Australian Paediatric Multi-Instrument Comparison (P-MIC) Study: Technical Methods Paper

Version 3 (26/04/2023), Data Cut 2 (10/08/2022)

Renee Jones [1,3], Brendan Mulhern [2], Nancy Devlin [1], Harriet Hiscock [3,4,5], Gang Chen [6], Rachel O'Louglin[1], Xiuqin Xiong [1], Mina Bahrampour [2], Kristy McGregor [3], Shilana Yip[3], and Kim Dalziel [1,3,*] on behalf of the Quality Of Life in Kids: Key evidence to strengthen decisions in Australia (QUOKKA) project team.

[1] Health Economics Unit, Centre for Health Policy, The University of Melbourne, Melbourne, Australia;

[2] Centre for Health Economics Research and Evaluation, University of Technology Sydney, Ultimo, Australia

[3] Health Services and Economics, Murdoch Children's Research Institute, Melbourne, Australia;

[4] Health Services Research Unit, Royal Children's Hospital, Melbourne, Australia

[5] Department of Paediatrics, The University of Melbourne, Melbourne, Australia

[6] Centre for Health Economics, Monash University, Melbourne, Australia

Table of contents:

1. Purpose and scope of this document 2
2. Study aim 2
3. Study design 2
4. Timelines 2
5. Participants 3
6. Recruitment 6
7. Instruments 8
8. Survey development, piloting, and testing 13
9. Data collection and survey administration 13
10. Participant reimbursement 16
11. Quotas 17
12. Quality monitoring 17
13. Decision to close samples 17
14. Psychometric Analysis Guide 17
15. Funding 29
16. Conflicts of interest 29
17. References 29
18. Appendix 32

1. Purpose and scope of this document

The purpose of this document is to outline the methods used to conduct the Australian Paediatric Multi-Instrument Comparison (P-MIC) Study. A study protocol for the P-MIC has previously been published, providing an overview of the planned P-MIC methodology.(1) This document provides a detailed description of the methodology used to conduct the P-MIC study, providing additional detail to that which is publishable via peer reviewed journals, both to be fully transparent about study design, and to help others who may be interested in undertaking similar studies in future. As recruitment is ongoing (see Table 1), data will be cut at certain time points and a new version of the technical methods paper will be produced for each data cut to transparently summarise the data available in each cut. Updated versions of this document will be available on the QUOKKA Research Program website. The data cut for this version of the technical methods paper, data cut 2, was taken on the 10th August 2022 and includes 6,787 children and their caregivers (see Table 1). This is expected to represent approximately 94% of the total planned P-MIC participants. For a summary of previous data cuts (data cut 1 06 May 2022), please see Appendix Table 1.

2. Study aim

The broad aim of this study is to compare the performance of a range of paediatric generic and condition specific Health Related Quality of Life (HRQoL) instruments in terms of validity, reliability, responsiveness, acceptability, feasibility, measurement relationship, and consistency across age, proxy and self-report, and health condition groups.(1) Within this overall aim, there are many specific aims that will be investigated. These will be reported in subsequent papers, reports and other dissemination activities.

3. Study design

The P-MIC study prospectively collected multiple generic and condition specific paediatric HRQoL instruments concurrently in a single online survey collected at two time points, initial and follow-up. Most participants receive the follow-up survey 4-weeks after completing the initial survey to assess change in health and instrument responsiveness, however, a small sub-set of children from the general population sample receive the follow-up survey at 2-days to assess test-retest reliability. A 4-week follow-up was selected to assess responsiveness as it was considered enough time for children with acute health conditions at the time of initial survey to change their health status for the follow-up survey and was also a short enough so as to not place pressure on follow-up survey completion rates. Recruitment was conducted across both a tertiary hospital and an online panel. This study was overseen by study investigators and guided by input and feedback from 1) the wider Quality Of Life in Kids: Key Evidence to Strengthen Decisions in Australia (QUOKKA) Project investigators, 2) a Consumer Advisory Group, made up of parents and caregivers of children with and without health conditions, and 3) a Decision Makers Panel, made up of industry and government stakeholders. This study was approved by The Royal Children's Hospital (RCH) Human Research Ethics Committee (HREC/71872/RCHM2021) on 20th May 2021 and registered with the Australia New Zealand Clinical Trials Registry (ACTRN12621000657820) on 31st May 2021.

4. Timelines

The P-MIC study received ethics approval in May 2021. Hospital recruitment began in June 2021 and online panel recruitment began in October 2021. Recruitment for some samples was ongoing at the time of this data cut (see Table 1).

5. Participants

5.1. Overview of P-MIC samples

The P-MIC study includes three key samples: Sample 1) recruited via hospital, Sample 2) general population recruited via online panel, and Sample 3) health condition-specific groups recruited primarily via online panels (see 6.3 for further information).

- Sample 1 includes participants recruited via hospital and has two subsamples:
  - Sample 1a, general hospital sample recruited via The RCH, Melbourne, Australia. Sample 1a includes any participant recruited via the hospital, the children were not required to have any particular condition nor were they required to be a patient of the hospital.
  - Sample 1b, specialised hospital sample recruited via The RCH, Melbourne, Australia or The Royal Women's Hospital (RWH), Melbourne Australia. Sample 1b includes five samples: 1) children receiving care in the intensive care unit (ICU), 2) children receiving care in the Short Stay Unit (SSU), 3) children born extremely premature, or 4) children with a rare genetic condition.
- Sample 2 is the general population sample not reporting one of the health condition groups recruited via online panels and includes two sub-samples:
  - Sample 2a, general population sample with a four-week follow-up (same as the rest of the samples), and
  - Sample 2b, general population sample with two-day follow-up.
- Sample 3 is the health condition groups primarily recruited via online panels and includes 11 sub-samples:
  - Sample 3a, attention deficit hyperactivity disorder (ADHD),
  - Sample 3b, anxiety or depression,
  - Sample 3c, autism spectrum disorder (ASD),
  - Sample 3d, asthma,
  - Sample 3e, eating disorders,
  - Sample 3f, epilepsy,
  - Sample 3g, recurrent abdominal pain,
  - Sample 3h, sleep problems,
  - Sample 3i, tooth problems,
  - Sample 3j, type 1 diabetes, and
  - Sample 3k, wetting problems.

Table 1: Summary of P-MIC samples, number recruited# to each sample and recruitment status for data cut 2.

Sample	Sub-sample	N	Recruitment status
Total	n/a	6,787	Ongoing
1) Recruited via hospital	1a) general hospital sample	916	Ongoing
	1b) specialised hospital sample, including the following five groups:	151	Ongoing
	ICU		
	SSU	27	Ongoing
	Born premature	25	Ongoing
	Rare genetic condition	26	Ongoing
		73*	Complete

2)	General population sample recruited via online panels	2a) general population sample with a four-week follow-up	1,642	Complete
		2b) general population sample with two-day follow-up	252	Complete
3)	Health condition-specific groups primarily recruited via online panels	3a) ADHD	533	Complete
		3b) Anxiety or depression	480	Complete
		3c) ASD	510	Complete
		3d) Asthma	487	Complete
		3e) Eating disorder	186	Ongoing
		3f) Epilepsy	289	Ongoing
		3g) Recurrent abdominal pain	392	Complete
		3h) Sleep problems	459	Complete
		3i) Tooth problems	490	Complete
		3j) Type 1 diabetes	0	Ongoing**
		3k) Wetting problems	0	Ongoing**

*Note: All ongoing samples are subject to change in future data cuts as recruitment is still open. Additionally, all ongoing samples may not have had all data quality checks completed. #A participant is considered recruited if they have consented, passed minimum quality eligibility criteria, and completed at least initial survey. * A total of 155 children in the dataset have a rare genetic condition, however only 73 of these were recruited via hospital, with the remaining 82 recruited via online panel. ** The samples of children with Type 1 diabetes and Wetting problems were decided to be included as additional samples in early 2023.*

A sample of children recruited via a large tertiary hospital (Sample 1) was selected to ensure children with a range of moderate to severe health conditions were included in the sample, enabling the assessment of instruments in children who are very unwell. Additionally, the specialised hospital sample (Sample 1b) was included to ensure children who are extremely unwell who likely have severe decrements in quality of life were represented in the sample. The condition groups for Sample 3 were chosen based on the following criteria:

1) evidence of reduced quality of life and documented change in quality of life over time,(2) this was based on an internal analysis of data from Longitudinal Study of Australian Children (LSAC) across child ages for 30 conditions able to be identified in the data,
2) being common conditions so as to be feasible for recruitment via an online panels,(3)
3) preference-based measures have not previously been extensively studied or validated extensively in the condition group,(4)
4) conditions that would give a balance of impacts across different common dimensions (e.g., pain, participation in usual activities, mental health, mobility), for example, two pain focussed conditions would not be selected, and
5) having a suitable validated condition-specific measure of quality of life or symptoms available.

5.2. Inclusion criteria

Any parent, caregiver, or guardian of a child(ren) aged 2–18 years (inclusive) at study enrolment. Additional criteria apply to Sample 3, health condition groups (see section 5.4).

5.3. Exclusion criteria

Any parent who is unable to communicate in written English, unable to answer or comprehend the survey questions or those who do not reside in Australia.

5.4. Screening for health condition samples

Additional eligibility criteria were applied to Sample 3, health condition groups. Screening questions were used to determine eligibility (see Table 2). Screening questions were designed to capture children currently experiencing the condition or if episodic, a recent episode of the condition, as diagnosed by a doctor or relevant health professional. Where possible, screening questions were derived from previous surveys such as the Longitudinal Study of Australian Children (LSAC) or through consultation with clinical experts.(5) The

age range eligibility for each health condition sample was based on the validated age range for the corresponding health condition instrument selected as well as expert clinician advice.

Table 2: Screening questions and eligibility for health condition samples (Sample 3)

Health condition sample	Screening and eligibility questions	Child age range in years
3a. ADHD	Do you have a child aged 4-18 years with attention deficit hyperactivity disorder (ADHD) as diagnosed by a health professional?(5) Yes - *inclusion* No - *exclusion*	4-18
3b. Anxiety or depression	Do you have a child aged 7-18 years with anxiety or depression as diagnosed by a health professional? (5) Yes - *inclusion* No - *exclusion*	7-18
3c. ASD	Do you have a child aged 5-18 years with autism spectrum disorder (ASD) as diagnosed by a health professional? (5) Yes - *inclusion* No - *exclusion*	5-18
3d. Asthma	Do you have a child aged 5-18 with asthma as diagnosed by a doctor? (5) Yes - *go to next question* No - *exclusion* Has your child had symptoms of asthma or used an asthma treatment in the last 12 months? Yes- *inclusion* No - *exclusion*	5-18
3e. Eating disorder	Do you have a child aged 14-18 with an eating disorder (such as anorexia, bulimia, or avoidant restrictive food intake disorder) as diagnosed by a health professional? (5) Yes - *inclusion* No - *exclusion*	14-18
3f. Epilepsy	Do you have a child with epilepsy, or a seizure disorder as diagnosed by a doctor? (5) Yes - *inclusion* No - *exclusion*	4-18
3g. Recurrent abdominal pain	Do you have a child with the ongoing condition 'recurrent abdominal pain'? (5) *Recurrent abdominal pain is at least three episodes of pain that occur over at least three months and affect the child's ability to perform normal activities.(6)* Yes - *inclusion* No - *exclusion*	5-18
3h. Sleep problems	Thinking about your child aged 3-16 with sleep problems, how much is their ongoing sleeping pattern or habits a problem for you? (5) Not a problem at all- *exclusion* A small problem- *exclusion* A moderate problem- *inclusion* A large problem - *inclusion*	3-16
3i. Tooth problems	Do you have a child who **currently has** or has **experienced in the last 3 months,** any of the following **tooth problems**? (5) *This includes problems that have been treated, untreated or are still undergoing treatment.* Yes, cavities, dental decay or hole(s) in teeth - *inclusion* Yes, tooth or teeth filled because of dental decay - *inclusion* Yes, teeth pulled because of dental decay - *inclusion* Yes, accident causing breakage or loss of teeth - *inclusion* Yes, crowded teeth - *inclusion* Yes, problems with bite (e.g., crossbite or overbite) - *inclusion*	5-18

Health condition sample	Screening and eligibility questions	Child age range in years
	No, my child has not experienced any of the above tooth problems - *exclusion*	
3j. Type 1 diabetes	Do you have a child aged 5-18 with Type 1 Diabetes (requiring insulin) as diagnosed by a doctor? Yes - inclusion No - exclusion	5-18
3k. Wetting problems	Do you have a child aged 6-17 years who currently has any of the following ongoing conditions? Tick all that apply Wetting self during the day – inclusion Bed wetting 4 or more nights a week – inclusion None of the above - exclusion	6-17

5.5. Caregivers with multiple eligible children

Where caregivers had multiple eligible children for any given sample, they were directed to respond to the survey questions based on the child with the highest health needs. Caregivers were directed to complete the survey in relation to one child only.

6. Recruitment

6.1. Sample 1, hospital sample recruitment:

6.1.1. Sample 1a, general hospital sample:

Research Assistants (RAs) approached caregivers for recruitment from a range of RCH departments, including outpatient clinics and surgical department waiting rooms. Poster advertisements with QR codes linking to the study were placed in high traffic areas of The RCH. Online advertisements with a link to the study were placed on RCH telehealth appointments virtual platform, appearing for any family attending a hospital appointment via telehealth. Additionally, the study advert was shared with caregivers from the onsite RCH childcare centre.

6.1.2. Sample 1b, specialised hospital sample:

In addition to the above recruitment strategies (6.1.1), which were also used to recruit children to the specialised hospital samples, several specific recruitment methods were also used:

- **ICU (intensive care unit):** ICU research staff approached potential participants for consent prior to the child's admission to ICU (e.g., pre-operative clinic visits or while in hospital). Elective admissions were the focus of active recruitment. This approach ensured avoiding approaching families in high stress or where an approach from the study was considered inappropriate. For example, where the child was unlikely to survive. The ICU research staff notified the study team when the consented participant was admitted to ICU and the study team then sent the family the survey link with a friendly reminder to complete the survey.
- **SSU (short stay unit):** Recruitment of SSU patients needed a specialised approach due to the COVID impacts on the RCH and the use of the SSU as a COVID-19 ward. This limited study research staff from physically attending these spaces to recruit. A strategy was used whereby advertisements were printed and handed to an attending doctor to hand out to families in SSU.
- **Born premature:** Participants from the study 'Preventing Chronic Lung Health condition in Extremely Preterm Infants Using Surfactant + Steroid' (PLUSS) trial (ACTRN12617000322336), an interventional trial of children born less than 28 weeks' gestation, were approached for recruitment to the study (if the child was 2 years or older, corrected for prematurity). Potential participants were approached for recruitment by a member of the PLUSS research team when they attended The RWH for the

developmental clinic/PLUSS study 2 year follow-up. Participants were provided with an advertisement inviting them to also participate in this survey.

- **Rare genetic condition:** Eligible participants (children currently aged 2-18 years old who are still alive) from Australian Genomics study cohorts who consented to be contacted for future research as part of their involvement in a previous study with Australian Genomics were sent an email from Australian Genomics inviting them to take part in the study.

