TAKE 6, and the Hvidoere Smiley Faces measure.

3.4 Development and Content Validity of Paediatric Pictorial PROMs

Table 3 reports on the development process of paediatric pictorial PROMs, highlighting the involvement of children in the process, the graphic approaches employed, and the pivotal role of pilot testing.

3.4.1 Content Generation Methods

In the development of the identified paediatric pictorial PROMs, various methods were used, including literature reviews, expert consultation, focus groups, and interviews. The breakdown of these approaches among the PROMs is as follows: almost half (43%) of the identified PROMs (12 out of 28) utilised literature reviews in their development process. Similarly, 43% of the PROMs (12 out of 28) incorporated focus groups during development. Expert consultation played a role in developing approximately one-third (32%) of the PROMs (9 out of 28). Interviews were the most commonly employed method, with 17 out of 28 PROMs (32%) using them as part of their content and item generation process. These approaches highlight the diverse methodologies employed in the development of pictorial paediatric PROMs, contributing to their depth and comprehensiveness.

3.4.2 Children's Involvement in the Development of Pictorial Elements

Thirteen studies emphasised the active participation of children in crafting the pictorial elements for HRQoL measures. PROMs in this category include CHU9D animation/

Table 3. Methodological approaches used for developing pictorial measures of quality of life

PROM	Children's input in developing pictorial elements	Graphic approaches for domain/item development	Children's participation in pilot testing pictorial elements	Handling recall periods graphically
CHILDSPLA Abrines Jaume [39]	Adapted from CHU9D Collaborative character design with child psychologist, animation filmmaker, and 38 school children (4–11 years) and 36 children from a specialist hospital Interactive group sessions with varying durations Children actively participated in 15 interactive group sessions Exploration of character attributes and feedback collection, including preferences Incorporation of engaging activities for feedback and character suitability	Universal character used Used various graphic methods to represent different domains of HRQoL Made adjustments based on children's feedback for different emotions: Sadness: Adjusted for a more relaxed character Annoyed: Clear animations well-received Worried: Adjustments due to misunderstanding; children often used sentences Pain: Focused on abdominal pain; interpretations varied Tired: Easily understood; no changes needed School Work: Quickly grasped; added a pen in animation Sleep: Represented sleep problems with decreasing sleep time Daily Routine: Final version included three activities, with declining character ability Able to Join in Activities: Adjusted for character's condition	Pilot test with 389 children (ages 4–14)	Details not provided
17D Apajasalo [48]	Adapted from 15D and 16D The study does not specify whether children were involved in the development phase of the pictorial elements for the 17D	A variety of graphic methods were used to represent the different domains of the 17D. For example, they used cartoons doing different activities to represent the 'School and hobbies' domain, and they used pictures of different facial expressions to represent the 'Emotional well-being' domain	Pilot study conducted with 79 healthy 7- to 10-year-olds	Details not provided
AUQUEI Nursery Manificat [45, 46]	Not stated	Not specified	Not stated	Details not provided
AUQUEI Primary Manificat [45, 46]	Not stated	Not specified	Not stated	Details not provided

Table 3. (continued)

PROM	Children's input in developing pictorial elements	Graphic approaches for domain/item development	Children's participation in pilot testing pictorial elements	Handling recall periods graphically
CHIP-CE CRF Riley [31]	Adapted from CHIP-AE Not stated whether children were involved in developing pictorial elements	Illustrations of a 'universal character' were used to anchor the response scale, created by a professional cartoonist Sixteen simple items representing common tasks or activities were developed	Children were engaged in assessing and providing feedback on the character used in the study Children as young as 5 were able to use the response scale effectively There was consistency in their responses between horizontal and vertical presentations, with some exceptions related to specific items Children were concrete in their identification with the character and provided feedback on details like the number of fingers on the character	Details not provided
CHRIS—general health module Kaplan [32, 36]	Not stated	Not specified	Not stated	Details not provided
COOP Charts Wasson [33]	Not stated	Not specified	Not stated	Details not provided
DUX-25 Koopman [52]	Adapted from DUCATQOL Not stated whether children were involved in developing pictorial elements	Used abstract faces with varying expressions or smileys for scoring items	Not stated	Details not provided
ExqoL Eiser [55]	Not stated	Not specified	Not stated	Details not provided
KidIQoL Gayral-Taminh [47]	Children participated in developing KidIQoL's pictorial elements Used ‘Smiley’ figurines to express responses Adjusted images based on children's feedback for clarity	Not specified	Children actively involved in testing and providing feedback on prototype tool Children’s feedback, particularly on images used in tool, was considered Children's comments crucial in adjusting items representation	Details not provided
PedsQL GCS—Young Child Varni [34]	Developed through focus groups and cognitive interviews (participants not specified) Possible involvement of children in focus groups and cognitive interviews	Not specified	Not stated	Details not provided
PedsQL SF15—Young Child Chan [35]	Short form derived from PedsQL GCS using quantitative techniques	Not specified	Not stated	Details not provided

Table 3. (continued)

PROM	Children's input in developing pictorial elements	Graphic approaches for domain/item development	Children's participation in pilot testing pictorial elements	Handling recall periods graphically
P-MYMOP Ishaque [49];	Adapted from adult MYMOP 24 children (aged 7–11) from local tertiary hospital involved in development of P-MYMOP's pictorial elements The children offered feedback on clarity and understandability of pictorial elements	Used face scales as graphics Faces scale chosen due to child-friendliness and ease of use compared to numeric rating scales, promoting self-completion Both children and their parents/guardians viewed the inclusion of a faces scale on the tool positively	A pilot feasibility study with 29 children aged 7–11 and their parents at tertiary hospital Examined practical aspects of implementing the P-MYMOP in a tertiary-care hospital's outpatient clinic P-MYMOP was accepted by clinicians, including nurses and doctors	Details not provided
QoL-C Thompson [41]	Adapted from EQ-5D Children (aged 4–9) were consulted during its adaptation of wording and layout Amended to make the measure child-friendly and relevant Developed by University of Exeter Medical School with the EuroQoL Group	QoL-C uses simple language and age-appropriate wording, such as 'moving around' instead of 'mobility' Employs horizontal visual scale resembling number lines, for children to rate their general health Emoticons used to represent response options	Not stated	Details not provided
TedQL Lawford [44]	TedQL developed based on the results of two preliminary studies with children In study 2 (TedQL.2), children were involved in developing pictorial elements for the measure	Various graphic methods used based on different domains Employed teddy bears and diverse response scales Tailored graphic methods to address specific content and challenges for young children	Children involved in pilot testing pictorial elements in TedQL.2 study Actively participated to develop and refine the measure Used teddy bears and various graphic methods to assess aspects of their lives in QoL domains	Details not provided
Animated computer program to assess quality of life in children with inflammatory bowel disease Buller [57]	No mention of children's involvement in program development Authors collaborated with artists to create program for children aged 5 and above Story-based with bear and clown, 35 questions	The program included a story of a bear and a clown with interactive elements, making use of triggers, and used faces and hands to express moods No graphic methods mentioned	Pilot study with 16 children (5–12 years) Effectively used by children in this age group Reduced interest in children older than 11, requiring an alternative approach	Details not provided
Bt-DUX [54]	No mention of children's involvement in developing pictorial elements Developed from DUX-25 10 items directly taken from DUX-25 Items generated from interviews with 10 patients (aged 10–19) and four healthcare experts	Not specified	Not stated	Details not provided
CAQ-A Christie [42]	Children involved in developing CAQ-A questionnaire's pictorial elements, suggesting 'Smiley' faces for expressing feelings about activities	Study does not mention various graphic methods for different domains; primarily uses 'Smiley' faces	Not stated.	Details not provided

Table 3. (continued)

PROM	Children's input in developing pictorial elements	Graphic approaches for domain/item development	Children's participation in pilot testing pictorial elements	Handling recall periods graphically
CAQ-B Christie [42]	Children involved in developing CAQ-B questionnaire's pictorial elements, suggesting 'Smiley' faces for expressing feelings about activities	Study does not mention various graphic methods for different domains; primarily uses 'Smiley' faces	Not stated.	Details not provided
CAQ-C Christie [42]	Children involved in developing CAQ-C questionnaire's pictorial elements, suggesting 'Smiley' faces for expressing feelings about activities	Study does not mention various graphic methods for different domains; primarily uses 'Smiley' faces	Not stated.	Details not provided
CDLQI Cartoon version Holme [43]	Adapted from CDLQI Developed by authors No children's involvement	A dog as universal cartoon character was chosen to avoid gender or age issues Japanese version uses cartoon children as dog motif is inappropriate	CDLQI Cartoon version pilot tested in children Three pilot testing studies conducted: Study 1: 101 children (aged 4–16) Study 2: 107 children Study 3: 546 children Validation against CDLQI Preferred by most parents and children Found easy to use	Details not provided
CDDUX van Doorn [53]	Not stated	Not stated	Not stated	Details not provided
CHRIS—preoperative anxiety Kaplan [36]	Children involved in focus groups to provide input on item content for measures Input from children helped shape item content	Artists and animators were included in children's focus groups Visual elements, potentially including graphics or animations, were used to engage children in providing feedback	Children and parents reviewed final animated versions of the measures during a testing phase	Details not provided
CHRIS—postoperative pain management Kaplan [36]	Children involved in focus groups to provide input on item content for measure Input from children helped shape item content	Artists and animators were included in children's focus groups Visual elements, potentially including graphics or animations, were used to engage children in providing feedback	Children and parents reviewed final animated versions of the measures during a testing phase	Details not provided
DISABKIDS Smileys TAKE 6 Chaplin [50, 58]	Discussions conducted with children to ensure question formulation Unclear if children were involved in design of pictorial element (smiley faces)	Not stated	Not stated	Details not provided
Hvidoere Smiley Faces Hoey [51]	Not stated	Not stated	Not stated	Details not provided
PAQLQ—Pictorial version Everhart [37]	Adapted from the PAQLQ Children aged 5–7 participated in pictorial development Used thermometers for child-friendly responses	Various graphic approaches for content domains Thermometers for bother levels, from 'not at all bothered' to 'extremely bothered'	Pilot testing with children aged 5–7 Feedback led to removal of seven problematic items	Details not provided

Table 3. (continued)

PROM	Children's input in developing pictorial elements	Graphic approaches for domain/item development	Children's participation in pilot testing pictorial elements	Handling recall periods graphically
Cochlear Implants—QOL Hoffman [38]	Children were involved in developing the pictorial elements through their participation in the cognitive testing phase of the instrument development	The study employed a multimodal approach with various graphic and audio methods. Pictorial representations were used to visually represent items and response options	While the study did not mention a separate pilot testing phase for pictorial elements, children were involved in the cognitive testing phase, which likely included feedback on the pictorial representations used in the child self-report version	Details not provided

17D 17-dimensional measure of health-related quality of life, *AUQUEI* Auto Questionnaire Enfant Image/Child Pictured Self Report, *Bt-DUX* Bone Tumour DUX-25, *CAQ* Childhood Asthma Questionnaire, *CDDUX* Celiac Disease DUX-25, *CDLQI* Children's Dermatology Life Quality Index, *CHILDSPLA* Children's Health States Preferences Learnt from Animation (computer animation program), *CHIP-CE CRF* Child Health and Illness Profile-Child Edition Child Report Form, *CHRIS* Child Health Rating Inventories, *CHU9D* Child Health Utility 9 Dimension, *Cochlear Implants—QOL* Cochlear Implant Quality of Life, *COOP Charts* Dartmouth Primary Care Cooperative Information Project Charts, *DUX-25* Dutch Children's AZL/TNO Quality of Life—Short Form, *ExqoL* Exeter Quality of Life, *HRQoL* health-related quality of life, *Hvidoere Smiley Faces* Hvidoere Smiley Faces International diabetes quality of life assessment tool for young children, *PAQLQ* Pediatric Asthma Quality of Life Questionnaire, *PedsQL GCS* Pediatric Quality of Life Inventory, Generic Core Scales, *PedsQL SF15* Pediatric Quality of Life Inventory, Short Form 15, *P-MYMOP* Paediatric Measure Yourself Medical Outcomes Profile, *PROM* patient-reported outcome measure, *QoL* quality of life, *QoL-C* Quality of Life Scale for Children

CHILDSPLA, KidIQoL, P-MYMOP, QoL-C, TedQL, PAQLQ—Pictorial version, Cochlear Implant Quality of Life (Cochlear Implants—QOL), CHRIS—postoperative pain management, CAQ-A, CAQ-B, CAQ-C, and DISABKIDS Smileys TAKE 6. This collaborative approach aimed to create instruments that are more child-friendly, engaging, and relevant to the unique experiences of children.

In contrast, 15 PROMs, including 17D, AUQUEI Nursery, AUQUEI Primary, CHIP-CE CRF, COOP Charts, DUX-25, ExqoL, PedsQL GCS—Young Child, PedsQL SF15—Young Child, CDLQI Cartoon version, CDDUX, Hvidoere Smiley Faces, Bt-DUX, CDDUX, and the animated computer program for children with IBD, did not explicitly mention the participation of children in the design of pictorial versions.

3.4.3 Diverse Graphic Methods for Varied Content Domains

Various graphic methods were employed depending on the content of different domains in the development of paediatric pictorial PROMs. This approach was taken to address the potential challenges in visual representation and ensure that the instruments could effectively capture children's experiences in diverse aspects of HRQoL.

For example, in domains related to emotional well-being and mood, some PROMs, such as CHU9D animation/CHILDSPLA, CAQ-A, CAQ-B, and CAQ-C, used 'smiley' faces or emoticons with different expressions (such as happy, sad, or neutral) to help children express their feelings. This graphic method made it easier for children to convey their emotional states.

In contrast, for domains related to physical health or daily activities, different graphic approaches were used. Some instruments employed interactive characters, animations, or cartoons to represent various aspects of a child's life, making it more engaging and relatable for them. For instance, the CHU9D animation/CHILDSPLA PROM utilised interactive group sessions with children and an animation filmmaker to create engaging characters that children could relate to.

Additionally, in domains where the assessment required quantifying bother levels or responses to specific items, thermometers or visual scales were used. These visual tools allowed children to indicate the extent to which they were bothered by certain issues, helping in a more precise assessment. For instance, the PAQLQ—Pictorial version PROM used thermometers to gauge bother levels, providing children with a clear and visual way to express their discomfort.

By adapting graphic methods to suit the specific content domains, instrument developers aimed to create measures that were child-friendly, comprehensible, and relevant. This approach recognised the importance of considering the unique challenges associated with visual representation when assessing HRQoL in children.