6.2. Samples 2 and 3, online panel sample recruitment:

The recruitment of online panel samples was managed by Pureprofile Pty Ltd Australia (www.pureprofile.com, accessed on 14th June 2022). Potential participants were randomly selected from this panel to take part in the study if they met eligibility criteria. Participants were selected based on quotas for age. As children may have multiple health conditions, entry to the different samples was managed on a 'least fill' basis, with samples filled from least to most prevalent (see Table 3). Where estimated prevalence was the same, the study team discussed and prioritised the condition they felt would be the hardest to fill. Hence, children with rarer conditions were invited to take part for the rarer condition, even if the child had another health condition. If a child had none of the health conditions (i.e., they did not meet the eligibility for Sample 3), they were invited to take part in the general population survey (Sample 2). Additionally, as quotas for the eating disorder (3e) and epilepsy (3f) samples were not able to be reached, two additional samples, type 1 diabetes (3j) and wetting (3k), were added in January 2023 to top up the condition specific sample group (Sample 3). The additional sample (3j and 3k) were recruited separately to other samples and hence have a separate hierarchy.

Table 3: Least fill hierarchy of health condition samples and estimated prevalence for age range (Sample 3)

Health condition sample	Estimated prevalence for age range	Least fill priority/ hierarchy
3a. ADHD	3-5% (3)	3a
3b. Anxiety or depression	5-10% (3, 7)	6a
3c. ASD	2-5% (3)	2a
3d. Asthma	10-15% (3, 7)	9a
3e. Eating disorders	4-16% (8)	4a
3f. Epilepsy	0.5-1% (3)	1a
3g. Recurrent abdominal pain	3-5% (3)	5a
3h. Sleep problems	10-15% (3)	8a
3i. Tooth problems	10-30% (3)	7a
3j. Type 1 diabetes	0.3-0.8% (3)	1b
3k. Wetting problems	5-20% (3, 9)	2b

6.3. Sample 3, hybrid recruitment for hard-to-fill health condition samples:

Two of the health condition samples, epilepsy and eating disorders, were not able to be filled to the desired sample size by the online survey panel company. Hence, these samples were recruited via a hybrid approach of online survey panels and supplementary recruitment methods managed by the study team.

The supplementary recruitment methods used by the study team included:

- **RCH Telehealth appointments:** We advertised in the virtual waiting room and at the end of all TH appointments. The advert included a short description of the study and a link to the PICF and survey.

- **Social media:** Facebook advert targeting families of children with an eating disorder.
- **Relevant newsletters/ email subscription lists:** The study advert was shared via e-newsletter, email subscription lists and notice boards of relevant organisations who are interested in sharing the study information with their subscribers (e.g. The Victorian Centre of Excellence in Eating Disorders (CEED) etc.). The newsletter adverts and emails were only sent to people who opted in to/ subscribed to receive the email/newsletter.
- **Opt-in letter of invitation from RCH clinics:** Using patient lists from relevant RCH clinical departments, a letter of invitation was sent to eligible participants. The letter was an opt-in style letter with a short description of the study and a QR code linking to the survey.

7. Instruments

7.1. Paediatric HRQoL instruments

Paediatric HRQoL instruments included in the P-MIC study were classified as 'core', included for all samples, or 'additional', included for only a portion of samples (see Table 5 for a summary of which samples were allocated which instruments). To minimise responder burden in the sample recruited via hospital (Sample 1), only core instruments were included in the survey. To minimise responder burden in the online panel samples (Samples 2 and 3), participants were randomised to receive one additional instrument block. The Paediatric Quality of Life Inventory (PedsQL) Core Generic Version 4.0, EQ-5D Youth 3 level (EQ-5D-Y-3L), EQ-5D Youth 5 level (EQ-5D-Y-5L), and Child Health Utility 9D (CHU-9D) were included as core instruments following a recent systematic review identifying these instruments as common, well performing, paediatric HRQoL instruments requiring further evidence regarding their psychometric performance.(4) The Toddler and Infant Questionnaire (TANDI) version 2 was also included as a core instrument as it is an experimental generic paediatric HRQoL instrument designed and validated for younger children, requiring further evidence on performance.(10) The Patient-Reported Outcome Measurement Information System 25 (PROMIS-25), Assessment of Quality of Life (AQoL-6D), Health Utilities Index Mark 2/3 (HUI 2/3), and EQ-5D-5L were included as additional instruments. The PROMIS-25 was included as an additional instrument because it is a new tool requiring further validation work with the adult version being routinely used as a PROM in some Australian hospitals. The AQoL-6D is a tool used less frequently internationally but was included as an additional instrument because of its use in Australian populations. The HUI 2/3 was included as an additional instrument because it has been used in Australian health technology assessment decision making for children, however, was not included as a core instrument as there is mixed evidence regarding its performance compared to other instruments.(4, 11) The EQ-5D-5L was included to build on a research agenda focused on transitions between EuroQol instruments across the lifespan. Table 4 summarises instrument characteristics. See Appendix Table 2 for a summary of instruments and instrument properties.

7.1.1. PedsQL generic core 4.0

The PedsQL generic core 4.0 is a proxy or self-report 23-item generic paediatric HRQoL instrument with 5 item response levels, a 1 month recall period, covering 4 domains: physical functioning, emotional functioning, social functioning, and school functioning.(12) Validated versions exist for children aged 2–18 years.(12) Respondents are asked to rate the frequency of each item over the past month on a 5-point scale from 0 (Never) to 4 (Almost always). The PedsQL generic core was developed specifically for a paediatric population through cognitive interviews and focus groups.(13) The PedsQL generic 4.0 was iteratively adapted from previous versions and was designed to ensure the core health dimensions outlined by the World Health Organisation were measured.(12)

7.1.2. TANDI

The TANDI is a proxy report 6-item generic paediatric HRQoL instrument designed for children <4 years of age with 3 item response levels, a 'today' recall period, covering 6 dimensions: movement, play, pain, social

interaction, communication, and eating.(10) The TANDI was developed from the structure of the EuroQol Youth version (EQ-5D-Y) using cognitive interviews with caregivers of young children and a Delphi study with experts to design the instrument for children <4 years of age.(10)

7.1.3.EQ-5D-Y (3L and 5L)

The EQ-5D-Y is a proxy or self-report 5-item generic paediatric HRQoL instrument.(14, 15) Both have a 'today' recall period and cover 5 dimensions: mobility, looking after self, usual activities, pain/discomfort, and worried/sad.(14) Respondents are asked to rate the severity of each item on a 3-point scale for the EQ-5D-Y-3L and on a 5-point scale for the EQ-5D-Y-5L. The EQ-5D-Y also includes a general health Visual Analogue Scale (VAS). The EQ-5D-Y-3L was adapted from the EQ-5D adult version using cognitive interviews and the EQ-5D-Y-5L was adapted from the EQ-5D-Y-3L.(14, 15) The EQ-5D-Y has been validated in children aged 4-18 years. Additionally, an adapted proxy version of the EQ-5D-Y for age 2-4 years with guidance notes is also trialled for children of this age.

7.1.4.CHU9D

The CHU9D is a proxy or self-report 9-item generic paediatric HRQoL instrument with 5 item response levels, a 'today' recall period, covering 9 dimensions: worried, sad, pain, tired, annoyed, schoolwork/homework, and sleep.(16, 17) Respondents are asked to rate the severity of each item on a 5-point scale. The CHU9D was developed specifically for use in younger children aged 6 to 11 years old, however, has been validated in children up to age 17.(16, 17) Additionally, a proxy version of the CHU9D with guidance notes available for under 5 years (method of development is unclear, but assumed to be adapted by instrument developers) and is being trialled for children of this age.

7.1.5.AQoL-6D Adolescent

The AQol-6D adolescent is a proxy or self-report 20-item generic adolescent HRQoL instrument with 4 to 6 item response levels, a 1 week recall period, covering 6 domains: independent living, mental health, coping, relationships, pain, and senses.(18, 19) Respondents are asked to rate the severity of each item on a 4- to 6-point scale. The adult AQoL-6D was adapted by instrument developers to develop the AQoL-6D for adolescents aged 12–18 years, however, has been used in children aged 11 years.(19, 20)

7.1.6.PROMIS-25 paediatric profile

The PROMIS-25 paediatric profile is a proxy or self-report 25-item generic paediatric HRQoL instrument with 5 item response levels (except for the pain item which is 10 levels), a 1 week recall period, covering 6 domains: physical function mobility, anxiety, depressive symptoms, fatigue, peer relationships, and pain interference.(21) Respondents are asked to rate the severity of 5-items and the frequency of items on a 5-point scale. Except for the pain item which is on a scale from 0-10. The PROMIS-25 was developed from the PROMIS-37 which was developed from the PROMIS-49. The PROMIS-25 is recommended for use in children aged 5 years and older.(21)

7.1.7.HUI 2/3

The Health Utilities Index Mark 2 and 3 (HUI2/3) is a proxy or self-report 15-item generic HRQoL instrument with 4 to 6 levels that can be used in paediatric populations.(22-24) The HUI2/3 instrument can be used to classify a participant's health according to either the HUI2 or HUI3 classification system. (22-24) The HUI3 classification system has 8 domains (vision, hearing, speech, ambulation, dexterity, emotion, cognition, and pain) and was developed to address issues in the HUI2 classification system which has 7 domains (sensation, mobility, emotion, self-care, cognition, pain, and fertility), however the fertility domain is dropped when being used in paediatric populations.(22-24)The HUI 2/3 has current a range of validated recall options. These recall options include 'current' recall versions (e.g., "during the past 1 week", or "during the past 2 weeks", or "during the past 4 weeks") or a 'usual' recall version. The usual recall version, which asks the participant to respond based on their usual health, was used for this study. Respondents are asked to rate the severity of each item

on a 4- to 6-point scale. The HUI 2/3 is HUI is recommended for use in children 5 years or older, however, some studies have used the instrument in children as young as 1 year old.(25)

7.1.8.EQ-5D-5L

The EQ-5D-5L is a 5-item generic adult HRQoL instrument with 5 item response levels, a 'today' recall period, covering 5 dimensions: mobility, self-care, usual activities, pain/discomfort, and anxiety/depression.(26) Respondents are asked to rate the severity of each item on a 5-point scale Although the EQ-5D-5L is generally self-report, a proxy report version is available. The EQ-5D-5L was adapted from the 3-level version, the EQ-5D-3L.(26)

7.2. Health condition-specific instruments

Due to the survey nature of this study, all health condition-specific instruments were required to be self or carer-reported (as opposed to clinician-reported or interview format). Additionally, the following criteria were applied to guide the choice of health condition-specific instrument: 1) well validated for children, 2) quality of life measure, 3) functional impairment measure, and 4) symptom measure. For example, if a condition-specific quality of life measure had been validated in children, this would be selected over a condition-specific functional or symptom measure that has been validated in children. Where there was ambiguity about the best choice, clinical experts were consulted.

7.2.1.ADHD

The Strengths and Weaknesses of Attention-Deficit/Hyperactivity Disorder Symptoms and Normal Behaviour Scale (SWAN) is a proxy-report ADHD symptom scale used as the health condition specific instrument for the ADHD group in this study, Sample 3a.(27) The SWAN has 18 items, 7 item response levels, a 1 month recall period, and covers 3 symptom areas: inattention, hyperactivity, and impulsivity.(27) The SWAN has been validated in children aged 6 to 18 years, however, has been used in children as young as 4 years.(28)

7.2.2.Anxiety or depression

The Revised Children's Anxiety and Depression Scale (RCADS-25) is a proxy or self-report anxiety and depression symptom scale used as the health condition specific instrument for the anxiety and depression group in this study, Sample 3b.(29) The RCADS-25 has 25 items, 4 item response levels, no specified recall period, and covers 6 domains: generalized anxiety disorder, major depressive disorder, obsessive compulsive disorder, panic disorder, separation anxiety disorder and social phobia.(29) The RCADS-25 has been validated in children aged 7 to 18 years.(29)

7.2.3.ASD

The KIDSCREEN-27 is a proxy or self-report generic paediatric HRQoL instrument used as the health condition specific instrument for the ASD group in this study, Sample 3c.(30, 31) Although the KIDSCREEN-27 is a generic HRQoL instrument, it was chosen as the health condition-specific instrument for the ASD group as no appropriate ASD-specific HRQoL instrument or symptom scale was available and the KIDSCREEN-27 has previously been recommended as a robust HRQoL instrument in children with ASD. The KIDSCREEN-17 has 27 items, 5 item response levels, a 1 week recall period and covers 5 domains: physical wellbeing, psychological wellbeing, autonomy/ parent relation, peer/social support, and school environment.(30, 31) The KIDSCREEN-27 is designed for use in children aged 8 to 18 years.(31)

7.2.4.Asthma

The PedsQL Asthma Module Version 3 is a proxy or self-report asthma paediatric HRQoL instrument used as the health condition specific instrument for the asthma group in this study, Sample 3d.(32) The PedsQL Asthma Module has 28 items (26 items in the 2–4-year-old version), 5 item response levels, a 1 month recall period, and covers 4 domains: asthma, treatment, worry, and communication.(32) The PedsQL asthma module has validated versions available for children aged 2 to 18 years.(32)

7.2.5. Eating disorders

The Eating Disorder Quality of Life Scale (EDQLS) is an adolescent and adult eating disorder specific quality of life instrument used as the health condition specific instrument for the eating disorder group in this study, Sample 3e.(33) The EDQLS has 40 items, 5 item response levels, a 1 week recall period, covering 12 domains: cognitive, education/vocation, family and close relationships, relationships with others, future outlook, appearance, leisure, psychological, emotional, values and beliefs, physical, and eating.(33) A self-reported version of the EDQLS has been validated in adolescents and adults aged 14-60 years old.(33) A proxy version was generated for the purposes of this study (see Section 7.4 for further information).

7.2.6. Epilepsy

The Quality of Life in Childhood Epilepsy Questionnaire (QOLCE-16) is a proxy report epilepsy specific paediatric HRQoL instrument used as the health condition specific instrument for the epilepsy group in this study, Sample 3f.(34) The QOLCE-16 has 16 items, 6 item response levels, a 4 week recall period, covering 4 domains: cognitive functioning, emotional functioning, social functioning, and physical functioning.(34) The QOLCE-16 has been validated in children with epilepsy aged 4-12 years, however, the QOLCE-57 has been validated in children up to the age of 18 years.(34)

7.2.7. Recurrent abdominal pain

Two pain visual analogue scales (VASs) asking about pain today and pain at last pain episode were used as the health condition specific instrument for the recurrent abdominal pain group in this study, Sample 3g. The pain VAS scales were adapted from the Paediatric Pain Questionnaire.(35)

7.2.8. Sleep problems

The Sleep Disturbance Scale for Children (SDSC) is a proxy report paediatric sleep disturbances and sleep behaviour instrument used as the health condition specific instrument for the sleep problem group in this study, Sample 3h.(36) The SDSC has 26 items, 5 item response levels, a 6 month recall period, and covers 6 domains: parasomnias, difficulty in initiating and maintaining sleep, sleep disordered breathing, disorders of excessive somnolence, sleep hyperhydrosis and non-restorative sleep.(36) The SDSC is validated in children aged 6 to 16 years, however, it has been used in children as young as 3 years.(37)

7.2.9. Tooth problems

The Child Perceptions Questionnaire (CPQ) 11-14 short form is paediatric oral HRQoL instrument used as the health condition specific instrument for the tooth problem group in this study, Sample 3i.(38, 39) The CPQ 11-14 short form has 16 items, 5 item response levels, a 3 month recall period, and covers 4 domains: oral symptoms, functional limitations, emotional well-being, and social well-being.(38, 39) The CPQ 11-14 was designed and validated in children aged 11 to 14 years, however, evidence suggests it may be applicable in children as young as 5.(40) A proxy version was generated for the purposes of this study (see Section 7.4 for further information).