3.4.4 Pilot Testing

Pilot testing was found to be an important component of the development phase of 11 of the identified PROMs, playing a vital role in refining the measures. For example, in the case of KidIQoL, children actively took part in the testing process and provided feedback on the instrument, leading to refinements in item representation. Similarly, in the case of P-MYMOP, a pilot feasibility study involving children aged 7–11 was conducted to identify and address challenges, ensuring alignment with children's comprehension and preferences. This thorough pilot testing process ensured that the measures effectively captured children's experiences.

3.4.5 Adaptation from Existing Measures

Six PROMs were adapted from existing ones: the CHU9D animation/CHILDSPLA (adapted from the CHU9D); the CDLQI Cartoon version (adapted from the CDLQI); the CHIP-CE CRF (adapted from the CHIP CE Parent Report Form [CHIP-CE/PRF]); the QoL-C (adapted from the EQ-5D); the PAQLQ—Pictorial version (adapted from the PAQLQ—Established version); and the P-MYMOP (adapted from the adult MYMOP). Adjustments were made to item and response options and the inclusion of pictures, cartoons, photographs, happy–sad faces, and illustrations to enhance children's understanding. Preferences of children for certain images were observed, emphasising the importance of visual elements in paediatric HRQoL assessment.

3.5 Design and Format of the PROM

3.5.1 Response Scales and the Use of Pictorial Representations

Tables 1 and 2 show that PROMs' response options and presentation formats were divided into Likert and VAS response scales or a combination of the two. Both scale types were supported by pictorial representations and are thus referred to as pictorial Likert or pictorial VAS hereafter. Almost all the identified PROMs ($n = 26$, [93%]) used a pictorial Likert scale to present response options. The pictorial Likert scales ranged from a 3-point scale (PedsQL GCS—Young Child, PedsQL SF15—Young Child, QoL-C, PAQLQ) to a 7-point scale (P-MYMOP). Most PROMs ($n = 17$, [61.0%]) used 4- or 5-point scales. Two PROMs employed the VAS to present response choices

Fig. 2 Example of a response scale using cartoons (CHU9D animation/CHILDSPLA). The example depicts 5 levels of the sad dimension CHILDSPLA computer animation program, from left to right, 1—'I don't feel sad', 2—'I feel a little bit sad', 3—'I feel a bit sad', 4—'I feel quite sad', and 5—'I feel very sad'. Reprinted from Abrines Jaume, N., et al., CHILDSPLA: a collaboration between children and researchers to design and animate health states. *Child: Care, Health & Development*, 2015. 41(6): p. 1140–51, with permission from John Wiley &Sons. *CHU9D* Child Health Utility 9 Dimension, *CHILDSPLA* Children's Health States Preferences Learnt from Animation

[55, 56], i.e. the ExqoL, whose VAS ranged from 'not like me' to 'exactly like me' [55]. Another PROM, the QoL-C, used a combination of the pictorial Likert scale and the VAS (a horizontal child-friendly health meter ranging from 0 to 10 anchored with happy–sad faces at the extreme ends and in the middle) [41].

Five pictorial presentation styles were identified, with the majority using the happy–sad faces: happy–sad faces (15), cartoons/pictures (9), graduated circles (1), lines (1), thermometers (1), and props (1). For example, the CHU9D animation/CHILDSPLA used a cartoon character, which was then animated and presented in an interactive application on a touchscreen device (Fig. 2). The CHIP-CE used two illustrations depicting each question's two end extreme health states. Each question had five possible response circles, which gradually increased to show the increasing severity as response options advance from 'never' to 'always' (Fig. 3). In another example, the TedQL used a response scale with circles and lines along with props (two teddy bears) presented by the interviewer to the child (Fig. 4) [44]. Finally, the PAQLQ—Pictorial version employed a response scale using thermometers (Fig. 5).

3.5.2 Recall Period

The ISPOR task force report (2013) [12] suggests that shorter recall periods of 24 h or less are favourable for paediatric PROMs used within a regulatory context. A recent systematic review on the design of childhood PROMs in preliminary studies recommends a recall period of 48 h or

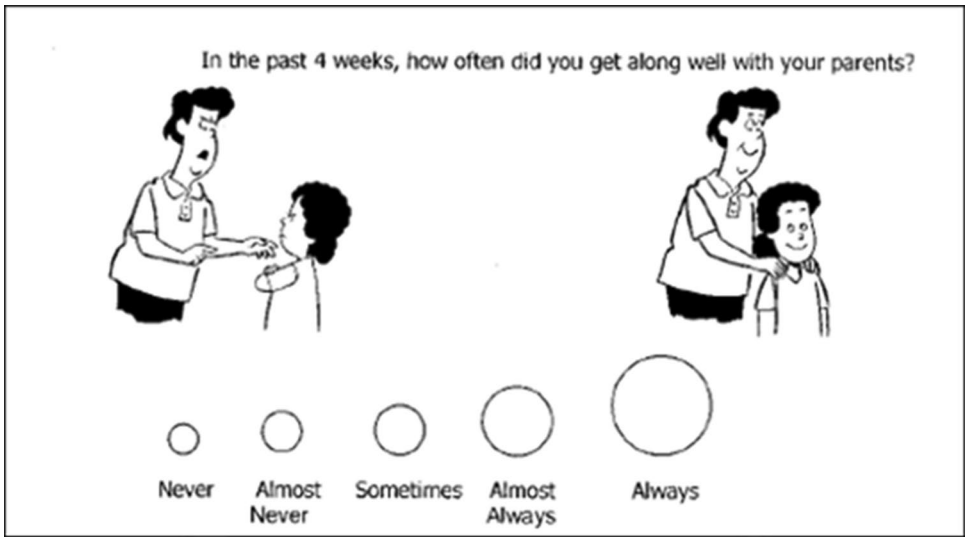

Fig. 3 Example of a response scale using illustrations and graduated circles (CHIP-CE). Reprinted from Riley, A.W., et al., The child report form of the CHIP-child edition: reliability and validity. *Medical care*, 2004: p. 221–231, with permission from Lippincott Williams & Wilkins. *CHIP-CE* Child Health and Illness Profile-Child Edition

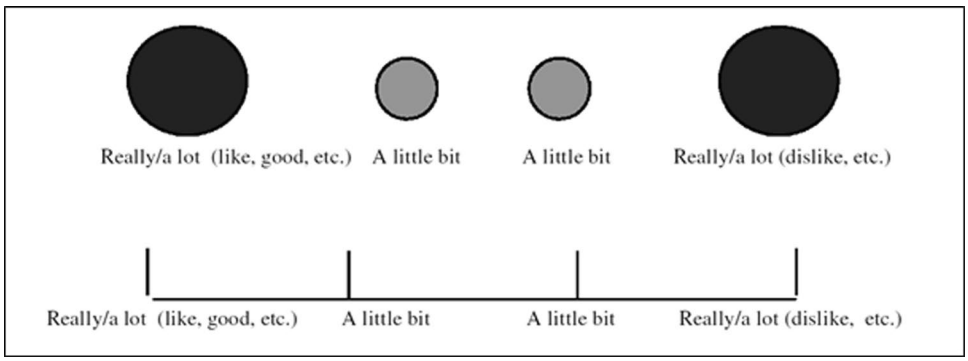

Fig. 4 Example of a response scale using circles and lines (TedQL). Reprinted from Lawford, J., Volavka N., and Eiser C., A generic measure of Quality of Life for children aged 3–8 years: results of two preliminary studies. *Pediatric Rehabilitation*, 2001. 4(4): p. 197–207, with permission from Taylor & Francis Ltd.