7.2.10. Type 1 diabetes

The PedsQL diabetes module version 3.2 is a paediatric diabetes HRQoL instrument used as the health condition specific instrument for the type 1 diabetes group in this study, Sample 3j.(41) The PedsQL diabetes module has 33 items, 5 response levels, a one month recall period, and covers 5 domains: symptoms, treatment barriers, treatment adherence, worry, and communication. The PedsQL diabetes module was designed and validated in children aged 2-25 years (see Section 7.4 for further information).(41)

7.2.11. Wetting problems

The PinQ is a paediatric bladder dysfunction HRQoL instrument used as the health condition specific instrument for the wetting problems group in this study, Sample 3k.(42, 43) The PinQ has 20 items, 5

response levels, no recall period, and covers two sub-scales: intrinsic and extrinsic. (42, 43) The PinQ was designed and validated in children aged 6-17 years (see Section 7.4 for further information). (42, 43)

7.3. Other instruments and survey questions

7.3.1.EQ-HWB-S

The EuroQol health and wellbeing short form (EQ-HWB-S) is an instrument assessing the impact of health and wellbeing being as a care recipient or caregiver.(44) The EQ-WHB-S has 9 items, 7 day recall period, and covers 8 domains: mobility, usual activities, energy, cognition, social relationships, control, anxiety/depression, and pain.(44, 45) A carer quality of life instrument was included in the study following advice from the study Consumer Advisory Group who noted the strong relationship between child and carer QoL in children who have chronic conditions. The EQ-HWB-S was chosen as the carer quality of life instrument to include in the study because it is a promising new instrument that requires further validation work.

7.3.2.SDQ

The strengths and difficulties questionnaire (SDQ) is a paediatric behavioural screening instrument.(46, 47) The SDQ has 25 items, 3 item response levels, a 1-month recall, and covers 5 domains: emotional symptoms, conduct problems, hyperactivity/inattention problems, peer relationship problems, and prosocial behaviour.(46, 47) Validated versions of the SDQ are available for children aged 2–17 years, with self-report available for children aged 11 years and older.(46, 47) The SDQ was included in the study to capture emotional wellbeing of child participants to enable the performance of HRQoL instruments to be compared across validated scales of emotional wellbeing.

7.3.3.Other survey questions

A core set of demographic questions was included in the initial survey and completed by the caregiver. Where possible, demographic questions were adapted from LSAC to allow for comparison with a nationally representative sample.(5)

For each health condition group in Sample 3, several health condition severity questions were added to the initial survey to help approximate self or carer reported health condition severity (see Appendix Table 3). Health condition severity questions were designed with clinical experts to be no more than 3 questions, where possible, questions were derived from previous research studies.

Questions regarding the impact of COVID-19 on both caregivers and children were added to initial and follow-up surveys. As the COVID-19 impact questions were added after recruitment for Sample 1 had begun and prior to recruitment for Sample 2 and 3 beginning, all online panel samples (Samples 2 and 3) received the COVID-19 impact questions. However, only a portion of the sample recruited via hospital (Sample 1) received the questions. These questions were designed to allow for testing of potential self-reported COVID-19 impacts on HRQoL and to aid with generalisability of results considering data were collected during periods of pandemic.

The following additional questions were added to the follow-up survey that were not in the initial survey to capture any change in health since the initial survey was completed:

- Caregiver report of child's change in general health between initial and follow-up survey, adapted from similar SF-36 question.(48)
- If relevant, caregiver report of child's change in main health condition between initial and follow-up survey, adapted from similar SF-36 question.(48)
- Caregiver report of any major health event between initial and follow-up survey and if this event made the child's health better worse or it had no change. Major health events asked about included new treatment or therapy, new medication, new accident or injury, new condition diagnosed, new illness, unplanned doctor visit, unplanned hospital visit.

7.4. Survey adaptations

Some minor adaptations were made to the wording of some health condition-specific instruments, these are outlined below in Table 4. Where a health condition-specific instrument only had a self-report version (CPQ 11-14 and EDQLS), a proxy-report version was generated for the purpose of this study. A proxy-report version was required for this study as the survey was designed so that a caregiver could proxy-report the entire survey if they felt the child was not currently able to self-report their HRQoL or the child was younger than 7 years of age.

Table 4. Health condition-specific instruments adaptations

Health condition-specific sample	Instrument	Report type	Adaptations
3a. ADHD	SWAN	Proxy only	Adapted wording of questions with permission from developer. Adaptations were made to ensure the wording of the instrument was appropriate for a caregiver to understand.
3b. Anxiety or depression	RCADS-25	Proxy and self-report	-
3c. ASD	KIDSCREEN-27	Proxy and self-report	-
3d. Asthma	PedsQL asthma module	Proxy and self-report	-
3e. Eating disorder	EDQLS	Proxy (adapted) and self-report (original)	Generated proxy version for the purpose of this study from self-report version.
3f. Epilepsy	QOLCE-16	Proxy only	-
3g. Recurrent abdominal pain	Pain VAS	Proxy and self-report	Two pain VAS' adapted from Paediatric Pain Questionnaire. (35)
3h. Sleep problems	SDSC	Proxy only	-
3i. Tooth problems	CPQ-11-14	Proxy (adapted) and self-report (original)	Generated proxy version for the purpose of this study from self-report version.
3j. Type 1 diabetes	PedsQL diabetes module version 3.2	Proxy and self-report	-
3k. Wetting problems	PinQ	Proxy (adapted) and self-report (original)	Generated proxy version for the purpose of this study as per developer instructions.(42)

8. Survey development, piloting, and testing

Six rounds of survey piloting and testing were conducted with colleagues, consumer advisors, associate investigators, decision makers, caregivers, and children prior to the final survey being launched. Survey piloting was instrumental in improving the design, length, and wording included in the survey. Additionally, all survey pathways were quality checked prior to launch to ensure no survey errors.

9. Data collection and survey administration

Participants completed surveys online via REDCap (Research Electronic Data Capture) hosted at The Murdoch Children's Research Institute (MCRI) (www.redcap.mcri.edu.au, accessed on 14th June 2022). All participants

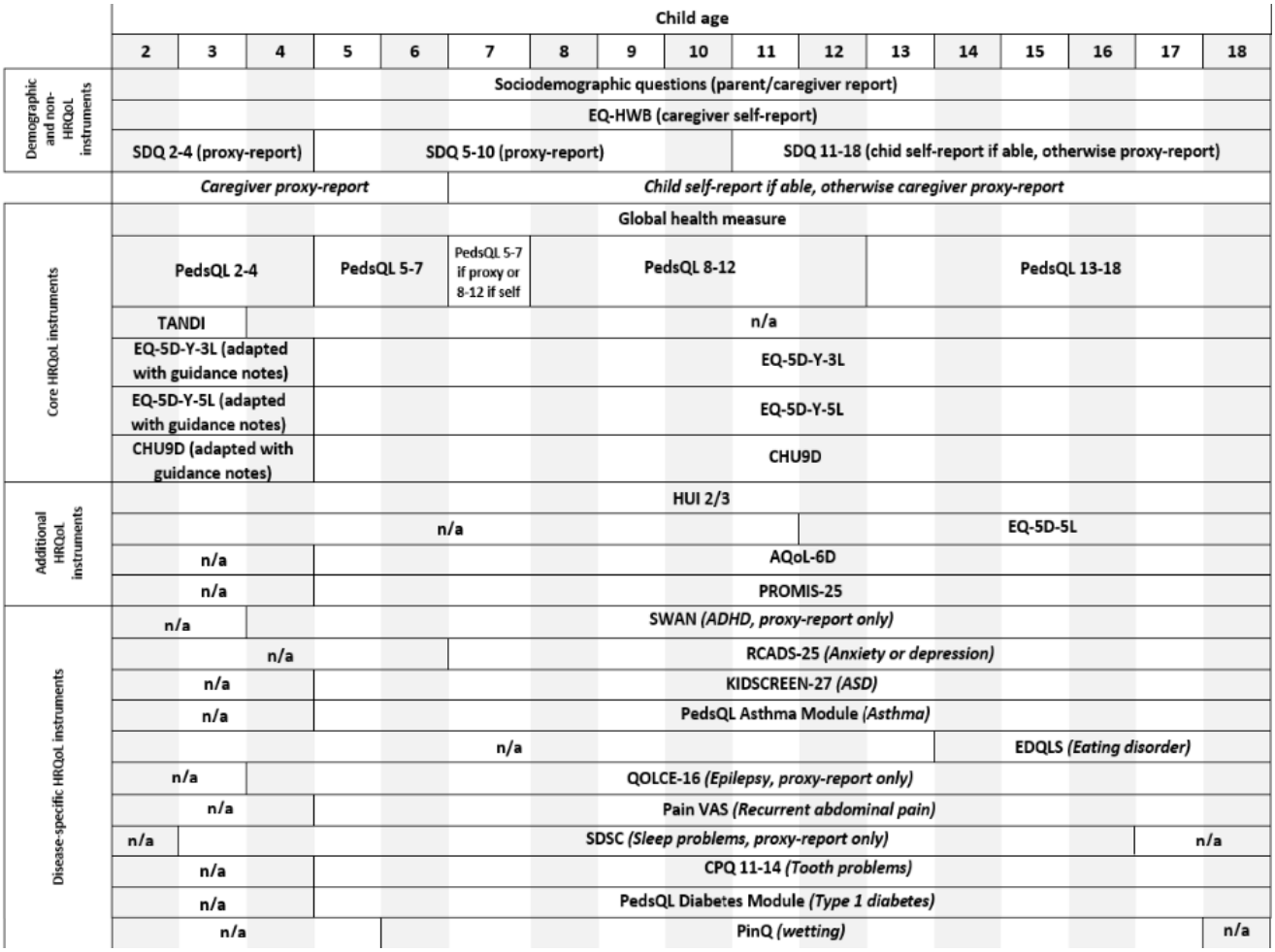
received a core set of questions and instruments, some samples also received additional instruments, instrument blocks, and/or questions. The schedule of instruments for each sample collected at the two time points is outlined in Table 5. As some instruments have different versions for different child ages, participants were allocated to receive the instrument version most appropriate for their child's age (see Figure 1). Children aged 7 years and older were asked to self-report the HRQoL instruments and health condition instruments if a self-report version was available and if the child was considered currently able to report on questions about their health and wellbeing by their caregiver.

The order of the demographics, EQ-HWB-S, SDQ, core HRQoL instruments, additional HRQoL instruments, and health condition-specific instruments was decided based on two criteria: 1) an order that minimises the survey being handed back and forth between caregivers and children, with only one handover point occurring if the child is 7 years or older and able to self-report, and 2) the order reflects the priority of questions as decided by the study team. For the initial survey, participants were first screened and consented into the survey, following this, participants completed the demographic questions (including health condition severity questions if relevant, see Appendix Table 3). The additional HRQoL instrument blocks were always presented to participants after the core HRQoL instruments. Condition-specific instruments that had both a proxy and self-report version available were always presented after the core and additional HRQoL instruments. Where only proxy-report versions of the condition-specific instruments were available, the condition-specific instrument was presented prior to the core HRQoL instruments, this was to prevent the caregiver and child having to hand the survey back and forth. The follow-up survey followed the same structure as the initial survey albeit with a smaller number of required instruments (see Table 5). Demographic questions included in the survey are summarised in Appendix Table 4.

Within the core HRQoL instruments, the order of instruments was randomized to minimize order and survey fatigue effects. Additionally, the EQ-5D-Y-3L, EQ-5D-Y-5L, and, if relevant the EQ-5D-5L, were presented with another HRQoL instrument separating them, given their similarities. Participants received the same order of instruments for both the initial and follow-up survey.

Except for Sample 2b (the online panel general population sample with a 2-day follow-up), all samples were followed up at 4 weeks and received up to three reminders at consistent time intervals. Sample 2b received a 2-day follow-up timeframe to allow for test-retest analysis. For consistency, 4-week follow-up time was decided for the remainder of the samples, this longer follow-up time was chosen to allow for analysis of instrument responsiveness to perceived change in health between time points, which is a key gap in the current literature.(4) A 4-week follow-up time was considered enough time to for children who were acutely unwell at the time of recruitment to recover before the follow-up survey and a short enough time frame to minimise attrition.

Figure 1. Instruments and questions by child age.



Abbreviations: AQoL-6D Assessment of Quality of Life, CHU9D Child Health Utility, CPQ Child Perceptions Questionnaire, RCADS Revised Children's Anxiety and Depression Scale, EDQLS Eating Disorder Quality of Life Scale, EQ-HWB EQ Health and Wellbeing Short Version, EQ-5D-Y EQ-5D Youth, HRQoL health-related quality of life, HUI2/3 Health Utilities Index Mark 2/3, PedsQL Paediatric Quality of Life Inventory, PROMIS-25 Patient-Reported Outcome Measurement Information System 25, QOLCE Quality of Life in Childhood Epilepsy Questionnaire, VAS Visual Analog Scale, SDSC Sleep Disturbance Scale for Children, SDQ Strengths and Difficulties Questionnaire, SWAN Strengths and Weaknesses of Attention-Deficit/Hyperactivity Disorder Symptoms and Normal Behavior Scale, TANDI Toddler and Infant Questionnaire.

Table 5: summary of instruments and questions by sample

Instrument	Sample 1, Recruited via hospital		Sample 2, General population		Sample 3, Health condition-specific groups	
	Initial survey	Follow-up survey	Initial survey	Follow-up survey	Initial survey	Follow-up survey
Demographic and non-HRQoL instruments						
Informed Consent	x		x		x	
Demographic Information	x		x		x	
EQ-HWB	x		x		x	
SDQ	x		x		x	
Core HRQoL instruments						
PedsQL	x	x	x	x	x	x
TANDI (if <=3yrs)	x	x	x	x	x	x
EQ-5D-Y-3L (inc VAS) & 5L original (if >= 5 years)	x	x	x	x	x	x
EQ-5D-Y-3L (inc VAS) & 5L adapted (if <=4 years)	x	x				
EQ-5D-Y-3L original (inc VAS) & adapted OR EQ-5D-Y-5L original (inc VAS) & adapted (if <=4 years)			x*	x*	x*	x*
CHU9D	x	x	x	x	x	x
Global Health Measure	x	x	x	x	x	x
Additional HRQoL instruments						
AQoL-6D (if =>5yrs)			x*	x*	x*	x*
HUI2 (if >=2yrs) & EQ-5D-5L (if >=12yrs)			x*	x*	x*	x*
PROMIS-25 (if =>5yrs)			x*	x*	x*	x*
Health condition-specific instruments						
SWAN *(ADHD)*					x*	
RCADS-25 *(Anxiety or depression)*					x*	
KIDSCREEN-27 *(ASD)*					x*	
PedsQL Asthma Module *(Asthma)*					x*	
EDQLS *(Eating disorder)*					x*	
QOLCE-16 *(Epilepsy)*					x*	
Pain VAS *(Recurrent abdominal pain)*					x*	
SDSC *(Sleep problems)*					x*	
CPQ 11-14 *(Tooth problems)*					x*	
PedsQL Diabetes Module *(Type 1 diabetes)*					x*	
PinQ *(wetting problems)*					x*	

*X- indicates the instrument will be collected from the sample/time point. *Participant will only receive, if allocated, instrument based on health condition group, and/or randomization to receive additional instrument, and/or randomization to receive EQ-5D-Y 3L original and adapted or EQ-5D-Y 5L original and adapted. Abbreviations: AQoL-6D Assessment of Quality of Life, CHU9D Child Health Utility, CPQ Child Perceptions Questionnaire, RCADS Revised Children's Anxiety and Depression Scale, EDQLS Eating Disorder Quality of Life Scale, EQ-HWB EQ Health and Wellbeing Short Version, EQ-5D-Y EQ-5D Youth, HRQoL health-related quality of life, HUI2/3 Health Utilities Index Mark 2/3, PedsQL Paediatric Quality of Life Inventory, PROMIS-25 Patient-Reported Outcome Measurement Information System 25, QOLCE Quality of Life in Childhood Epilepsy Questionnaire, VAS Visual Analog Scale, SDSC Sleep Disturbance Scale for Children, SDQ Strengths and Difficulties Questionnaire, SWAN Strengths and Weaknesses of Attention-Deficit/Hyperactivity Disorder Symptoms and Normal Behavior Scale, TANDI Toddler and Infant Questionnaire.*

10. Participant reimbursement

Participants from Sample 1, sample recruited via hospital, were reimbursed with a $15 online gift voucher once they had completed the follow-up survey. Participants from Samples 2 and 3, the online panel samples,

were reimbursed for their time by Pureprofile Australia. Participants in both the online panel general population sample and health condition-specific groups sample (Samples 2 and 3) were reimbursed $3-$5 for completing the initial survey and $3-$4 for the second. Total reimbursement ranged from $6-$9.