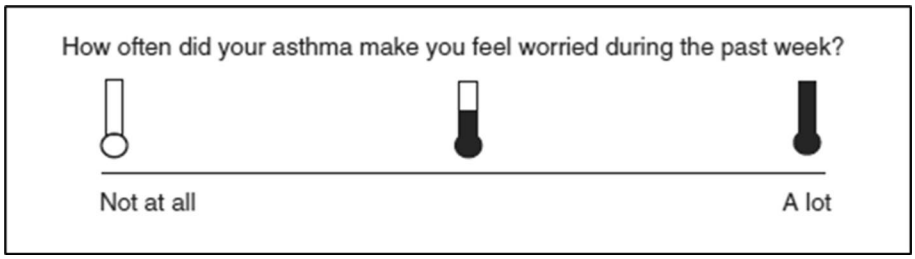

Fig. 5 Example of a response scale using thermometers (PAQLQ—Pictorial version). Reprinted from Everhart, R.S. and B.H. Fiese, Development and initial validation of a pictorial quality of life measure for young children with asthma. *J Pediatric Psychology*, 2009. 34(9): p. 966–76, with permission from Oxford University Press. *PAQLQ* Pediatric Asthma Quality of Life Questionnaire

less for children under 8 years old and 14 days and under for children above 8 [24]. Eight [29%] of the PROMs in this review had a recall period of today or current health: the CAQ-A, CAQ-B, CAQ-C, CHU9D animation/CHILD-SPLA, DUX-25, TedQL, 17D, and QoL-C. Three [11%] measures, ExqoL, KidIQoL, and CDDUX, used a general recall period (did not specify a recall period). Another eight [29%] PROMs had recall periods ranging from the past week (CDLQI, P-MYMOP, PAQLQ—Pictorial version, Cochlear Implants—QOL) to the past month (CHIP-CE CRF, COOP Charts, PedsQL GCS, PedsQL SF15). The remaining nine PROMs [32%] did not state the recall period.

The available information for the paediatric pictorial PROMs does not offer specific details about how recall periods were dealt with graphically in these studies.

3.5.3 Mode of Administration

In terms of mode of administration, 71% of paediatric pictorial PROMs were designed for self-administration, while 29% were intended for interviewer-assisted administration using electronic and non-electronic data collection methods. Among these, seven PROMs (25%) used electronic data collection methods.

In terms of incorporating pictorial representations in the identified PROMs, there is considerable variability among the studies. The CHU9D animation/CHILDSPLA study notably utilised pictorial versions in multiple modes, including animated presentations, touch-screen interfaces with still images, and hard-copy questionnaires. Conversely, the KidIQoL study did not explicitly specify the use of pictorial versions in either administration method.

Other studies adopted unique approaches. The P-MYMOP study integrated faces scales within interviews, while the QoL-C study employed diverse administration modes, such as one-on-one interviews, classroom activities, and distribution through parents. The TedQL study conducted individual interviews, while the animated computer program for children with IBD employed computer-based animations.

Some studies employed VAS, as demonstrated in the Bt-DUX study. However, others like CAQ-A, CAQ-B, and CAQ-C did not provide explicit information regarding the use of pictorial versions in administration. The CDLQI Cartoon version study lacked specific details on administration methods, and the CDDUX study did not address this aspect at all.

Furthermore, the PAQLQ—Pictorial version study did not specify the mode of administration for pictorial versions, and the Cochlear Implants—QOL study utilised computer-administered methods without providing detailed specifics. Finally, the DISABKIDS Smileys TAKE 6 study did not mention the mode of administration for pictorial versions.

In summary, the incorporation of pictorial versions into the mode of administration for PROMs varied significantly across studies, with some employing diverse methods, while others either did not specify or lacked detailed information on this aspect.

3.5.4 Length of PROM

All the measures contained multiple items, ranging from four (P-MYMOP) to 45 for the CHIP-CE CRF. The estimated time to complete the PROMs ranged from 3 to 5 min (CHU9D animation/CHILDSPLA) to 20–30 min (17D)[1]. Comparing the number of items across versions of the measures showed slightly more items in the versions of the PROMs for older children than those for younger children. For instance, the AUQUEI Nursery version for ages 3–5 years has 26 items, whilst the AUQUEI Primary version for ages 6–11 has 31 items (a difference of 16%). Similarly, the CAQ-A for ages 4–7 years has 14 items, the CAQ-B for ages 8–11 years has 23 items, and the CAQ-C for 12- to 16-year-olds has 41 items.

3.5.5 Scoring of PROMs

The PROMs were scored in one of three main ways: (1) multi-item PROMs only yielding a summary score; (2) multi-item, multidomain scales producing both domain-specific and summary score; and (3) preference-weighted measures that use a value set based on stated preferences for the health states. The scores for the former two are typically determined directly through summary scoring of responses to individual items in the measure.

3.5.6 Domains

Table 4 outlines the domains covered by the PROMs that employ pictorial representations. We combined similar constructs of PROMs under common domain definitions developed for this review. The final ten PROM domains were physical health, social functioning, mental health, emotional health, school, environment, autonomy, pain, disease-specific concepts, and overall QoL.

Physical health refers to PROM items that examine physical functioning status, activity impairment, physical conditions, and related symptoms. *Social functioning* refers to items that address the ability to engage in social interactions. Items assessing acceptance, friendship, family life and relationships, and social support were also classified under this domain. *Mental health* refers to items that capture mental and cognitive health conditions. *Emotional health* refers to items describing one's experience, perception, and expression of emotions. It includes items such as sadness, worry, and distress. *School* has facets that address school attendance and participation in schoolwork and activities. *Environment* covers different aspects of the immediate environmental living conditions. *Autonomy* refers to items that address a child's ability to manage their everyday tasks, become more independent, and make their own decisions, e.g. daily routine and looking after myself. *Pain* relates to items about bodily pain and discomfort. *Disease-specific concepts* address disease-specific issues, e.g. disease severity, diet, and device management. This domain was limited to condition-specific PROMs only. *Overall QoL* refers to a single overview item asking respondents to rate their QoL.

The most examined domain was physical health (measured by 21 PROMs), followed by emotional health and

[1] This information comes from developers but may not have been tested under a comparable environment.