11. Quotas

For the online panel general population sample (Sample 2), participants were selected based on quotas for age. Other characteristics such as child gender, family income, regionality, state, number of caregivers in the home, and caregiver education were monitored to ensure a diverse sample was obtained from the online panel.

12. Quality monitoring

Minimum quality eligibility criteria were applied to each sample to ensure that legitimate responses were being obtained. Respondents were ineligible if they:

- Child age outside of eligibility
- Caregiver age less than 18 years
- Survey completed in less than 1/3 of the median time, for both initial and follow-up surveys
- Caregiver not reporting child health condition they have screened for in initial survey condition list (Sample 3 only, online panel condition group sample)
- Child age reported at follow-up is not consistent with child age reported in initial survey.

Several additional quality criteria were monitored to ensure these occurrences were minimal, however, participants were ineligible based on these criteria. Additional monitoring criteria were:

- Child self-report during school hours
- Child does not screen as having a special healthcare need but does report having a condition where this would be expected

Duplicate records in the online panel samples (Sample 2 and 3) were identified using the unique online panel identifier. Duplicate records in the hospital sample (Sample 1) were identified using the email address entered by caregivers in the survey. In deciding which record to keep, the following criteria was applied:

1) Keep the most complete record.
2) If both records were equally complete, keep the record that was completed first.

A total of 99 respondents were removed from the current data cut for being a duplicate survey response.

Where a caregiver had completed the survey more than once for different children, this was noted so duplicate caregivers could be removed from relevant analysis such as for the EQ-HWB.

13. Decision to close samples

Samples were closed if target sample sizes were reached.(1) If sample quotas were not reached, such as for eating disorder (3f) and epilepsy (3g), samples were closed after all additional avenues of recruitment had been exhausted.

14. Psychometric Analysis Guide

The purpose of this statistical analysis plan is to provide an overview of key sub-groups, statistical tests, assumptions, and thresholds for interpretation alongside the justification for these decisions. Manuscripts publishing psychometric analysis using PMIC study data will be guided by this statistical analysis plan. Given

the breadth of manuscripts and corresponding statistical analysis that will be produced using PMIC study data, this psychometric analysis guide is only intended as an overview and additional details will be published alongside corresponding manuscripts. This analysis guide may be used for the analysis of any instrument (see Section 7) and/or population (see Section 5) included in the PMIC study, the instrument(s) and population(s) included in each analysis will be published in each manuscript. Additionally, this will be a living psychometric analysis guide and other methods such as Item Response Theory (IRT), factor analysis, and structural equation modelling, will be added over time as manuscripts progress.

14.1. Level of analysis and scoring

The performance of one or more of the instruments(s) included in the PMIC study can be assessed and compared at the item level, the domain level, and/or the total instrument level (see below for further details). A manuscript may assess the performance of one or more instrument(s) at any of these levels, depending on the focus of the manuscript. Each published manuscript will outline and provide further details on the level of analysis and if applicable, any instrument scoring used, however, a brief overview of how this may be done is detailed below. The choice of analytical test and/or corresponding threshold may vary depending on the level of analysis and hence this has been described below for each test if applicable.

Item level

- Description: When assessed at the item level, instrument items will be assessed using the original ordinal item levels. See Appendix Table 2 for the number of levels for each instrument.
- Purpose: The purpose of assessing instruments at the item level is to understand and compare the performance of individual instrument items. This may also be useful when a difference in performance of instrument items may be hypothesised for certain sub-groups.
- Scoring: No scoring or transformations will be applied unless specified by the instrument instructions from instrument developers.

Domain level

- Description: When assessed at the domain level, instrument domains will be assessed using the domains specified by instrument developers. See Appendix Table 2 for the domains specified by instrument developers for each instrument.
- Purpose: The purpose of assessing instruments at the domain level is to understand and compare the performance of instrument domains. This may also be useful when a difference in performance of instrument domains may be hypothesised for certain sub-groups.
- Scoring: Instrument domains will be scored according to the methodology of the instrument developer unless another method is justified for the focus of the manuscript. Each manuscript including analysis at the domain level will specify the method for calculating domain scores.

EQ Visual Analogue Scale (VAS) level

- Description: This level of analysis only applies to EuroQol instruments (EQ-TIPs, EQ-5D-Y-3L, EQ-5D-Y-5L, and EQ-5D-5L). The VAS is a general health scale ranging from 0 to 100.
- Purpose: The purpose of assessing EuroQol instruments at the VAS level is to understand and compare the performance of the VAS. The VAS is not captured by assessing the item, domain or total instrument and hence has been included as a separate level of analysis.

Total instrument level

- Description: When assessed at the instrument level, all items in an instrument will be combined to generate an overall instrument total score.
- Purpose: The purpose of assessing instruments at the total instrument level is to provide an overall score that can be compared to an overall score to other instruments. This may be

particularly useful when a large number of instruments are being compared in the one manuscript or as an additional level of analysis to instrument and domains levels for a manuscript that may be focused on a smaller number of instruments.

- Scoring: There are a range of scoring methods that may be applied to obtain a total instrument score. The method used to obtain a total instrument score will be specified and justified in each manuscript. The main methods used to obtain a total instrument score are a level sum scoring approach or a utility scoring approach (see details below). Some instruments are designed to be scored in a level sum score approach, such as the PedsQL. However, other instruments, such as the EQ-5D-Y or CHU9D are designed to be scored using the utility scoring approach. The main concern with using a level sum score approach on instruments designed to be scored using a utility scoring approach is that it may miss the complexity and relative importance of different domains captured by utility scores, according to Devlin et al 2020.(49) However, it is important to note that despite this, there is support for using both the level sum and utility scoring approaches in different contexts, according to Feng et al 2022.(50) Given the lack of consensus regarding the best total scoring approach for psychometric testing, the intension of the PMIC study is to eventually look at and compare instrument performance using both methods and not to just rely on one or the other to make final conclusions about instrument performance. As more manuscripts are published using PMIC study data, the study team will continue to refer to the literature regarding this as it evolves.
    - Level sum score: Unless otherwise specified, the level sum score is calculated by summing together all raw instrument item responses for that instrument. NB: For the PedsQL a level sum score is calculated by reverse scoring and linearly transforming raw instrument item responses.
    - Utility score: Value sets (weightings from the general public) are applied to give a total score between 0 and 1. Each manuscript will specify the choice of value set, a justification for this and any implications regarding this choice (i.e., different countries). The wider QUOKKA Research Program has been funded to create value sets for some of the instruments included in the PMIC Study, so in the coming few years we expect the number of value sets available for these instruments to increase and we want to retain the capacity to use these as they emerge. We would prioritise Australian value sets if available. It is anticipated the PMIC study will become a valuable dataset for testing value sets as they emerge.

14.2. Sub-group categories

Where applicable, sub-group analysis may be completed to understand if instrument performance varies for certain sub-groups. Additional sub-groups not listed may also be explored if relevant to certain manuscripts, the details of these additional sub-groups will be published in the corresponding manuscript.

Child age

- Categorisation: 2-4 years, 5-12 years, 13-18 years (or 13-16 years depending on instrument of focus)
- Justification: Aligns with Australian developmental milestones: preschool (including kindergarten and day care), primary school and high school.
- Note: Different age bands may be used in individual manuscripts to align with the focus of the manuscript or the different starting ages of some of the online condition groups (See Table 2 in Section 5 for a summary of child ages for each condition). Self-report by the child starts at age 7 years, so some manuscripts may consider starting an age band at age 7 years

to have greater consistency in the younger age group (See Figure 1 in Section 9 for a summary of instruments by child age).

Child health

By Sample

- Categorisation: participants recruited via a tertiary paediatric hospital (Sample 1), participants from the online panel general population sample (Sample 2), and participants from the online panel condition groups sample (Sample 3)). Additionally, participants may be categorised by the 11 conditions within Sample 3: ADHD, anxiety and/or depression, ASD, asthma, eating disorder, epilepsy, recurrent abdominal pain, sleep problems, tooth problems, type 1 diabetes, and wetting. See Section 5 for further details on these samples.
- Justification: Participants were recruited and screened into these samples with either a self-reported known health condition (online panel health condition groups, Sample 3) or through receiving care from a tertiary paediatric hospital (hospital sample, Sample 1). See Section 5 for further details on sample screening and Section 6 for further details on sample recruitment.

By Child Special Healthcare Needs Screener (CSHNS)

- Categorisation: Children with a special healthcare need versus children without as per CSHNS (from any sample).
- Justification: Validated screening tool.(51)

By global health question

- Categorisation: Global health reported as 1) excellent, 2) very good, 3) good, 4) fair and poor
- Justification: Capacity to capture more groupings of health rather than a binary categorisation. However, may be limited by sample size.
- Note: Only to be used as an alternative to the above health status categorisations if there is a good justification, or to be used as a supplementary sub-group to the above categorisations.

By severity questions or condition specific instruments (only for online panel condition groups sample, Sample 3)

- Categorisation: Cut offs points will be based on the literature, using the corresponding condition specific instruments relevant to the condition group (See Section 7 for details on condition specific instruments). Severity may also be categorised according to condition-specific questions detailed in Appendix Table 3.
- Justification: Capacity to capture differences within health condition groups.
- Note: Only to be used for the online panel condition groups sample, Sample 3.

Report type

- Categorisation: self-report and proxy report.
- Justification: Capacity to capture differences in instrument performance between child self-report and caregiver proxy report.

Caregiver education

- Categorisation: Highest level of caregiver education is bachelor degree or above versus not.
- Justification: Based on response distribution.
- Note: Optional additional sub-group, may be explored but not required.

Socioeconomic Status (SES)

- Categorisation: SES will be categorised according to postal code of residence using the index of relative socioeconomic advantage and disadvantage.
- Justification: As per recommendation of Australian Bureau of Statistics.
- Note: Optional additional sub-group, may be explored but not required.

14.3. Acceptability and feasibility

Variables used to assess:

Instrument acceptability and feasibility will be measured by assessing participant self-reported difficulty completing each instrument and time to complete each instrument. Participants were required to answer all questions in the PMIC survey and hence item missingness is not able to be assessed.

Self-reported difficulty of each instrument was measured after each instrument, rated on a 5-point scale from 1 'very difficult' to 5 'very easy'. Time to complete each instrument was automatically captured via the online survey administration platform.

Analytical approach:

Self-reported difficulty will be assessed descriptively and differences in difficulty across instruments were assessed using Pearson's chi-squared test. Time to complete each instrument will be assessed descriptively.

14.4. Ceiling and floor effects

Item level:

Analysis will simply describe the distribution or frequencies across response levels without applying thresholds or criteria. Discussion of ceiling or floor effects at the item level will only be discussed if relevant to the focus of the manuscript, for example, testing of new or experimental instruments. We will look at ceiling effects where possible in condition specific samples or sub-groups of children where children are expected to be more unwell.

Instrument level:

Ceiling effect flag >15% of respondents reporting lowest severity or frequency of category across all items (e.g., 'No problems').

Floor effect flag >15% of respondents reporting the highest severity or frequency of category across all items (e.g., 'Extreme problems/ unable to').

If more than 15% of participants report the lowest or highest severity level for all instrument items, then this is a flag for a ceiling or floor effect that requires further exploration and discussion. Such exploration and discussion should consider the characteristics of the sample. For example, more than 15% of participants reporting the level 1 for all instrument items might be appropriate in a general population sample. However, more than 15% of participants reporting level 1 for all instrument items in a clinical sample and/or sample of children with a chronic or ongoing health condition could be considered a ceiling effect issue (see child health status sub-group categories listed above). When ceiling effects are being assessed in participants from the online panel condition groups sample (Sample 3), sub-group analysis may be completed to assess ceiling effects by severity of condition, allowing for differentiation between children with a health condition that is well-managed versus children with a condition that is not currently well managed.

The 15% threshold was derived from the following key sources: 1) Terwee et al 2007 and 2) McHorney et al 1994. (52, 53)

14.5. Test-retest reliability

Sample

Participants from the online panel general population sample randomly allocated to receive follow-up survey 2 days after initial survey will be used for the test-retest analysis. Although participants were sent a reminder to complete the follow-up survey 2 days after completing the actual time between initial and follow-up survey completion varied, with a median of 3 days (IQR 2-4.5 days). Given the number of instruments participants are asked to complete in the initial survey, it was considered very unlikely participants would recall their responses at 2-days. Additionally, the research team were aware that participants were unlikely to all complete the follow-up survey on the day they were first reminded, hence wanted to begin reminders at 2 days to ensure follow-up surveys would largely be completed within 7 days. Finally, participants are only included if they do not report change in health at follow-up. Change is based on health status as measured by global change in health or change in health due to a stated health condition, since completing the last survey (see Section 7.3.3 for further details on change in health questions).

If the 2-day test-retest sample is too small for certain sub-group or specific population analysis, the use of the 4-week follow-up group may be considered. Again, this would only include participants who reported no change in health. If both the 2-day and 4-week follow-up time points are included this will be specified in the manuscript including clarity on the sample sizes used at each follow-up time points.

Given the recall period of instruments varies, each manuscript may discuss how this might impact on test-retest reliability.

Analytical approach and thresholds:

Item level

- ***Weighted Kappa (linear weighted)*** is the preferred choice of statistical test for ordinal outcomes (e.g., item levels). A weighted Kappa of 0.2 indicates poor agreement, 0.21–0.40 indicates fair agreement, 0.41–0.6 indicates moderate agreement, 0.61–0.80 indicates substantial agreement and >0.81 indicates almost perfect agreement. This threshold was derived from Landis et al 1997.(54) The decision to use linear weighting was derived from Al-Janabi et al 2015.(55)

Transforming a numerical outcome into an ordinal outcome for reliability assessment –Using the VAS categorisation as test-retest method considered experimental but potentially useful. As the EQ VAS is scaled from 0 to 100 even small changes might have a big impact when showing and comparing health changes therefore to avoid attributing too much weight to relatively small VAS changes, e.g., from 70 to 72, categorisation can be used. It is expected that data will be clustered around 10s and 5s responses. If this is the case, it is suggested EQ VAS results can be categorized in 10 groups and then kappa coefficients can be used to compare the results of initial survey and follow up (2 days). This can be added as method in addition of comparing instruments level sum scores as same level sum score might present different health states (e.g., EQ-5D health status 11112 has the same level sum score of 12111 of 6).