Table 4. Pictorial patient-reported outcome measures (PROMs) domains

Domain name	Domain definition	PROM name	Domains	Number of PROMs
Physical health	Items that examine physical functioning status, activity impairment, physical conditions, and their related symptoms	1. CHILDSPLA 2. 17D 3. AUQUEI Nursery 4. AUQUEI Primary 5.CHRIS 6. COOP Charts 7. DUX-25 8. ExqoL 9. KidIQoL 10. PedsQL GCS—Young Child 11. PedsQL SF15—Young Child 12. P-MYMOP 13. QoL-C 14. TedQL 15. CAQ-A 16. CDDUX 17. CDLQI 18. DISABKIDS Smileys TAKE 6 19. Hvidoere Smiley Faces diabetes QoL 20. PAQLQ—Pictorial version 21. Cochlear implants—QOL	1. Able to join in activities, tired 2. Mobility, vision, hearing, breathing, sleeping, eating, speech, excretion 3. Health 4. Health 5. Physical health, energy 6. Symptoms 7. Physical 8. Symptoms, physical activity 9. Physical and psychological health 10. Physical functioning 11. Physical functioning 12. Symptom 1, symptom 2, activity impairment, general well-being (child specifies symptoms and activity impairment affecting her most) 13. Moving 14. Physical competence 15. Physical functioning 16. Communication 17. Symptoms, sleep, clothes, playing, sports 18. Physical 19. Functional status, disease state, and physical symptoms 20. Physical impairment due to asthma 21. Fatigue	21
Social functioning	Items that address ability to engage in social interactions. Items addressing acceptance, friendship, family life and relationships, and social support were also categorised under this domain	1. 17D 2. AUQUEI Nursery 3. AUQUEI Primary 4. CHRIS 5. COOP Charts 6. DUX-25 7. ExqoL 8. KidIQoL 9. PedsQL GCS—Young Child 10. PedsQL SF15—Young Child 11. TedQL 12. CDLQI 13. DISABKIDS Smileys TAKE 6 14. Hvidoere Smiley Faces diabetes QoL 15. PAQLQ—Pictorial version 16. Cochlear implants—QOL	1. Friends 2. Family life, social life 3. Family life, social life 4. Social 5. Perceptions of social support 6. Social functioning 7. Social well-being, family relationships 8. Family life, social environment 9. Social functioning 10. Social functioning 11. Peer acceptance, maternal acceptance 12. Friendships, bullying 13. Social functioning 14. Psychological and social functioning 15. Social impairment due to asthma 16. Social functioning, child acceptance	16

Table 4. (continued)

Domain name	Domain definition	PROM name	Domains	Number of PROMs
Mental health	Items that capture mental and cognitive health conditions	1. 17D 2. CHRIS general health 3. DISABKIDS Smileys TAKE 6 4. TedQL	1. Mental function 2. Mental health, cognitive 3. Mental 4. Cognitive functioning	5
Emotional health	Items describing one's experience, perception, and expression of emotions. It includes items such as sadness, worry, and distress	1. CHILDSPLA 2. 17D 3. CHIP-CE CRF 4. COOP Charts 5. DUX-25 6. ExqoL 7. KidIQoL 8. PedsQL GCS 9. PedsQL SF15—Young Child 10. TedQL 11. QoL-C 12. Bt-DUX 13. CAQ-B 14. CAQ-C 15. CDLQI 16. CHRIS—preoperative anxiety 17. PAQLQ—Pictorial version 18. Cochlear Implants—QOL	1. Sad, annoyed, worried 2. Distress, vitality, concentration, depression 3. Satisfaction, comfort, resilience, risk avoidance 4. Emotional feelings 5. Emotion 6. Worry 7. Psychological health 8. Emotional functioning 9. Emotional functioning 10. Psychological functioning, 11. Feeling worried, sad, or unhappy 12. Emotional functioning 13. Distress 14. Distress, reactivity 15. Embarrassment 16. Preoperative anxiety 17. Emotional impairment due to asthma 18. Emotional functioning	16
School	Includes facets that address school attendance and participation	1. CHILDSPLA 2. 17D 3. AUQUEI Nursery 4. AUQUEI Primary 5. ExqoL 6. KidIQoL 7. PedsQL GCS—Young Child 8. PedsQL SF15—Young Child 9. CDLQI 10. Cochlear Implants—QOL	1. Schoolwork 2. School and hobbies 3. Children's activities at school and leisure 4. Children's activities at school and leisure 5. School achievements 6. School life 7. School functioning 8. School functioning 9. School 10. Academic functioning	10
Environment	Covers different aspects of the immediate environmental living conditions	1. KidIQoL 2. Cochlear Implants—QOL	1. Social and physical environment 2. Noisy environments	2

Table 4. (continued)

Domain name	Domain definition	PROM name	Domains	Number of PROMs
Autonomy	Includes items that address a child's ability to manage their everyday tasks, become more independent, and make their own decisions	1. CHILDSPLA 2. 17D 3. AUQUEI Nursery 4. AUQUEI Primary 5. CHIP-CE CRF 6. DUX-25 7. P-MYMOP 8. QoL-C 9. CHRIS	1. Daily routine 2. Appearance 3. Activities at school and leisure 4. Activities at school and leisure 5. Achievement 6. Home functioning 7. Activity impairment 8. Looking after myself, doing usual activities 9. Role	9
Overall QoL	A single overview item asking respondents to rate their QoL	1. CHRIS 2. CAQ-A 3. CAQ-B 4. CAQ-C 5. P-MYMOP	1. Overall QoL 2. Quality of life 3. Active QoL, passive QoL 4. Active QoL, teenage QoL 5. General well-being	5
Pain	Items about bodily pain and discomfort	1. CHILDSPLA 2. CHRIS 3. 17D 4. QoL-C 5. CHRIS—postoperative pain	1. Pain 2. Pain 3. Discomfort and symptoms 4. Having pain 5. Postoperative pain management	4
Disease-specific concepts	Items addressing disease-specific issues	1. Bt-DUX 2. CDDUX 3. CAQ-B 4. CAQ-C 5. CDLQI 6. Cochlear Implants—QOL	1. Cosmetic 2. Diet, having celiac disease 3. Disease severity 4. Disease severity 5. Impact of treatment 6. Device management	6

17D 17-dimensional measure of health-related quality of life, *AUQUEI* Auto Questionnaire Enfant Image/Child Pictured Self Report, *Bt-DUX* Bone Tumour DUX-25, *CAQ* Childhood Asthma Questionnaire, *CDDUX* Celiac Disease DUX-25, *CDLQI* Children's Dermatology Life Quality Index, *CHILDSPLA* Children's Health States Preferences Learnt from Animation (computer animation program), *CHIP-CE CRF* Child Health and Illness Profile-Child Edition Child Report Form, *CHRIS* Child Health Rating Inventories, *Cochlear Implants—QOL* Cochlear Implant Quality of Life, *COOP Charts* Dartmouth Primary Care Cooperative Information Project Charts, *DUX-25* Dutch Children's AZL/TNO Quality of Life—Short Form, *ExqoL* Exeter Quality of Life, *Hvidoere Smiley Faces* Hvidoere Smiley Faces International diabetes quality of life assessment tool for young children, *PAQLQ* Pediatric Asthma Quality of Life Questionnaire, *PedsQL GCS* Pediatric Quality of Life Inventory, Generic Core Scales, *PedsQL SF15* Pediatric Quality of Life Inventory, Short Form 15, *P-MYMOP* Paediatric Measure Yourself Medical Outcomes Profile, *QoL* quality of life, *QoL-C* Quality of Life Scale for Children

social functioning (assessed by 16 PROMs each). The least measured domain was environment, captured by only two PROMs.

3.6 Cross-Cultural Issues

PROMs are generally designed for widespread usage. As such, it is important to address differences due to culture and language during the development of the PROM, as they may impact use and acceptance.

Three measures addressed cultural issues in the PROM development stage. The animated computer program to assess QoL in children with IBD was designed to be culturally acceptable to all Dutch children [57], while the DISABKIDS Smileys TAKE 6 was developed across seven European countries (Austria, France, Germany, Greece, the Netherlands, Scotland, and Sweden) [50, 58]. The Hvidoere Smiley Faces questionnaire was developed in 17 countries as part of a multi-collaborative study [51].

4 Discussion

This systematic literature review has identified PROMs designed for children that incorporate pictorial approaches to assess HRQoL. Furthermore, this systematic review has provided a comprehensive description of the development and key characteristics of these instruments, focusing on aspects highlighted in the ISPOR PRO Good Research Practices report recommendations, specifically considering developmental differences, content validity, appropriate design and formatting, and cross-cultural issues.

The systematic review identified 22 condition-specific, generic preference-weighted, and non-preference-weighted childhood PROMs that use pictorial formats to assess HRQoL. Twenty-eight versions of the PROMs were included. These PROMs were developed in several countries, with the USA and the UK contributing significantly. The targeted age groups for these instruments varied, encompassing children as young as 3 years old up to young people aged 12–18 years. Many PROMs were designed with several age-specific formats to accommodate the diverse developmental stages of children.

To our knowledge, this study is the first to systematically identify and describe existing generic and condition-specific, preference-based and non-preference-based PROMs that use pictorial formats to assess HRQoL in children using good research guidance recommended by the ISPOR task force report. This review considered measures where the child is the respondent to self-reported or interviewer-administered HRQoL PROMs for children with or without health conditions 18 years and below. This review is a valuable resource for establishing current knowledge of pictorial paediatric PROMs. It underscores the need for future research on pictorial versions of preference-weighted paediatric PROM development, broadening the array of condition-specific measures and addressing cross-cultural relevance.

4.1 Developmental Differences and Age-Based Criteria for PROM Administration

4.1.1 Target Age Groups

The review shows a diverse range of target age groups, spanning from children under 5 to adolescents aged 12–18. Specific PROMs are designed for each age group, reflecting their developmental needs. Notably, several PROMs exhibit versatility by accommodating multiple age groups, highlighting their adaptability. This diversity underscores the significance of selecting age-appropriate instruments to ensure meaningful assessments of HRQoL for children and adolescents.

4.2 Establish Content Validity of Paediatric PROMs

Establishing the content validity of a PROM is an important stage in its development [59, 60]. Thirteen studies highlighted the active participation of children in crafting the pictorial elements, aiming to create instruments that were more child-friendly, engaging, and relevant to the unique experiences of children. Various graphic methods were employed based on the content domain, including the use of smiley faces, cartoons, and other illustrations, e.g. thermometers. These approaches were tailored to facilitate children's expression of their feelings and experiences effectively. Despite growing evidence of children and adolescents being effective content experts [61, 62], only half of the PROMs identified involved children in the item or domain generation. However, over 70% of the PROMs were pilot-tested in child populations. And while some PROMs used a wide age range in content validation, most PROMs used narrow age ranges, as recommended by ISPOR and others [12]. Thus, it is necessary to evaluate the extent to which domains of standardised pictorial PROMs for children accurately capture the most relevant concepts.

4.3 Ensure that the Instrument is Designed and Formatted Appropriately for the Target Age Group

The design and formatting of child PROMs will usually reflect the age and cognitive abilities of the respondents [8, 55, 63]. The length of the PROMs ranged from four to 45 items, with the estimated completion time varying from 3 to 5 min to 20–30 min. The scoring methods differed, with some PROMs yielding only summary scores, while others provided both domain-specific and summary scores.

Additionally, preference-based measures used value sets based on the valuation of health states.

4.3.1 Domains

Implicitly, HRQoL, by definition, is multi-dimensional. Some authors suggest that, at a minimum, paediatric HRQoL measures should address psychological, biological, and social concepts [64]. This study found that while some core domains are included in most pictorial PROMs for children (i.e. physical, emotional, mental, and social well-being, and school functioning), there is no universal consensus on which domains to include. Individual PROM developers included dimensions they consider relevant for a given population. This finding is expected since individual measures are developed for different purposes.

4.3.2 Recall Period

The recall period is another important consideration in paediatric PROMs. The appropriate duration for the recall period depends on the child's memory and understanding of past time. A short recall period of 24 h is recommended for PROMs for regulatory decision-making [12]. Eight of the PROMs evaluated in this study had a recall period of today or current health, aligning with the ISPOR task force recommendations for shorter recall periods in paediatric PROMs. However, there was variation in the specified recall periods across the remaining PROMs, with some not explicitly stating the recall period. It is worth noting that the available information did not provide specific details about how recall periods were graphically represented in these studies.

4.4 Consideration of Cross-Cultural Issues

PROMs are generally designed for widespread usage. As such, it is important to address differences due to culture and language during the development of the PROM, as they may impact use and acceptance. Only three PROMs addressed cultural issues in their development, which suggests the need for this aspect to be considered more in the development of pictorial PROMs as they can be limited if not validated in a particular cultural setting. Two of the three PROMs (DISABKIDS Smileys TAKE 6, and the Hvidoere Smiley Faces questionnaire) addressed cross-cultural effects and multiple translations through multi-country collaborations, which should be viewed in a positive light.

4.4.1 Practical Recommendations

Drawing from the insights gained in this review, we present a set of practical recommendations to guide the selection of appropriate instruments for children and adolescents with specific characteristics. First, it is imperative to take into account the age and cognitive abilities of the intended respondents, choosing instruments that are developmentally suitable. Some PROMs have demonstrated reliable psychometric properties even in children as young as 5 years, underlining the feasibility of utilising pictorial PROMs for younger age groups. Thus, matching the PROM to the target age group is a fundamental consideration. Second, active involvement of children in the development process is encouraged, as it results in child-friendly, engaging, and relevant instruments. It is particularly beneficial when children contribute to the creation of items or domains within the instrument. Third, the design and scoring of child PROMs should align with the cognitive abilities and age of the respondents, emphasising the importance of appropriate length, scoring methods, and recall periods. Ensuring that the design is tailored to the target age group while providing meaningful scoring insights is critical. Finally, for PROMs intended for widespread use, addressing cross-cultural factors and language differences is vital. In this regard, PROMs that have undergone cross-cultural validation, multi-country collaborations, and translations to ensure broader applicability and acceptance are preferable.

4.5 Strengths and Limitations

4.5.1 Strengths

The review follows a systematic approach in identifying, describing, and evaluating a wide range of paediatric PROMs that use pictorial formats to assess HRQoL. It provides a thorough overview of the existing literature in this area. This systematic review provides comprehensive evidence of existing self-report and interviewer-administered PROMs that use pictorial illustrations to enhance child self-reports of HRQoL. The review distinguishes between generic and domain- and condition-specific PROMs and between preference- and non-preference-based PROMs. The PROMs were assessed relative to ISPOR guidelines to inform the choice of PROMs that will accurately measure children's QoL and be considered in regulatory decisions. The review provides comprehensive information about the identified PROMs, including their origin, target populations, development methodologies, response scales, administration formats, domains, and measurement properties. Furthermore, it advocates child-centred practice in line with best practice recommendations safeguarding the right of children to have a voice in things that affect them.

4.5.2 Limitations

This review did not comprehensively critique the psychometric properties of identified PROMs, which is recommended

for PROM selection in research and decision-making [12], as this is the subject of a planned and separate systematic search and synthesis.

The review is based on the literature available up to its last search date, 31 January 2023. Newer studies or developments after that date are not included, potentially affecting the comprehensiveness and accuracy of the review's findings at publication.

4.5.3 Future Research Directions

Our research suggests the need to focus on a thorough evaluation of the psychometric properties of pictorial paediatric PROMs, which should involve rigorous assessments of their reliability, validity, responsiveness, and sensitivity to change. Conducting such evaluations is vital to solidify the credibility and robustness of these tools in both research and clinical applications. A subsequent systematic review is already in the planning stages, and it will be conducted by our team to further explore these aspects.