Instrument or domain level

- ***Intraclass correlation coefficient (ICC)*** with corresponding 95% confidence intervals (CIs) is the preferred choice of statistical test for numerical outcomes. ICCs will be calculated using a two-way mixed-effects model for a single instrument, based on absolute agreement. These model parameters are derived from Koo et al 2016.(56) Whilst acknowledging no accepted thresholds exist for interpreting ICC results, Koo et al 2016 state that as a rule of thumb, ICC values <0.5 indicate poor reliability, 0.50-0.74 moderate reliability, 0.75-0.90 good reliability, >0.90 excellent reliability (56). These thresholds stem from the book 'Foundations of Clinical Research: Applications to Practice'.(57) Other thresholds also exist for ICC values, with Cicchetti 1994 stating that ICC values below 0.4 indicate poor agreement, 0.40-0.59 fair agreement, 0.60-0.74 good agreement, and 0.75-1 excellent agreement.(58) Primary analysis will be done using Koo et al 2016 thresholds, given the recency and clarity of how thresholds were derived, however, sensitivity analysis will be completed using the Cicchetti et al 1994 threshold.
- ***Bland Altman Plots*** are preferred for utility comparisons and will only be included as a supplementary test where the analysis is focussed on test-retest reliability. Bland Altman Plots will be visually inspected and described. Half widths of the 95% limits of agreement will be calculated using 1.96 SD to define the 'limits' within which 95% of the differences should lie. Interpretation and thresholds were derived from Giavarina 2015.(59)

14.6. Known group validity

Key known groups:

- The following are known groups hypothesised to have poorer HRQoL compared to their counterparts and may be used to assess the known group validity of instruments. Determining if instrument(s) demonstrate known group validity will require the assessment of both children who are 'well' compared to 'unwell' and children who have 'mild severity' of a condition compared to children who have 'moderate or severe severity' of a condition, hence a range of known groups are described below. The known groups used for a given manuscript will depend on the focus of that manuscript.

  The strength of the PMIC study is that there are multiple ways in which we can assess known group validity. If the instrument can differentiate across multiple known groups, then that will be considered evidence of known group validity.

  Some known groups may be more appropriate for certain instruments (i.e., longer versus shorter instruments, generic versus condition specific instruments, and preference weighted versus not preference weighted instruments) and if relevant, this will be discussed in the manuscript. Additionally, some known group categorisations may allow for the well child group to be contaminated with unwell children and vice versa, hence, effects may be underestimated and if relevant, this will be discussed in the manuscript.

Children with a special healthcare need as per CSHCN:

- Categorisation: Children with a special healthcare need compared to children without a special healthcare need as per to CSHCN. See further details in Section 1.2 (Sub-groups).
- Justification: Expected that children with a special healthcare need as per the CSHCN screener will have poorer HRQoL compared to children without a special healthcare need. Children with a special healthcare need have previously been demonstrated to have poorer HRQoL compared to their counterparts by Chen et al 2011.(60)

Children with a chronic health condition (lasted or expected to last more than 6 months):

- Categorisation: Children whose caregiver has reported they have one or more medical condition(s) or disability/disabilities that is expected to last or has lasted more than 6 months compared to children whose caregiver has not reported this (see Appendix Table 4, question 12).
- Justification: Expected that children with a chronic health condition that lasts or is expected to last more than 6 months would have poorer HRQoL compared to those without.

Children reported as having fair or poor health on the global health question:

- Categorisation: general health reported as fair or poor compared to general health reported as excellent, very good, or good compared. See further details in Section 14.2 (Sub-groups).
- Justification: Expected that children whose general health is reported as fair or poor will have poorer HRQoL compared to children whose general health is reported as good, very good or excellent.

Children with one of the conditions from the online panel condition group sample (Sample 3):

- Categorisation: Participants from the online panel condition groups sample (Sample 3) compared to children from the online panel general population sample (Sample 1). See further details in Section 14.2 (Sub-groups).
- Justification: These are children screened into the sample based on wording previously used to establish the presence of such a health condition. These conditions were selected for inclusion in the study sample as they have been previously shown to demonstrate decrements on HRQoL.

Condition severity known groups (only applicable to Sample 3):

- Categorisation: Known groups may be formed using the corresponding condition specific instruments relevant to the condition group (See Section 7 for details on condition specific instruments). Severity may also be categorised according to condition-specific questions detailed in Appendix Table 3. Cut offs points will be based on the literature, using the condition specific instruments relevant to each condition group. Specific cut points will be detailed as manuscripts emerge.
- Justification: Expected that children more severe presentations or more symptoms of a condition would have poorer HRQoL compared to those with less severity or less symptoms.

"Healthy" child reference group:

- In addition to the reference groups described in the categorisations above, additional sensitivity analysis may be completed using a "healthy" child reference group.
- Categorisation: Participants from the online panel general population sample (Sample 1), who do not report a chronic or ongoing health condition and who have an EQ VAS score of at least 70.
- Justification: As described above, some known group categorisations may allow for the well child reference group to be contaminated with unwell children, hence this "healthy" child reference group may be included as a sensitivity analysis to explore the effects of minimising such contamination. Categorisation is obtained from Richardson et al 2014.(61)

Additional known groups for sensitivity analysis:

The following are additional known groups that may be used in addition to key known groups as sensitivity analysis to assess the known group validity of instruments:

EQ VAS score ≤80:

- Categorisation: An EQ VAS cut point of 80 has been used previously in the literature by Peasgood et al 2022 for assessment of known group validity and hence this same cut point was applied.(62)
- Justification: It is expected that children with a lower EQ VAS score (≤80) will have poorer quality of life compared to children with a higher score (>80).
- Note: This EQ VAS cut point will not be used to assess the known group validity of the EQ-5D-Y-3L, EQ-5D-Y5L, or EQ-5D.

PedsQL total score cut points:

- Categorisation: Total score known group cut points for children expected to have poorer quality of life from Varni et al 2003.(63)
    - ≤74.2, based on child self-reported mean from a sample of children with chronic conditions.
    - ≤73.1, based on proxy reported mean from a sample of children with chronic conditions.
    - ≤69.7, based on one standard deviation below the child self-reported population mean for children aged 5-18 years.
    - ≤65.4 based on one standard deviation below the proxy-reported population mean for children aged 2-18 years.
    - ≤74.9 based on one standard deviation below the proxy-reported population mean for children aged 2-4 years.
- Justification: It is expected that children with a lower PedsQL total score would have poorer HRQoL compared to children with a higher score.
- Note: These PedsQL cut points will not be used to assess the known group validity of the PedsQL.

Analytical approach and thresholds

Instrument level

- ***Comparing mean difference for each group:*** A mean difference in the expected direction between groups is also considered an indication of known group validity. Mean differences may be calculated using T-test (p-value will be reported) and/or ANOVA (F statistic reported).
- ***Effect sizes estimated using Cohen's D:*** Effect sizes of 0.2-0.49 is considered small, 0.5-0.79 moderate, and ≥0.8 large. These thresholds are obtained from Cohen 1992.(64)

Item or domain level

- Assessment of known group validity level at the instrument item or domain level may be explored for new or experimental instruments. This will be assessed descriptively, by exploring the distribution of items between groups.

14.7. Convergent & divergent validity

Instruments are considered to display convergent validity if they are correlated with other instruments where we expect them to be correlated. Instruments are considered to display divergent validity if they are not correlated with other instruments where we expect them not to be correlated. As there is no gold standard instrument for measuring health related quality of life in children, convergent and divergent validity is only ever assessed between instruments.

Both convergent and divergent validity will be assessed in all instruments, however, divergent validity is considered particularly important when condition specific instruments are being considered, where we expect these instruments to be picking up on different constructs to the generic instruments.

Sample and instruments for comparison

The sample and instruments for comparison will vary depending on the focus of the manuscript. The below details help provide any overview of what may be covered; however, additional detail will be provided in each manuscript.

Analytical approach and thresholds:

- Spearman's correlation is recommended due to non-normally distributed data. A correlation of 0.1-0.29 is considered weak, 0.3-0.49 moderate, and ≥0.5 strong. These thresholds are obtained from Cohen 1992.(64)

Hypothesised correlations for convergent validity:

- ***Hypothesised item correlations (for all samples)*** were based on similarity of item content and set a priori where at least a moderate correlation is expected between corresponding generic instrument items. This involved 6 experts reporting where they hypothesised item combinations would have at least a moderate correlation based on similarity of item content. These reports were summarised into a final series of tables via a consensus approach whereby any item combinations reported by a study member but not by another were discussed as a team and finalised via consensus. For example, the PedsQL pain item is hypothesised to be at least moderately correlated with the EQ-5D-Y-3L pain item. See the green cells highlighted in Appendix 5 below for all generic instrument items hypothesised to be at least moderately correlated with one another. The actual Spearman correlations will be compared back to those hypothesised correlations to be assess for convergent validity.
- ***Hypothesised item pool construct correlations (for all samples)*** were set a priori by the study team. This involved 4 experts assessing each generic instrument and reporting which items, based on item content, related to the following common constructs: physical mobility, emotion, pain, daily routine, school/cognition, and social/relationships. These constructs were chosen as they are common constructs included in HRQoL instruments. By creating these item construct pools, we can compare if these common constructs are convergent across instruments. Although some instruments have existing domains that cover some of these constructs, these domains were largely ignored and only items agreed on by experts to be included in the construct pool were included. These reports were summarised into a final table via a consensus approach whereby any items reported by a study member for a given construct but not by another were discussed as a team and finalised via consensus. All items for each generic instrument considered to be related to each construct will be pooled together (using a sum score approach). It is hypothesised the item pool construct for each instrument will be at least moderately correlated with the item pool for the same construct for other generic instruments. See Appendix below for all generic instrument hypothesised item pool constructs (Appendix 6).
- ***Hypothesised correlations in condition-specific groups (for Sample 3)*** were more targeted towards that specific condition area. Hypothesised convergence between condition specific instruments and generic instruments will be reported in individual manuscripts.

Hypothesised non-correlations for divergent validity:

- ***Items hypothesised not to be correlated (for all samples)*** were based on dissimilarity of item content. These were set a priori, where corresponding generic instrument items were hypothesised not to be correlated. This involved 2 experts reporting where they hypothesised item combinations would not be correlated (or have a very weak correlation) based on similarity of item content. See the red cells highlighted in Appendix 5 below for all generic instrument items hypothesised not to be correlated

with one another. The actual Spearman correlations will be compared back to those hypothesised correlations to be assess for divergent validity.

- ***Hypothesised correlations in condition-specific groups (for Sample 3)*** were more targeted towards that specific condition area. Hypothesised divergence between condition specific instruments and generic instruments will be reported in individual manuscripts.

14.8. Responsiveness

Sample

Analysis will be restricted to participants allocated to receive the follow-up survey at four-weeks, this will enable participants enough time to meaningfully change health. Furthermore, only participants who report a change in health (using variables described below) will be included in responsiveness analysis.

Determining change in health:

Reported change in health – primary analysis

- Caregivers were asked to report the child's change in **general health** at follow-up since the initial survey over 5 levels: 1) much better, 2) somewhat better, 3) about the same, 4) somewhat worse, or 5) much worse. This was an adapted version of an SF-36 item. See Section 7.3.3 for further details. For responsiveness analysis, improved health is defined as a report of 1) much better and worsening health is defined as a response of 4) somewhat worse or 5) much worse (levels combined for sample size). It was decided that taking the more extreme categories that still had a large enough sample size would make this the clearest way to assess a true change in health.
- Caregivers were also asked to report change in child's **main health condition** (if they have one) at follow-up since initial survey over 5 levels: 1) much better, 2) somewhat better, 3) about the same, 4) somewhat worse, or 5) much worse. SF-36 item. See Section 7.3.3 for further details. For responsiveness analysis, improved health is defined as a report of 1) much better and worsening health is defined as a response of 4) somewhat worse or 5) much worse (levels combined for sample size).

Clinically important differences – supplementary analysis

- At this stage, the PedsQL is the only instrument with a known clinically important difference cut point. Hence, a difference of 4.4 in child self-reported and 4.5 in proxy reported PedsQL total score is considered a minimal clinically important difference, as per Varni et al 2003.(63)

Instrument level

Analytical approach and thresholds

- ***Effect sizes estimated using Standardised Response Mean (SRM***). Effect sizes of 0.2-0.49 were considered small, 0.5-0.79 moderate, and ≥0.8 large. These thresholds are obtained from Cohen 1992.(64)
- ***Comparing mean at initial and follow-up.*** The mean difference in the expected direction between initial and follow-up with a statistically significant is also considered an indication of responsiveness. Mean differences may be calculated using paired T-test (p-value will be reported) and/or ANOVA (F statistic reported).
- Other analysis such as ***Receiver Operating Characteristic (ROC) and Paretian classification*** may be used for manuscripts that have a focus on responsiveness.

14.9. Shannon index

The discriminatory power of an instrument refers to its ability to distinguish between different health states and detect changes in health status over time. The Shannon Index, by showing the distribution of responses in each dimension, provides a measure of its discriminatory power and informativity. In the context of measuring the informativity of instruments, the Shannon Index can be used to evaluate the amount of information captured by each instrument. To measure Shannon index, the formula bellow is used:

$$H' = -\sum_{i=1}^{L} p_i \log_2 \quad p_i$$

Where H' represents the absolute amount of informativity captured, L is the number of possible levels, and ,pi$pi = n_i \quad / _N$ where n_i is the observed number of responses in i^{th} level (i=1,……,L) and N is the total sample size

A higher Shannon Index indicates that the instrument can obtain more information, and therefore provides more information about health status. Conversely, a lower Shannon Index indicates that the instrument is less capable of measuring a wide range of information and may not capture important aspects of health status. the Shannon Index can be used to compare the performance of different HRQoL instruments and identify the most informative instrument for a particular population or research question.

14.10. Sample size considerations

Sample sizes will be considered for each psychometric property assessed. Relevant sample size suggestions from the 2019 COSMIN study design checklist have been summarised below.(65) Any sample sizes considered 'inadequate' or 'doubtful' as per the COSMIN guidelines will be flagged and discussed.

- Acceptability and feasibility: no standardised sample size recommendations available.
- Floor and ceiling effects: no standardised sample size recommendations available. However, minimum sample sizes are relatively consistent for assessing other psychometric attributes and this will be considered when making assessments on sample sizes for floor and ceiling effects.
- Test-retest reliability: n≥100 very good; n=50-99 adequate; n=30-49 doubtful; n <30 inadequate
- Inter-rater reliability: n≥100 very good; n=50-99 adequate; n=30-49 doubtful; n <30 inadequate
- Known group validity: n≥100 per group very good; n=50-99 per group adequate; n=30-49 per group doubtful; n<30 per group inadequate
- Convergent validity: n≥100 very good; n=50-99 adequate; n=30-49 doubtful; n <30 inadequate
- Responsiveness (construct approach/ hypotheses testing; comparison with other outcome measurements): n≥100 very good; n=50-99 adequate; n=30-49 doubtful; n <30 inadequate
- Responsiveness (criterion approach/ comparison to 'gold standard'; correlations between change scores or ROC analysis): n≥50 in the *smallest group* very good; n=30-50 in the *smallest group* adequate; n<30 in *biggest group* doubtful

14.11. Parametric and non-parametric tests

By default, all analysis will be parametric unless a sample size is considered 'inadequate' or 'doubtful' as described in Section 14.10, whereby the impact of non-parametric tests will be explored. The need for non-parametric tests and the corresponding non-parametric tests used will be described in each manuscript.

15. Funding

This research was funded by the Australian Government Medical Research Futures Fund, grant number 1200816 and EuroQol Research Foundation grant 361-RA.

16. Conflicts of interest

K.D., N.D., B.M., H.H., R.J. have all received previous or current funding from the EuroQol Foundation who is the developer and copyright holder of some instruments included in this study. N.D. and B.M. are members of the EuroQol Group. The EuroQol Foundation are providing some direct funding for this research.

17. References

1. Jones R, Mulhern B, McGregor K, Yip S, Loughlin R, Devlin N, et al. Psychometric Performance of HRQoL Measures: An Australian Paediatric Multi-Instrument Comparison Study Protocol (P-MIC). Children. 2021;8(8):714.
2. Xiong X, Dalziel, K., Huang, L. et al. How do common conditions impact health-related quality of life for children? Providing guidance for validating pediatric preference-based measures. Health Qual Outcomes. 2023.
3. Liu T, Lingam R, Lycett K, Mensah FK, Muller J, Hiscock H, et al. Parent-reported prevalence and persistence of 19 common child health conditions. Archives of Disease in Childhood. 2018;103(6):548.
4. Rowen D, Keetharuth AD, Poku E, Wong R, Pennington B, Wailoo A. A Review of the Psychometric Performance of Selected Child and Adolescent Preference-Based Measures Used to Produce Utilities for Child and Adolescent Health. Value in Health. 2021;24(3):443-60.
5. Edwards B. Growing Up in Australia: the Longitudinal Study of Australian Children: entering adolescence and becoming a young adult. Family Matters. 2014(95):5-14.
6. Reust CE, Williams A. Recurrent Abdominal Pain in Children. Am Fam Physician. 2018;97(12):785-93.
7. Australian Bureau of Statistics (ABS). National Health Survey: first results, 2017–18. ABS cat. no. 4363.0.55.001. Report. Canberra, Australia; 2019a.
8. Hay P, Girosi F, Mond J. Prevalence and sociodemographic correlates of DSM-5 eating disorders in the Australian population. Journal of Eating Disorders. 2015;3(1):19.
9. Caldwell PHY, Hodson E, Craig JC, Edgar D. 4. Bedwetting and toileting problems in children. Medical Journal of Australia. 2005;182(4):190-5.
10. Verstraete J, Ramma L, Jelsma J. Validity and reliability testing of the Toddler and Infant (TANDI) Health Related Quality of Life instrument for very young children. Journal of Patient-Reported Outcomes. 2020;4(1):94.
11. Bailey C, Dalziel K, Cronin P, Devlin N, Viney R, The Quality Of Life in Kids: Key Evidence to Strengthen Decisions in Australia Project T. How are Child-Specific Utility Instruments Used in Decision Making in Australia? A Review of Pharmaceutical Benefits Advisory Committee Public Summary Documents. PharmacoEconomics. 2022;40(2):157-82.
12. 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. Medical Care. 2001;39(8).
13. Varni JW, Seid M, Rode CA. The PedsQL: measurement model for the pediatric quality of life inventory. Med Care. 1999;37(2):126-39.
14. Wille N, Badia X, Bonsel G, Burström K, Cavrini G, Devlin N, et al. Development of the EQ-5D-Y: a child-friendly version of the EQ-5D. Quality of Life Research. 2010;19(6):875-86.

15. Kreimeier S, Åström M, Burström K, Egmar A-C, Gusi N, Herdman M, et al. EQ-5D-Y-5L: developing a revised EQ-5D-Y with increased response categories. Quality of life research : an international journal of quality of life aspects of treatment, care and rehabilitation. 2019;28(7):1951-61.
16. Stevens K. Assessing the performance of a new generic measure of health-related quality of life for children and refining it for use in health state valuation. Applied Health Economics and Health Policy. 2011;9(3):157-69.
17. Stevens K. Developing a descriptive system for a new preference-based measure of health-related quality of life for children. Qual Life Res. 2009;18(8):1105-13.
18. Ratcliffe J, Stevens K, Flynn T, Brazier J, Sawyer MG. Whose values in health? An empirical comparison of the application of adolescent and adult values for the CHU-9D and AQOL-6D in the Australian adolescent general population. Value Health. 2012;15(5):730-6.
19. Moodie M, Richardson J, Rankin B, Iezzi A, Sinha K. Predicting time trade-off health state valuations of adolescents in four Pacific countries using the Assessment of Quality-of-Life (AQoL-6D) instrument. Value Health. 2010;13(8):1014-27.
20. Keating CL, Moodie ML, Richardson J, Swinburn BA. Utility-Based Quality of Life of Overweight and Obese Adolescents. Value in Health. 2011;14(5):752-8.
21. Northwestern University. Intro to PROMIS® [Online]. 2021. Available From (accessed on 14th June 2022): https://www.healthmeasures.net/explore-measurement-systems/promis/intro-to-promis.
22. Feeny D, Furlong W, Torrance GW, Goldsmith CH, Zhu Z, DePauw S, et al. Multiattribute and single-attribute utility functions for the health utilities index mark 3 system. Med Care. 2002;40(2):113-28.
23. Torrance GW, Feeny DH, Furlong WJ, Barr RD, Zhang Y, Wang Q. Multiattribute Utility Function for a Comprehensive Health Status Classification System: Health Utilities Index Mark 2. Medical Care. 1996;34(7).
24. Horsman J, Furlong W, Feeny D, Torrance G. The Health Utilities Index (HUI®): concepts, measurement properties and applications. Health and Quality of Life Outcomes. 2003;1(1):54.
25. Gemke RJ, Bonsel GJ. Reliability and validity of a comprehensive health status measure in a heterogeneous population of children admitted to intensive care. J Clin Epidemiol. 1996;49(3):327-33.
26. Herdman M, Gudex C, Lloyd A, Janssen MF, Kind P, Parkin D, et al. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). Quality of Life Research. 2011;20(10):1727-36.
27. Swanson JM, Schuck S, Porter MM, Carlson C, Hartman CA, Sergeant JA, et al. Categorical and Dimensional Definitions and Evaluations of Symptoms of ADHD: History of the SNAP and the SWAN Rating Scales. Int J Educ Psychol Assess. 2012;10(1):51-70.
28. Serrien DJ, Sovijärvi-Spapé MM, Rana G. Developmental changes in motor control: insights from bimanual coordination. Dev Psychol. 2014;50(1):316-23.
29. Ebesutani C, Korathu-Larson P, Nakamura BJ, Higa-McMillan C, Chorpita B. The Revised Child Anxiety and Depression Scale 25-Parent Version: Scale Development and Validation in a School-Based and Clinical Sample. Assessment. 2017;24(6):712-28.
30. Ravens-Sieberer U, Auquier P, Erhart M, Gosch A, Rajmil L, Bruil J, et al. The KIDSCREEN-27 quality of life measure for children and adolescents: psychometric results from a cross-cultural survey in 13 European countries. Quality of Life Research. 2007;16(8):1347-56.
31. Ravens-Sieberer U, Herdman M, Devine J, Otto C, Bullinger M, Rose M, et al. The European KIDSCREEN approach to measure quality of life and well-being in children: development, current application, and future advances. Quality of Life Research. 2014;23(3):791-803.
32. Varni JW, Burwinkle TM, Rapoff MA, Kamps JL, Olson N. The PedsQL™ in Pediatric Asthma: Reliability and Validity of the Pediatric Quality of Life Inventory™ Generic Core Scales and Asthma Module. Journal of Behavioral Medicine. 2004;27(3):297-318.
33. Adair CE, Marcoux GC, Cram BS, Ewashen CJ, Chafe J, Cassin SE, et al. Development and multi-site validation of a new condition-specific quality of life measure for eating disorders. Health and quality of life outcomes. 2007;5:23-.
34. Goodwin SW, Ferro MA, Speechley KN. Development and assessment of the Quality of Life in Childhood Epilepsy Questionnaire (QOLCE-16). Epilepsia. 2018;59(3):668-78.

35. Varni JW, Thompson KL, Hanson V. The Varni/Thompson Pediatrie Pain Questionnaire. I. Chronic musculoskeletal pain in juvenile rheumatoid arthritis. PAIN. 1987;28(1).
36. Bruni O, Ottaviano S, Guidetti V, Romoli M, Innocenzi M, Cortesi F, et al. The Sleep Disturbance Scale for Children (SDSC). Construction and validation of an instrument to evaluate sleep disturbances in childhood and adolescence. J Sleep Res. 1996;5(4):251-61.
37. Romeo DM, Bruni O, Brogna C, Ferri R, Galluccio C, De Clemente V, et al. Application of the Sleep Disturbance Scale for Children (SDSC) in preschool age. European Journal of Paediatric Neurology. 2013;17(4):374-82.
38. Thomson WM, Foster Page LA, Robinson PG, Do LG, Traebert J, Mohamed AR, et al. Psychometric assessment of the short-form Child Perceptions Questionnaire: an international collaborative study. Community Dent Oral Epidemiol. 2016;44(6):549-56.
39. Jokovic A, Locker D, Guyatt G. Short forms of the Child Perceptions Questionnaire for 11-14-year-old children (CPQ11-14): development and initial evaluation. Health and quality of life outcomes. 2006;4:4-.
40. Foster Page LA, Boyd D, Thomson WM. Do we need more than one Child Perceptions Questionnaire for children and adolescents? BMC Oral Health. 2013;13:26.
41. Varni JW, Delamater AM, Hood KK, Raymond JK, Chang NT, Driscoll KA, et al. PedsQL 3.2 Diabetes Module for Children, Adolescents, and Young Adults: Reliability and Validity in Type 1 Diabetes. Diabetes Care. 2018;41(10):2064-71.
42. Bower WF, Sit FKY, Bluyssen N, Wong EMC, Yeung CK. PinQ: A valid, reliable and reproducible quality-of-life measure in children with bladder dysfunction. Journal of Pediatric Urology. 2006;2(3):185-9.
43. Bower WF, Wong EMC, Yeung CK. Development of a validated quality of life tool specific to children with Bladder dysfunction. Neurourology and Urodynamics. 2006;25(3):221-7.
44. Peasgood T, Mukuria C, Carlton J, Connell J, Devlin N, Jones K, et al. What is the best approach to adopt for identifying the domains for a new measure of health, social care and carer-related quality of life to measure quality-adjusted life years? Application to the development of the EQ-HWB? The European journal of health economics : HEPAC : health economics in prevention and care. 2021;22(7):1067-81.
45. Brazier J, Peasgood T, Mukuria C, Marten O, Kreimeier S, Luo N, et al. The EQ-HWB: Overview of the Development of a Measure of Health and Wellbeing and Key Results. Value in Health. 2022;25(4):482-91.
46. Goodman R, Ford T, Simmons H, Gatward R, Meltzer H. Using the Strengths and Difficulties Questionnaire (SDQ) to screen for child psychiatric disorders in a community sample. British Journal of Psychiatry. 2000;177(6):534-9.
47. Stone LL, Otten R, Engels RCME, Vermulst AA, Janssens JMAM. Psychometric Properties of the Parent and Teacher Versions of the Strengths and Difficulties Questionnaire for 4- to 12-Year-Olds: A Review. Clinical Child and Family Psychology Review. 2010;13(3):254-74.
48. Stansfeld SA, Roberts R, Foot SP. Assessing the validity of the SF-36 General Health Survey. Quality of Life Research. 1997;6(3):0-.
49. Devlin N, Parkin D, Janssen B. Analysis of EQ-5D Profiles. In: Devlin N, Parkin D, Janssen B, editors. Methods for Analysing and Reporting EQ-5D Data. Cham: Springer International Publishing; 2020. p. 23-49.
50. Feng YS, Jiang R, Pickard AS, Kohlmann T. Combining EQ-5D-5L items into a level summary score: demonstrating feasibility using non-parametric item response theory using an international dataset. Qual Life Res. 2022;31(1):11-23.
51. Bethell CD, Read D, Stein RE, Blumberg SJ, Wells N, Newacheck PW. Identifying children with special health care needs: development and evaluation of a short screening instrument. Ambul Pediatr. 2002;2(1):38-48.
52. Terwee CB, Bot SDM, de Boer MR, van der Windt DAWM, Knol DL, Dekker J, et al. Quality criteria were proposed for measurement properties of health status questionnaires. Journal of Clinical Epidemiology. 2007;60(1):34-42.
53. McHorney CA, Tarlov AR. Individual-patient monitoring in clinical practice: are available health status surveys adequate? Qual Life Res. 1995;4(4):293-307.

54. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics. 1977;33(1):159-74.
55. Al-Janabi H, Flynn TN, Peters TJ, Bryan S, Coast J. Test-retest reliability of capability measurement in the UK general population. Health Econ. 2015;24(5):625-30.
56. Koo TK, Li MY. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. J Chiropr Med. 2016;15(2):155-63.
57. Portney LG, Watkins MP. Foundations of clinical research: applications to practice: Pearson/Prentice Hall Upper Saddle River, NJ; 2009.
58. Cicchetti DV. Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. US: American Psychological Association; 1994. p. 284-90.
59. Giavarina D. Understanding Bland Altman analysis. Biochem Med (Zagreb). 2015;25(2):141-51.
60. Chen H-Y, Cisler RA. Assessing Health-Related Quality of Life Among Children With Special Health Care Needs in the United States. Children's Health Care. 2011;40(4):311-25.
61. Richardson J, Iezzi A, Khan MA, Maxwell A. Validity and reliability of the Assessment of Quality of Life (AQoL)-8D multi-attribute utility instrument. Patient. 2014;7(1):85-96.
62. Peasgood T, Mukuria C, Brazier J, Marten O, Kreimeier S, Luo N, et al. Developing a New Generic Health and Wellbeing Measure: Psychometric Survey Results for the EQ-HWB. Value in Health. 2022;25(4):525-33.
63. Varni JW, Burwinkle TM, Seid M, Skarr D. The PedsQL 4.0 as a pediatric population health measure: feasibility, reliability, and validity. Ambul Pediatr. 2003;3(6):329-41.
64. Cohen J. A power primer. Psychol Bull. 1992;112(1):155-9.
65. Mokkink LB, Prinsen C, Patrick DL, Alonso J, Bouter LM, De Vet H, et al. COSMIN Study Design checklist for Patient-reported outcome measurement instruments [Online]. Amsterdam, The Netherlands. 2019:1-32. Available from (accessed 26th April 2023): https://www.cosmin.nl/wp-content/uploads/COSMIN-study-designing-checklist_final.pdf
66. World Health Organization (WHO). Oral Health Surveys, Basic Methods. 5th ed. Geneva, Switzerland: WHO Press, World Health Organization; 2013.
67. Australian Bureau of Statistics (ABS). 2007.0 - Census of Population and Housing: Consultation on Topics, 2021 [Online]: Sex and Gender Canberra, Australia: 2018. Available from (accessed on 11th November 2021): https://www.abs.gov.au/ausstats/abs@.nsf/Lookup/2007.0main+features62021.
68. Inglehart R, Haerpfer C, Moreno A, Welzel C, Kizilova K, Diez-Medrano J, et al. World Values Survey: Round Six - Country-Pooled Datafile Version Madrid [Online]. JD Systems Institute; 2014. Available from (accessed 15th August 2022): https://www.worldvaluessurvey.org/WVSDocumentationWV6.jsp.
69. White M, Pelly R, Le J, Dove L, Connolly S, Morgan A, et al. Feasibility of single question mental health surveillance in chronic disease. Archives of Disease in Childhood. 2022:archdischild-2022-324000.

18. Appendix

Appendix Table 1: Previous data cuts- Summary of P-MIC samples, number recruited to each sample and recruitment status for data cut 1 06 May 2022.

Sample	Sub-sample	N	Recruitment status
Total	n/a	6,247	Ongoing
4) Recruited via hospital	1a) general hospital sample	883	Ongoing
	1b) specialised hospital sample, including the following five groups:	121	Ongoing
	ICU		
	SSU	20	Ongoing
	Born premature	16	Ongoing
	Rare genetic condition	20	Ongoing

			65	Ongoing
5)	General population sample recruited via online panels	2a) general population sample with a four-week follow-up	1,624	Ongoing
		2b) general population sample with two-day follow-up	251	Complete
6)	Health condition-specific groups primarily recruited via online panels	3a) ADHD	517	Ongoing
		3b) Anxiety or depression	470	Ongoing
		3c) ASD	521	Ongoing
		3d) Asthma	370	Ongoing
		3e) Eating disorder	140	Ongoing
		3f) Epilepsy	196	Ongoing
		3g) Recurrent abdominal pain	370	Ongoing
		3h) Sleep problems	376	Ongoing
		3i) Tooth problems	408	Ongoing
		3j) Type 1 diabetes	n/a	n/a sample added later
		3k) Wetting problems	n/a	n/a sample added later

Appendix Table 2: Summary of HRQoL instruments included and key instrument characteristics.