5 Conclusion

In conclusion, this review serves as a valuable resource for understanding the current landscape of paediatric PROMs using pictorial formats. It highlights the need for further research in developing pictorial preference-weighted PROMs, expanding the scope of condition-specific measures, and considering cross-cultural applicability. Additionally, ensuring age-appropriate design and rigorous content validation are crucial for the development of effective and reliable paediatric pictorial PROMs. Further investigation is needed to validate the potential benefits of pictorial PROMs in assessing HRQoL and facilitating self-report, particularly in younger children and children with developmental delays or severe chronic illness. An important observation is that only half of the existing pictorial PROMs for children actively involved children in their development, indicating room for improvement in research practices.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s40273-024-01357-z.

Declarations

Conflict of interest ND and DGTW are members of the EuroQol Group who are developers of HRQoL instruments (EQ-5D instruments), albeit not included in this study. ND, KD, and DGTW have received research grants from the EuroQol Research Foundation. DGTW has received honoraria for reviewing EuroQol foundation grant submissions. All other authors declare that they have no conflict of interest.

Funding Open Access funding enabled and organized by CAUL and its Member Institutions. This research was funded by an Australian Government Medical Research Futures Fund (MRFF) grant (1200816), project title 'Quality of Life in Kids: Key Evidence to Strengthen Decisions in Australia (QUOKKA)'.

Author contributions Conceptualisation and methodology: All authors. Database search, study selection, and data extraction: CMK, LL, and NBB. Writing the first draft: CMK. Draft reviewing, editing, and approval of the final amendments: All authors.

Data Availability Not applicable.

Role of the funder/sponsor The funder had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; or decision to submit the manuscript for publication.

Disclosure This article is published in a journal supplement wholly funded by the Australian Government MRFF funded QUOKKA research grant 1200816, the University of Melbourne, the EuroQol Research Foundation, the University of Technology Sydney and Flinders University.

References

1. Black N. Patient reported outcome measures could help transform healthcare. BMJ. 2013;346: f167.
2. Dawson J, et al. The routine use of patient reported outcome measures in healthcare settings. BMJ. 2010;340: c186.
3. Kwon J, et al. Systematic review of conceptual, age, measurement and valuation considerations for generic multidimensional childhood patient-reported outcome measures. Pharmacoeconomics. 2022;40(4):379–431.
4. Brazier J, et al. Measuring and valuing health benefits for economic evaluation. 2nd ed. London: Oxford University Press; 2017.
5. Chiou CF, et al. Development of the multi-attribute Pediatric Asthma Health Outcome Measure (PAHOM). Int J Qual Health Care. 2005;17(1):23–30.
6. Kreimeier S, et al. EQ-5D-Y-5L: developing a revised EQ-5D-Y with increased response categories. Qual Life Res. 2019;28(7):1951–61.
7. Haverman L, et al. Paediatric health-related quality of life: what is it and why should we measure it? Arch Dis Child. 2017;102(5):393–400.

8. Connolly MA, Johnson JA. Measuring quality of life in paediatric patients. Pharmacoeconomics. 1999;16(6):605–25.
9. Varni JW, Burwinkle TM, Lane MM. Health-related quality of life measurement in paediatric clinical practice: an appraisal and precept for future research and application. Health Qual Life Outcomes. 2005;3:34.
10. Raat H, et al. Reliability and validity of comprehensive health status measures in children: the Child Health Questionnaire in relation to the Health Utilities Index. J Clin Epidemiol. 2002;55(1):67–76.
11. Ravens-Sieberer U, et al. The KIDSCREEN-52 quality of life measure for children and adolescents: psychometric results from a cross-cultural survey in 13 European countries. Value Health. 2008;11(4):645–58.
12. Matza LS, et al. Pediatric patient-reported outcome instruments for research to support medical product labeling: report of the ISPOR PRO good research practices for the assessment of children and adolescents task force. Value Health. 2013;16(4):461–79.
13. Khanna D, et al. Are we agreed? Self-versus proxy-reporting of paediatric health-related quality of life (HRQoL) using generic preference-based measures: a systematic review and meta-analysis. Pharmacoeconomics. 2022;40(11):1043–67.
14. Khadka J, et al. Mind the (inter-rater) gap. An investigation of self-reported versus proxy-reported assessments in the derivation of childhood utility values for economic evaluation: A systematic review. Soc Sci Med. 2019;240:112543.
15. Jiang M, et al. A comparison of self-reported and proxy-reported health utilities in children: a systematic review and meta-analysis. Health Qual Life Outcomes. 2021;19(1):45.
16. Mpundu-Kaambwa C, et al. A review of preference-based measures for the assessment of quality of life in children and adolescents with cerebral palsy. Qual Life Res. 2018;27(7):1781–99.
17. Grange A, et al. Adequacy of health-related quality of life measures in children under 5 years old: systematic review. J Adv Nurs. 2007;59(3):197–220.
18. Varni JW, Limbers CA, Burwinkle TM. How young can children reliably and validly self-report their health-related quality of life? An analysis of 8591 children across age subgroups with the PedsQL™ 4.0 Generic Core Scales. Health Quality Life Outcomes. 2007;5(1):1.
19. Solans M, et al. Health-related quality of life measurement in children and adolescents: a systematic review of generic and disease-specific instruments. Value Health. 2008;11(4):742–64.
20. Mahakwe G, et al. A systematic review of self-report instruments for the measurement of anxiety in hospitalized children with cancer. Int J Environ Res Public Health. 2021;18(4):1911.
21. Houts PS, et al. The role of pictures in improving health communication: a review of research on attention, comprehension, recall, and adherence. Patient Educ Couns. 2006;61(2):173–90.
22. Schubbe D, et al. Using pictures to convey health information: a systematic review and meta-analysis of the effects on patient and consumer health behaviors and outcomes. Patient Educ Couns. 2020;103(10):1935–60.
23. Green MJ, Myers KR. Graphic medicine: use of comics in medical education and patient care. BMJ. 2010;340: c863.
24. Coombes L, et al. Enhancing validity, reliability and participation in self-reported health outcome measurement for children and young people: a systematic review of recall period, response scale format, and administration modality. Qual Life Res. 2021;30(7):1803–32.
25. Cremeens J, Eiser C, Blades M. Factors influencing agreement between child self-report and parent proxy-reports on the Pediatric Quality of Life Inventory 4.0 (PedsQL) generic core scales. Health Qual Life Outcomes. 2006;4:58.
26. Arsiwala T, et al. Measuring what matters for children: a systematic review of frequently used pediatric generic PRO instruments. Ther Innov Regul Sci. 2021;55(5):1082–95.
27. Moher D, et al. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. Syst Rev. 2015;4(1):1–9.
28. Liberati A, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. PLoS Med. 2009;6:e1000100.
29. Google. Google Translate. 2023. https://translate.google.com. Accessed 10 Oct 2023.
30. Cohen J. A coefficient of agreement for nominal scales. Educ Psychol Measur. 1960;20(1):37–46.
31. Riley AW, et al. The child report form of the CHIP-child edition: reliability and validity. Med Care. 2004;42:221–31.
32. Kaplan SH, et al. Assessing functional status and health-related quality of life among school-aged children: reliability and validity of a new self-reported measure. Qual Life Res. 1995;4(5):444.
33. Wasson JH, et al. Adolescent health and social problems. A method for detection and early management. The Dartmouth Primary Care Cooperative Information Project (COOP). Arch Fam Med. 1995;4(1):51–6.
34. Varni JW, Seid M, Kurtin PS. PedsQL™ 4.0: reliability and validity of the Pediatric Quality of Life Inventory™ Version 4.0 Generic Core Scales in healthy and patient populations. Med Care. 2001;39:800–12.
35. Chan KS, et al. The PedsQL: reliability and validity of the short-form generic core scales and Asthma Module. Med Care. 2005;43(3):256–65.
36. Kaplan SH, et al. Development and initial validation of self-report measures of general health, preoperative anxiety, and postoperative pain in young children using computer-administered animation. Paediatr Anaesth. 2021;31(2):150–9.
37. Everhart RS, Fiese BH. Development and initial validation of a pictorial quality of life measure for young children with asthma. J Pediatr Psychol. 2009;34(9):966–76.
38. Hoffman MF, Cejas I, Quittner AL. Health-related quality of life instruments for children with cochlear implants: development of child and parent-proxy measures. Ear Hear. 2019;40(3):592–604.
39. Abrines Jaume N, et al. CHILDSPLA: a collaboration between children and researchers to design and animate health states. Child Care Health Dev. 2015;41(6):1140–51.
40. Eiser C, et al. Health-related quality-of-life measures for children. Int J Cancer Suppl. 1999;12:87–90.
41. Thompson HL, et al. The quality of life scale for children (QOL-C). J Childrens Serv. Res. Inf. Policy Pract. 2014;9(1):4–17.
42. Christie MJ, et al. Development of child-centered disease-specific questionnaires for living with asthma. Psychosom Med. 1993;55(6):541–8.
43. Holme SA, et al. The Childrens Dermatology Life Quality Index: validation of the cartoon version. Br J Dermatol (1951). 2003;148(2):285–90.
44. Lawford J, Volavka N, Eiser C. A generic measure of Quality of Life for children aged 3–8 years: results of two preliminary studies. Pediatr Rehabil. 2001;4(4):197–207.
45. Manificat S, Dazord A. Children's quality of life assessment: preliminary results obtained with the AUQUEI questionnaire. Qual Life Newsl. 1998;19:2–3.
46. Manificat S, Dazord A. Infant, child and adolescent quality of life: surveys performed in a European context. Expert Rev Pharmacoecon Outcomes Res. 2002;2(6):589–96.
47. Gayral-Taminh M, et al. Auto-évaluation de la qualité de vie d'enfants de 6 à 12 ans : construction et premières étapes de