Instrument	Descriptions	Number of items	Item response levels	Recall	Domains/dimensions/scales
Core HRQoL instruments					
PedsQL generic core 4.0 (12, 13)	Generic paediatric HRQoL instrument.	23 items	5 levels	1 month	Physical functioning, emotional functioning, social functioning, and school functioning.
EQ-5D-Y-3L(14)	Generic paediatric HRQoL instrument.	5 items	3 levels	Today	Mobility, looking after self, usual activities, pain/discomfort, and worried/sad.
EQ-5D-Y-5L (14, 15)	Generic paediatric HRQoL instrument.	5 items	5 levels	Today	Mobility, looking after self, usual activities, pain/discomfort, and worried/sad.
CHU9D (16, 17)	Generic paediatric HRQoL instrument.	9 items	5 levels	Today	Worried, sad, pain, tired, annoyed, schoolwork/homework, daily routine, activities, and sleep.
TANDI (10)	Generic toddler and infant HRQoL instrument.	6 items	3 levels	Today	Movement, play, pain, social interaction, communication, and eating.
Additional HRQoL instruments					
AQoL-6D (18)	Generic adolescent HRQoL instrument.	20 items	4 to 6 levels	1 week	Independent living, mental health, coping, relationships, pain, and senses
PROMIS-25 paediatric profile (21)	Generic paediatric HRQoL instrument.	25 items	5 levels, except for the pain item which is 10 levels.	1 week	Physical function mobility, anxiety, depressive symptoms, fatigue, peer relationships, and pain interference.
HUI 2/3 (22-24)	Generic HRQoL instrument that can be used in paediatric populations.	15 items	4 to 6 levels	The HUI 2/3 has 'current' recall versions with a specified recall time period or a 'usual' recall version. The usual recall version was	The HUI3 classification system: vision, hearing, speech, ambulation, dexterity, emotion, cognition, and pain. The HUI2 classification: sensation, mobility, emotion, self-care, cognition, pain, and fertility. However, the fertility domain is

Instrument	Descriptions	Number of items	Item response levels	Recall	Domains/dimensions/scales
				used for this study.	dropped when being used in paediatric populations.
EQ-5D-5L (26)	Generic adult HRQoL instrument.	5 items	5 levels	Today	Mobility, self-care, usual activities, pain/discomfort, and anxiety/depression.
Health condition-specific instruments					
SWAN (ADHD) (27)	ADHD symptom scale	18 items	7 levels	1 month	Inattention, hyperactivity, and impulsivity.
RCADS-25 (Anxiety/ depression) (29)	Anxiety and depression symptom scale	25 items	4 levels	n/a	Generalized anxiety disorder, major depressive disorder, obsessive compulsive disorder, panic disorder, separation anxiety disorder, and social phobia.
KIDSCREEN-27 (ASD) (30, 31)	Generic paediatric HRQoL instrument	27 items	5 levels	1 week	Physical wellbeing, psychological wellbeing, autonomy/ parent relation, peer/social support, and school environment.
PedsQL asthma module (Asthma) (32)	Asthma paediatric HRQoL instrument	28 items (26 items in 2-4 year old version)	5 levels	1 month	Asthma, treatment, worry, and communication.
EDQLS (Eating disorder) (33)	Eating disorder adolescent and adult quality of life instrument	40 items	5 levels	1 week	Cognitive, education/vocation, family and close relationships, relationships with others, future outlook, appearance, leisure, psychological, emotional, values and beliefs, physical, and eating.
QOLCE-16 (Epilepsy) (34)	Epilepsy specific paediatric HRQoL instrument	16 items	6 levels	4 weeks	Cognitive functioning, emotional functioning, social functioning, and physical functioning.
Pain VAS (Recurrent abdominal pain) (35)	Pain VAS adapted from the Paediatric Pain Questionnaire	2 items	VAS scale	Today and last pain episode.	n/a
SDSC (Sleep problems) (36, 37)	Paediatric sleep disturbances and sleep behaviour instrument	26 items	5 levels	6 months	Parasomnias, difficulty in initiating and maintaining sleep, sleep disordered breathing, disorders of excessive somnolence, sleep hyperhydrosis and non-restorative sleep.
CPQ-11-14 short form (Tooth problems) (38, 39)	Paediatric oral HRQoL instrument	16 items	5 levels	3 months	Oral symptoms, functional limitations, emotional well-being, and social well-being
PedsQL diabetes module (Type 1 diabetes) (41)	Paediatric diabetes HRQoL instrument	33-items	5 levels	1 month	Symptoms, treatment barriers, treatment adherence, worry, and communication
PinQ (Wetting) (42, 43)	Paediatric bladder dysfunction	20-items	5 levels	none	Intrinsic and extrinsic

Instrument	Descriptions	Number of items	Item response levels	Recall	Domains/dimensions/scales
	HRQoL instrument				
Other instruments					
EQ-HWB-S (44, 45)	Instrument assessing the impact of health and wellbeing being as a care recipient or caregiver.	9 items	5 levels	7 day	Mobility, usual activities, energy, cognition, social relationships, control, anxiety/depression, and pain.
SDQ (46, 47)	Paediatric behavioural screening questionnaire.	25 items	3 levels	1 month	Emotional symptoms, conduct problems, hyperactivity/inattention problems, peer relationship problems, and prosocial behaviour

Appendix Table 3: Health condition-specific questions to approximate health condition severity.

Health condition-specific sample	Severity questions	Source
3a. ADHD	1. Does your child currently take regular medication for their ADHD? Yes No 2. Thinking about your child's ADHD and its impact on school, would you say their ADHD has: No Little impact Some impact A large impact 3. Thinking about your child's ADHD and its impact on home, would you say their ADHD has: No Little impact Some impact A large impact 4. Thinking about your child's ADHD and its impact on social life, would you say their ADHD has: No Little impact Some impact A large impact	Consultation with clinical expert.
3b. Anxiety or depression	N/A, severity measured using SDQ.	N/A
3c. ASD	1. What type of school does the Study Child attend? A special school Does not attend school Mainstream school with integration support funding Mainstream school with no integration support funding	Consultation with clinical and research experts. Derived from severity question used in ASD study at MCRI, iSAID project.
3d. Asthma	1. Since they were first diagnosed, has your child ever required an overnight hospital stay for their asthma? Yes- *go to a and b* No- *go to 2* a) How many times have they required an overnight hospital stay for their asthma?	Consultation with clinical expert.

Health condition-specific sample	**Severity questions**	**Source**
	b) When was their most recent overnight hospital stay for their asthma? 2. Since they were first diagnosed, has your child ever had to attend an Emergency Department for their asthma? Yes- *go to a and b* No- *go to 3* a) How many times have they attended an Emergency department for their asthma? b) When was their most recent attendance to an Emergency department for their asthma? 3 months/6 months/12 months/ more than 12 months/ I'm not sure 3. Does your child currently have a prescription for an oral corticosteroid (also called a 'preventer') medication for their asthma? *This includes medications such as Flixotide®, Pulmicort®, Alvesco® and Symbicort®* Yes No	
3e. Eating disorder	1. Since they were first diagnosed, has your child ever required an overnight hospital stay for their eating disorder? Yes- *go to a and b* No- *go to 2* a) How many times have they required an overnight hospital stay for their eating disorder? b) When was their most recent overnight hospital stay for their eating disorder? 2. Since they were first diagnosed, has your child ever had to attend an Emergency Department for their eating disorder? Yes- *go to a and b* No- *go to 3* a) How many times have they attended an Emergency department for their eating disorder? b) When was their most recent attendance to an Emergency department for their eating disorder? 3. Is your child regularly meeting with a health care provider for their eating disorder (e.g. counsellor or mental health professional, eating disorder service, CAMHS, paediatrician, GP, headspace, dietician)? Yes No	Consultation with clinical expert.
3f. Epilepsy	1. How old was your child when they had their first seizure? 2. When was your child's last seizure? 3. How frequently does your child experience seizures? 4. How many daily medications does your child take for their epilepsy?	Consultation with clinical expert.
3g. Recurrent abdominal pain	1. Overall, would you describe the child's recurrent abdominal pain condition as mild, moderate or severe? 2. Would you describe the child's last abdominal pain episode as mild, moderate or severe?	Adapted from LSAC.(5)
3h. Sleep problems	1. Thinking about your child with sleep problems, how much is their ongoing sleeping pattern or habits a problem for you? Not a problem at all	Adapted from LSAC.(5)

Health condition-specific sample	**Severity questions**	**Source**
	A small problem A moderate problem A large problem	
3i. Tooth problems	1. Which of the following tooth problems has the study child **experienced in the last 3 months?** Cavities, dental decay or hole(s) in teeth Tooth or teeth filled because of dental decay Teeth pulled because of dental decay Accident causing breakage or loss of teeth Crowded teeth Problems with bite (e.g., crossbite or overbite) a) Has your child been hospitalised for this problem? b) Has the problem been treated? c) *If Yes to b,* How long ago was this problem treated? 2. How would you describe the health of your child's teeth and gums? *(Respond on the following scale for both teeth and gums)* Excellent Very good Good Average Poor Very poor Don't know	Consultation with clinical expert and adapted from World Health Organisation (WHO) oral health questionnaire.(66)
3j. Type 1 diabetes	1. Since they were first diagnosed, has your child ever required an overnight hospital stay for their diabetes? Yes- go to a and b No- go to next a. How many times have they required an overnight hospital stay for their diabetes? b. When was their most recent overnight hospital stay for their diabetes? 2. Since they were first diagnosed, has your child ever had to attend an Emergency Department for their diabetes? Yes- go to a and b No- go to next a. How many times have they attended an Emergency department for their diabetes? b. When was their most recent attendance to an Emergency department for their diabetes? 3 months/6 months/12 months/ more than 12 months/ I'm not sure 3. In the last month, how many times has your child's blood glucose levels been out of the desired range?	Consultation with clinical expert.
3k. Wetting	1. Would you describe the child's Wetting self during the day as mild, moderate or severe? 2. How often does the child experience night wetting? 4 nights a week, 5 nights a week, 6 nights a week or 7 nights a week	Adapted from LSAC.(5)

Appendix Table 4: Demographic Questions.

Question and outcome options	Source
Study Child Demographic Questions	
1. Is the Study Child currently receiving care at The Royal Children's Hospital? Yes No 1a. Which hospital department is your child currently receiving care from? Emergency Department Short Stay Unit Intensive Care Unit (ICU) or expected to be in the care of ICU soon Outpatient care – Adolescent Medicine Outpatient care- Neurodevelopment and Disability Outpatient care- Neurology Outpatient care- Gynaecology Outpatient care- Centre for community child health Outpatient care- General Medicine Outpatient care- Endocrinology and Diabetes Outpatient care- Complex Care Hub Asthma Outpatient care- Metabolic Medicine Outpatient care- Ophthalmology Outpatient care - Rheumatology Outpatient care- Other Surgery- Colorectal Surgery- Facial Surgery- Day surgery I'm not sure Other *(drop to free text)* 1b. <u>If yes to any Outpatient departments</u>, Is this your first appointment with this department? Yes, first appointment No, not first appointment 1c. <u>If yes to any surgery departments</u>, What type of appointment did you most recently have or are about to have with this department? Initial or preoperative appointment (before surgery) Review appointment (after surgery) Other 1c. <u>If Yes to ICU,</u> Has the Study Child recently had or currently having any of the following surgeries? Cardiac (heart) surgery Spinal surgery Ear, nose, or throat surgery Other (please specify) No	Based on recruiting hospital departments.
2. Is the Study Child currently receiving care at The Royal Women's Hospital? Yes No 2a. <u>If yes to 2</u>, Was the Study Child born premature? Yes No 2b. <u>If yes to 2a,</u> How many weeks gestation was he Study Child born? 2c. <u>If yes to 2a,</u> Is the Study Child also a participant in the PLUSS study? Yes No I don't know	Based on recruiting hospital departments.
3. Who lives with Study Child in the primary home? (Please tick all that apply, including yourself) *If the Study Child has multiple homes please answer based on the home the child spends the most time in.*	Adapted from LSAC (5)

Question and outcome options	Source
Parent Parents Sibling(s) Grandparent(s) Other relative(s) Someone not related to them	
4. What was the Study Child's age at last birthday?	Adapted from LSAC (5)
5. What is the gender of the Study Child? Male Female Transgender Female Transgender Male Not described (please specify) Prefer not to say	(67)
6. Is Study Child of Aboriginal or Torres Strait Islander origin? No Yes, Aboriginal Yes, Torres Strait Islander Yes, both	Adapted from LSAC (5)
7. Does Study Child speak a language other than English at home? (If more than one, record main language) Australian Standard Classification of Languages (ASCL) code	Adapted from LSAC (5)
8. What is the postcode of the Study Child's primary home?	
9. Does the Study Child have a disability? No Yes	
10. Does the Study Child currently need or use medicine prescribed by a doctor (other than vitamins)? Yes - Go to Question 10a No - Go to Question 11 10a. Is this because of ANY medical, behavioural or other health condition? Yes - Go to Question 10b No - Go to Question 11 10b. Is this a condition that has lasted or is expected to last for at least 12 months? Yes No	(51)
11. Does the Study Child need or use more medical care, mental health or educational services than is usual for most children of the same age? Yes - Go to Question 11a No - Go to Question 12 11a. Is this because of ANY medical, behavioural or other health condition? Yes - Go to Question 11b No - Go to Question 12 11b. Is this a condition that has lasted or is expected to last for at least 12 months? Yes No	(51)
12. Does the Study Child have any medical conditions or disabilities that have lasted or are likely to last for six months or more? Yes No	Adapted from LSAC (5)
13. Does Study Child have any of these common ongoing conditions? Tick ALL that apply, including conditions you have already told us about. We understand there are many more ongoing conditions not listed here, if your child's condition is not listed we ask you select *Other illness* and specify the condition in the pop out box.	(3, 5) Also based on samples or sub-samples for the study.