validation du KidIQol, outil générique présenté sur ordinateur. Santé Publ. 2005;17(2):167–77.
48. Apajasalo M, et al. Quality of life in pre-adolescence: a 17-dimensional health-related measure (17D). Qual Life Res. 1996;5(6):532–8.
49. Ishaque S, et al. Adaptation/content validation of measure yourself medical outcomes profile (MYMOP) questionnaire for 7-11 year old children. In: Quality of life research. Van Godewijckstraat 303311 GX Dordrecht, Netherlands: Springer; 2018. vol 27, pp S114–S115.
50. Chaplin JE, Koopman HM, Schmidt S. DISABKIDS smiley questionnaire: the TAKE 6 assisted health-related quality of life measure for 4 to 7-year-olds. Clin Psychol Psychother. 2008;15(3):173–80.
51. Hoey H, et al. Hvidoere smiley faces: international diabetes quality of life assessment tool for young children. Pediatr Diabetes. 2018;19(3):553–8.
52. Koopman HM, Theunissen NCM, Vogels TGC. The DUX-25: a short form questionnaire for measuring health related quality of life of children with chronic illness. Qual Life Res. 1998;7:619.
53. van Doorn RK, et al. CDDUX: a disease-specific health-related quality-of-life questionnaire for children with celiac disease. J Pediatr Gastroenterol Nutr. 2008;47(2):147–52.
54. Bekkering WP, et al. The Bt-DUX: development of a subjective measure of health-related quality of life in patients who underwent surgery for lower extremity malignant bone tumor. Pediatr Blood Cancer. 2009;53(3):348–55.
55. Eiser C, Vance Y, Seamark D. The development of a theoretically driven generic measure of quality of life for children aged 6–12 years: a preliminary report. Child Care Health Dev. 2000;26(6):445–56.
56. Thompson HL, et al. The quality of life scale for children (QoL-C). J Childrens Serv Res Inf Policy Pract. 2014;9(1):4–17.
57. Buller H. Assessment of quality of life in the younger child: the use of an animated computer program. J Pediatr Gastroenterol Nutr. 1999;28(4):S53–5.
58. Schmidt S, et al. The DISABKIDS Questionnaires: quality of life questionnaires for children with chronic conditions. Lengerich: Pabst Science Publishers; 2006.
59. Patrick DL, et al. Content validity—establishing and reporting the evidence in newly developed patient-reported outcomes (PRO) instruments for medical product evaluation: ISPOR PRO Good Research Practices Task Force report: part 2–assessing respondent understanding. Value Health. 2011;14(8):978–88.
60. Patrick DL, et al. Content validity—establishing and reporting the evidence in newly developed patient-reported outcomes (PRO) instruments for medical product evaluation: ISPOR PRO good research practices task force report: part 1—eliciting concepts for a new PRO instrument. Value Health. 2011;14(8):967–77.
61. Riesch SK, et al. Evaluating content validity and test–retest reliability of the children's health risk behavior scale. Public Health Nurs. 2006;23(4):366–72.
62. Tomlinson D, et al. Understandability, content validity, and overall acceptability of the Children's International Mucositis Evaluation Scale (ChIMES): child and parent reporting. J Pediatr Hematol Oncol. 2009;31(6):416–23.
63. Matza LS, et al. Assessment of health-related quality of life in children: a review of conceptual, methodological, and regulatory issues. Value Health. 2004;7(1):79–92.
64. Fayed N, et al. Generic patient-reported outcomes in child health research: a review of conceptual content using World Health Organization definitions. Dev Med Child Neurol. 2012;54(12):1085–95.
65. Herdman M, et al. Expert consensus in the development of a European health-related quality of life measure for children and adolescents: a Delphi study. Acta Paediatr. 2002;91(12):1385–90.
66. Keck JF, et al. Reliability and validity of the Faces and Word Descriptor Scales to measure procedural pain. J Pediatr Nurs. 1996;11(6):368–74.
67. Mellor D, Moore KA. The use of likert scales with children. J Pediatr Psychol. 2013;39(3):369–79.

Authors and Affiliations

Christine Mpundu-Kaambwa[1] · Norma B. Bulamu[2] · Lauren Lines[1] · Gang Chen[3] · David G. T. Whitehurst[4] · Kim Dalziel[5,6] · Nancy Devlin[6] · Julie Ratcliffe[1] on behalf of the Quality of Life in Kids: Key Evidence for Decision Makers in Australia (QUOKKA) project team

✉ Christine Mpundu-Kaambwa
christine.mpundu-kaambwa@flinders.edu.au

[1] Health and Social Care Economics Group, Caring Futures Institute, College of Nursing and Health Sciences, Flinders University, Sturt Road, Bedford Park, Adelaide, SA 5042, Australia

[2] Flinders Health and Medical Research Institute, College of Medicine and Public Health, Flinders University, Adelaide, SA, Australia

[3] Centre for Health Economics, Monash Business School, Monash University, Melbourne, VIC, Australia

[4] Faculty of Health Sciences, Simon Fraser University, Burnaby, BC, Canada

[5] Murdoch Children's Research Institute, Parkville, Melbourne, VIC, Australia

[6] Health Economics Unit, Melbourne School of Population and Global Health, University of Melbourne, Melbourne, VIC, Australia