Question and outcome options	**Source**
'Ongoing conditions' exist for some period of time (weeks, months or years) or re-occur regularly. They do not have to be diagnosed by a doctor. Anaemia Anxiety disorder Asthma Attention Deficit Disorder (ADD or ADHD) Autism spectrum disorder or Aspergers Arthritis Bedwetting after age 16 Behavioural, cognitive & emotional problems Bone, joint or muscle problem Born premature Chronic Fatigue Chronic sinusitis Constipation Dental decay Depression Developmental delay Diabetes Diarrhoea or colitis Ear infections Eating disorder Eczema Epilepsy or seizure disorder Eyes or seeing problems Food or digestive allergies Frequent headaches Genetic condition (if ticked-Is this a rare genetic condition? Yes/No) Hay fever Hearing problems Irritable bowel Overweight/obesity Palpitations Physical disabilities Problems with psychological development Recurrent abdominal pain Recurrent back pain Recurrent chest pain Recurrent pain in other parts of the body Sleep problems Soiling after age 4 Tonsillitis Tooth problems in the last 3 months Undiagnosed condition Wetting self during day after age 6 Wetting self during night Wheezing that lasts more than a week Other Illness (please specify) Additional items added to list for Epilepsy sample only: Learning disability Intellectual disability Language delay Hemiplegia/quadriplegia/dystonia	
14. Is the Study Child in High School? *(only for children aged 16years+)* Yes No	
15. All things considered, how happy would you say the Study Child is usually? Very happy Happy Neither happy nor unhappy Not very happy	Adapted from the World Value Survey.(68)

Question and outcome options	Source
Very unhappy	
16. Thinking about your (child's/ teenager's) mental health over the past 4 weeks, are they thriving/coping/struggling/always overwhelmed? Thriving Coping Struggling Always overwhelmed	(69)
Caregiver Demographic Questions	
17. How are you related to the Study Child? Parent Grandparent Sibling Carer unrelated to child Other relative	Adapted from LSAC (5)
18. What was your age at last birthday?	Adapted from LSAC (5)
19. What is your gender? Male Female Transgender Female Transgender Male Not described (please specify) Prefer not to say	(67)
20. Before income tax is taken out how much does the household usually receive from all sources of income weekly? Less than $500 per week ($25,999 or less per year) $500-$999 per week ($26,000-$51,999 per year) $1,000-$1,999 per week ($52,000-$103,9799 per year) $2,000 or more per week ($104,000 or more per year)	Adapted from LSAC (5)
21. How many adults usually reside in your household (including you)?	Adapted from LSAC (5)
22. How many children usually reside in your household? (including Study Child)	Adapted from LSAC (5)
23. Do you currently have a government Health Care Card for yourself? Yes No	Adapted from LSAC (5)
24. What is your highest level of education? Bachelor Degree or above Certificate III/IV or Diploma (including Advanced Diploma) Year 12 Year 9-11 Certificate I/II Year 8 or below Never attended school and no non-school qualification Still at high school	Adapted from LSAC (5)

Appendix 5 – Hypothesised item correlations

Appendix 5 1. Items hypothesised to be correlated (indicated by green cell) and not to be correlated (indicated by red cells) between PedsQL and EQ-5D-Y-3L/5L

Domain	**PedsQL**	**EQ-5D-Y-3L & 5L**				
		Mobility	Looking after self	Usual activities	Pain/ discomfort	Sad/worried
Physical functioning	Walking					
	Running					
	Participating in sports activities or exercise					

Domain	**PedsQL**	**EQ-5D-Y-3L & 5L**				
		Mobility	Looking after self	Usual activities	Pain/ discomfort	Sad/worried
	Lifting something					
	Bathing					
	Doing chores around the house					
	Having hurts or aches					
	Low energy levels					
Emotional functioning	Feeling afraid or scared					
	Feeling sad or blue					
	Feeling angry					
	Trouble sleeping					
	Worrying					
Social functioning	Playing/ getting along with other children					
	Other children not wanting to play/ be friends with them					
	Getting teased					
	Not able to do things with other children their age can do					
	Keeping up when playing with other children					
School functioning	Paying attention					
	Forgetting things					
	Keeping up with schoolwork					
	Missing school because not well					
	Missing school to go to doctor or hospital					

Appendix 5.2. Items hypothesised to be correlated (indicated by green cell) and not to be correlated (indicated by red cells) between PedsQL and CHU9D.

Domain	**PedsQL**	**CHU9D**								
		Worried	Sad	Pain	Tired	Annoyed	School	Sleep	Daily routine	Activities
Physical functioning	Walking									
	Running									
	Participating in sports activities or exercise									
	Lifting something									
	Bathing									
	Doing chores around the house									

Domain	**PedsQL**	**CHU9D**								
		Worried	Sad	Pain	Tired	Annoyed	School	Sleep	Daily routine	Activities
	Having hurts or aches									
	Low energy levels									
Emotional functioning	Feeling afraid or scared									
	Feeling sad or blue									
	Feeling angry									
	Trouble sleeping									
	Worrying									
Social functioning	Playing/ getting along with other children									
	Other children not wanting to play/ be friends with them									
	Getting teased									
	Not able to do things with other children their age can do									
	Keeping up when playing with other children									
School functioning	Paying attention									
	Forgetting things									
	Keeping up with schoolwork									
	Missing school because not well									
	Missing school to go to doctor or hospital									

Appendix 5.3. Items hypothesised to be correlated (indicated by green cell) and not to be correlated (indicated by red cells) between PedsQL and AQoL-6D.

PedsQL		**AQoL-6D**																				
Domain		Physical ability				Social and family relationships			Mental health				Coping			Pain			Vision, hearing and communication		
	Item	Help needed with jobs	Getting around	Walk/run	Washing self	Happiness from close friendships	Health affect on relationships	Health affect on participation	How often feel despair	How often feel worried	How often feel sad	How often calm/ stressed	How much energy	How often do you manage life well	Cope with life problems	How often experience serious physical pain	How much physical pain experience	How often physical pain interfere	Vision	Hearing	Communicate
Physical functioning	Walking																				
	Running																				
	Participating in sports activities or exercise																				
	Lifting something																				
	Bathing																				
	Doing chores around the house																				
	Having hurts or aches																				

PedsQL		**AQoL-6D**																			
Domain		Physical ability				Social and family relationships			Mental health				Coping			Pain			Vision, hearing and communication		
	Item	Help needed with jobs	Getting around	Walk/run	Washing self	Happiness from close friendships	Health affect on relationships	Health affect on participation	How often feel despair	How often feel worried	How often feel sad	How often calm/ stressed	How much energy	How often do you manage life well	Cope with life problems	How often experience serious physical pain	How much physical pain experience	How often physical pain interfere	Vision	Hearing	Communicate
	Low energy levels																				
Emotional functioning	Feeling afraid or scared																				
	Feeling sad or blue																				
	Feeling angry																				
	Trouble sleeping																				
	Worrying																				
Social functioning	Playing/ getting along with other children																				
	Other children																				

PedsQL		**AQoL-6D**																		
Domain		Physical ability				Social and family relationships			Mental health				Coping			Pain			Vision, hearing and communication		
	Item	Help needed with jobs	Getting around	Walk/run	Washing self	Happiness from close friendships	Health affect on relationships	Health affect on participation	How often feel despair	How often feel worried	How often feel sad	How often calm/ stressed	How much energy	How often do you manage life well	Cope with life problems	How often experience serious physical pain	How much physical pain experience	How often physical pain interfere	Vision	Hearing	Communicate
	not wanting to play/ be friends with them																				
	Getting teased																				
	Not able to do things other children their age can do																				
	Keeping up when playing with other children																				

PedsQL		**AQoL-6D**																			
Domain		Physical ability				Social and family relationships			Mental health				Coping			Pain			Vision, hearing and communication		
	Item	Help needed with jobs	Getting around	Walk/run	Washing self	Happiness from close friendships	Health affect on relationships	Health affect on participation	How often feel despair	How often feel worried	How often feel sad	How often calm/ stressed	How much energy	How often do you manage life well	Cope with life problems	How often experience serious physical pain	How much physical pain experience	How often physical pain interfere	Vision	Hearing	Communicate
School functioning	Paying attention																				
	Forgetting things																				
	Keeping up with schoolwork																				
	Missing school because not well																				
	Missing school to go to doctor or hospital																				

Appendix 5.4. Items hypothesised to be correlated (indicated by green cell) and not to be correlated (indicated by red cells) between HUI 3 and PedsQL

Domain	**PedsQL**	**HUI 3**							
		Vision	Hearing	Speech	Ambulation	Dexterity	Emotion	Cognition	Pain
Physical functioning	Walking								
	Running								
	Participating in sports activities or exercise								
	Lifting something								
	Bathing								
	Doing chores around the house								
	Having hurts or aches								
	Low energy levels								
Emotional functioning	Feeling afraid or scared								
	Feeling sad or blue								
	Feeling angry								
	Trouble sleeping								
	Worrying								
Social functioning	Playing/ getting along with other children								
	Other children not wanting to play/ be friends with them								
	Getting teased								
	Not able to do things other children their age can do								
	Keeping up when playing with other children								
School functioning	Paying attention								
	Forgetting things								
	Keeping up with schoolwork								
	Missing school because not well								
	Missing school to go to doctor or hospital								

Appendix 5.5. Items hypothesised to be correlated (indicated by green cell) and not to be correlated (indicated by red cells) between EQ-5D-Y-3L/5L and EQ-5D-Y-5L.

EQ-5D-Y-3L/5L	**EQ-5D-Y-5L**				
Item	Mobility	Looking after self	Usual activities	Pain/ discomfort	Sad/worried
Mobility					
Looking after self					
Usual activities					
Pain/ discomfort					
Sad/worried					

Appendix 5.6. Items hypothesised to be correlated (indicated by green cell) and not to be correlated (indicated by red cells) between EQ-5D-Y-3L/5L and CHU9D.

	CHU9D								
EQ-5D-Y-3L/5L	Worried	Sad	Pain	Tired	Annoyed	School	Sleep	Daily routine	Activities
Mobility									
Looking after self									
Usual activities									
Pain/ discomfort									
Sad/worried									

Appendix 5.7. Hypothesised item correlations (indicated by green cell) between EQ-5D-Y-3L/5L and AQoL-6D.

AQoL-6D		**EQ-5D-Y-3L**				
Domain	**Item**	Mobility	Looking after self	Usual activities	Pain/ discomfort	Sad/worried
Physical ability	Help needed with jobs					
	Getting around					
	Walk/run					
	Washing self					
Social and family relationships	Happiness from close friendships					
	Health affect on relationships					
	Health affect on participation					
Mental Health	How often feel despair					
	How often feel worried					
	How often feel sad					
	How often calm/ stressed					
Coping	How much energy					
	How often do you manage life well					
	Cope with life problems					
Pain	How often experience serious physical pain					

AQoL-6D		EQ-5D-Y-3L				
Domain	**Item**	Mobility	Looking after self	Usual activities	Pain/ discomfort	Sad/worried
	How much physical pain experience					
	How often physical pain interfere					
Vision, hearing and	Vision					
	Hearing					
	Communicate					

Appendix 5.8. Items hypothesised to be correlated (indicated by green cell) and not to be correlated (indicated by red cells) between EQ-5D-Y-3L/5L and HUI 3.

	HUI 3							
EQ-5D-Y-3L/5L	Vision	Hearing	Speech	Ambulation	Dexterity	Emotion	Cognition	Pain
Mobility								
Looking after self								
Usual activities								
Pain/ discomfort								
Sad/worried								

Appendix 5.9. Items hypothesised to be correlated (indicated by green cell) and not to be correlated (indicated by red cells) between CHU9D and AQoL-6D.

AQoL-6D		CHU9D								
Domain	**Item**	Worried	Sad	Pain	Tired	Annoyed	School	Sleep	Daily routine	Activities
Physical ability	Help needed with jobs									
	Getting around									
	Walk/run									
	Washing self									
Social and family relationships	Happiness from close friendships									
	Health affect on relationships									
	Health affect on participation									
Mental Health	How often feel despair									
	How often feel worried									
	How often feel sad									
	How often calm/ stressed									
Coping	How much energy									
	How often do you manage life well									
	Cope with life problems									

AQoL-6D		CHU9D								
Domain	**Item**	Worried	Sad	Pain	Tired	Annoyed	School	Sleep	Daily routine	Activities
Pain	How often experience serious physical pain									
	How much physical pain experience									
	How often physical pain interfere									
Vision, hearing and	Vision									
	Hearing									
	Communicate									

Appendix 5.10. Items hypothesised to be correlated (indicated by green cell) and not to be correlated (indicated by red cells) between CHU9D and HUI 3.

	HUI 3							
CHU9D	Vision	Hearing	Speech	Ambulation	Dexterity	Emotion	Cognition	Pain
Worried								
Sad								
Pain								
Tired								
Annoyed								
School								
Sleep								
Daily Routine								
Activities								

Appendix 6 - Hypothesised item pool construct correlations

Appendix 6.1. Hypothesised item pool construct correlations for PedsQL

PedsQL domains	**PedsQL items**	Item included in **physical activity** construct pool	Item included in **emotion** construct pool	Item included in **pain** construct pool	Item included in **daily routine** construct pool	Item included in **school/cognition** construct pool	Item included in **social/relationships** construct pool
Physical functioning	Walking	✓					
	Running	✓					
	Participating in sports activities or exercise	✓					
	Lifting something	✓					

PedsQL domains	**PedsQL items**	Item included in **physical activity** construct pool	Item included in **emotion** construct pool	Item included in **pain** construct pool	Item included in **daily routine** construct pool	Item included in **school/cognition** construct pool	Item included in **social/relationships** construct pool
	Bathing				✓		
	Doing chores around the house				✓		
	Having hurts or aches			✓			
	Low energy levels						
Emotional functioning	Feeling afraid or scared		✓				
	Feeling sad or blue		✓				
	Feeling angry		✓				
	Trouble sleeping						
	Worrying		✓				
Social functioning	Playing/ getting along with other children						✓
	Other children not wanting to play/ be friends with them						✓
	Getting teased						✓
	Not able to do things other children their age can do				✓		
	Keeping up when playing with other children						✓
School functioning	Paying attention					✓	
	Forgetting things					✓	
	Keeping up with schoolwork					✓	
	Missing school because not well				✓		
	Missing school to go to doctor or hospital				✓		

Appendix 6.2. Hypothesised item pool construct correlations for EQ-5D-Y-3L & 5L

EQ-5D-Y- 3L and 5L items	Item included in **physical mobility** construct pool	Item included in **emotion** construct pool	Item included in **pain** construct pool	Item included in **daily routine** construct pool	Item included in **school/cognition** construct pool	Item included in **social/relationships** construct pool
Mobility	✓					

Looking after self				✓		
Usual activities				✓		
Pain/ discomfort			✓			
Sad/worried		✓				

Appendix 6.3. Hypothesised item pool construct correlations for CHU9D

CHU9D items	Item included in **physical mobility** construct pool	Item included in **emotion** construct pool	Item included in **pain** construct pool	Item included in **daily routine** construct pool	Item included in **school/cognition** construct pool	Item included in **social/relationships** construct pool
Worried		✓				
Sad		✓				
Pain			✓			
Tired						
Annoyed		✓				
School				✓	✓	
Sleep						
Daily routine (things like eating, having a bath/shower, getting dressed)				✓		
Activities (things like playing out with your friends, doing sports, joining in things)	✓					✓

Appendix 6.4. Hypothesised item pool construct correlations for AQoL-6D

AQOL Domain	**AQOL Item**	Item included in **physical mobility** construct pool	Item included in **emotion** construct pool	Item included in **pain** construct pool	Item included in **daily routine** construct pool	Item included in **school/cognition** construct pool	Item included in **social/relationships** construct pool
Physical ability	Help needed with jobs				✓		
	Getting around	✓					
	Walk/run	✓					
	Washing self				✓		
Social and family relationships	Happiness from close friendships		✓				✓
	Health affect on relationships						✓
	Health affect on participation						✓
Mental Health	How often feel despair		✓				
	How often feel worried		✓				
	How often feel sad		✓				
	How often calm/ stressed		✓				
Coping	How much energy						

AQOL Domain	**AQOL Item**	Item included in **physical mobility** construct pool	Item included in **emotion** construct pool	Item included in **pain** construct pool	Item included in **daily routine** construct pool	Item included in **school/cognition** construct pool	Item included in **social/relationships** construct pool
	How often do you manage life well					✓	
	Cope with life problems		✓			✓	
Pain	How often experience serious physical pain			✓			
	How much physical pain experience			✓			
	How often physical pain interfere			✓			
Vision, hearing and	Vision						
	Hearing						
	Communicate						✓

Appendix 6.5. Hypothesised item pool construct correlations for HUI 3

HUI 3 domains	Item included in **physical mobility** construct pool	Item included in **emotion** construct pool	Item included in **pain** construct pool	Item included in **daily routine** construct pool	Item included in **school/cognition** construct pool	Item included in **social/relationships** construct pool
Vision						
Hearing						
Speech						✓
Ambulation	✓					
Dexterity						
Emotion		✓				
Cognition					✓	
Pain			✓			