Universal Library of Medical and Health Sciences Methodological Guide

ISSN: 3064-6545 | Volume 4, Issue 3

Open Access | PP: 01-45

DOI: https://doi.org/10.70315/uloap.ulmhs.2026.0403001

Adapting Patient-Reported Outcome Measures Across Languages and Cultures in Pediatric Rheumatology - An Integrated Methodology for Validation and Implementation

Ankita Rana, DPT, OCS

Board Certified Orthopedic Clinical Specialist

Abstract

This monograph develops an integrated methodological framework for the cross-cultural adaptation and validation of patient-reported outcome (PRO) measures in pediatric rheumatology. The framework, here named the Cross-Cultural PRO Adaptation Framework for Pediatric Rheumatology (CC-PRO-PR), arose from the observation that the methodological literature on cross-cultural PRO adaptation, while well developed at the general level, leaves the specific demands of pediatric rheumatology insufficiently treated. Dual-perspective reporting from the youth and the caregiver, the developmental variability of the pediatric respondent, the disease-specific content of juvenile rheumatic conditions, and the small samples typical of rare-disease validation each require structural rather than residual methodological treatment. The monograph consolidates these considerations into a single workflow and positions that workflow alongside the established general-purpose guidelines that the field has developed over the past three decades.

The monograph is intended for working specialists in pediatric rheumatology PRO measurement: researchers planning new adaptation studies, clinicians using or evaluating adapted instruments in routine care, instrument developers whose engagement supports adaptation by other teams, and institutions and funding bodies that evaluate adaptation proposals. The reader is expected to be familiar with the general vocabulary of PRO measurement and pediatric rheumatology; technical concepts are defined where introduced, and a glossary in the back matter consolidates the principal terms.

The author declares no competing financial interests. The work received no external funding. Ethical approval for the validation study referenced as Case 1 in Chapter 4 was granted by the relevant institutional review board.

Table of Contents

Citation: Ankita Rana, "Adapting Patient-Reported Outcome Measures Across Languages and Cultures in Pediatric Rheumatology - An Integrated Methodology for Validation and Implementation", Universal Library of Medical and Health Sciences, 2026; 4(3): 01-45. DOI: https://doi.org/10.70315/uloap.ulmhs.2026.0403001.

INTRODUCTION

Relevance

Patient-reported outcome (PRO) measures occupy a central place in contemporary pediatric rheumatology. Over the past three decades, the field has moved from clinician-rated joint counts and laboratory markers as the principal indicators of treatment success toward a multidimensional view in which the patient's own account of pain, function, daily activity, treatment burden, and psychosocial impact is treated as primary outcome data. The conceptual foundation for this shift was formalised by Wilson and Cleary (1995), who proposed that health-related quality of life (HRQOL) is the integrating endpoint that links biological and physiological variables, symptom status, functional status, and general health perceptions in a single explanatory model. Within pediatric rheumatology, that foundation has been operationalised through disease-specific instruments such as the Pediatric Quality of Life Inventory Rheumatology Module 3.0 (Varni et al., 2002), the Childhood Health Assessment Questionnaire (Singh et al., 1994), the Juvenile Arthritis Multidimensional Assessment Report developed under the Paediatric Rheumatology International Trials Organisation (Consolaro et al., 2011), and the Pediatric Rheumatology Quality of Life Scale (Filocamo et al., 2010). These instruments now inform clinical decisions, clinical-trial endpoints, and increasingly regulatory submissions for new therapies in juvenile rheumatic disease.

Regulatory and methodological bodies have reinforced this trajectory. The United States Food and Drug Administration's PRO guidance, the International Society for Quality of Life Research minimum standards (Reeve et al., 2013), the International Society for Pharmacoeconomics and Outcomes Research good-practice principles (Wild et al., 2005), and the consensus-based Standards for the Selection of Health Measurement Instruments (COSMIN) initiative (Mokkink et al., 2010, 2018) together define how PRO instruments should be developed, evaluated, and selected for use. The cumulative effect is that an instrument cannot enter routine clinical practice or research evidence without demonstrated content validity, internal consistency, test-retest reliability, construct validity, and an explicit account of its measurement properties.

The same period has revealed the linguistic and cultural reach of these requirements. Pediatric rheumatic conditions, including juvenile idiopathic arthritis classified by the International League of Associations for Rheumatology (Petty et al., 2004), juvenile systemic lupus erythematosus, juvenile dermatomyositis, and juvenile fibromyalgia, occur across every world region. The instruments that measure their impact on HRQOL, however, were predominantly developed in English. Translation alone does not produce a usable instrument in a new population: conceptual equivalence, cultural appropriateness, linguistic clarity at the developmental stage of the respondent, and psychometric performance in the target language must each be established (Beaton et al., 2000; Wild et al., 2005).

The scale of the gap is substantial. Hindi is among the world's most widely spoken languages, yet validated Hindi versions of several leading PRO instruments in pediatric rheumatology have until recently been absent or limited. Comparable gaps exist for Bengali, Tamil, Swahili, Vietnamese, and other major languages spoken by hundreds of millions of children and their caregivers. The practical consequence is that clinicians and researchers serving these populations either administer instruments in a language not native to the respondent, which compromises data quality and equity of care, or rely on ad hoc translations that lack the validation evidence required by current guidance.

A second consequence is invisible in single-centre work but visible in the literature. Cross-cultural prevalence estimates of disease impact, treatment-response comparisons across regions, and the generalisability of trial findings rely on PRO data that are measured equivalently across the linguistic and cultural settings in which the instrument is used. When adaptation is inadequate, the resulting measurement error contaminates downstream evidence synthesis (Reeve et al., 2013).

General-purpose cross-cultural adaptation guidelines exist. Beaton and colleagues (2000) formalised the five-stage forward and backward translation model; the Mapi Institute Linguistic Validation Protocol extended that model with reconciliation and patient testing; the ISPOR Task Force on Translation and Cultural Adaptation issued good-practice principles applicable across health-state measures (Wild et al., 2005); and the World Health Organization developed its own translation process. These frameworks are valuable, but they are general. They address adult populations and chronic-disease measures broadly. None of them is calibrated for the specific demands of pediatric rheumatology, where four characteristics shape adaptation in ways general protocols do not anticipate.

First, pediatric rheumatology depends on dual-perspective

reporting: a youth self-report from the affected adolescent and a parallel caregiver proxy-report. Adaptation of one form without the other, or adaptation that fails to preserve the alignment between forms, undermines the instrument's structure. Second, the cognitive and developmental variability of the pediatric respondent demands age-appropriate language, idiomatic clarity, and cognitive-interview techniques tuned for adolescents rather than adults (Eiser and Morse, 2001). Third, the disease-specific content domains in juvenile rheumatic conditions, including pain and hurt, treatment burden, communication with the care team, and worry about the future course of disease, require domain-specific cultural equivalence checks not covered in generic protocols. Fourth, validation studies in this population are typically conducted with small samples because the underlying diseases are rare; small-sample psychometric strategies are addressed only superficially in general adaptation guidance.

These four characteristics, taken together, define the gap that the present monograph addresses. The proposed methodology is not an extension of an existing protocol with an added section on children. It is an integrated, end-to-end workflow in which pediatric considerations are structural rather than supplementary.

Aim and Objectives

The aim of the monograph is to develop an integrated methodological framework for the cross-cultural adaptation and validation of patient-reported outcome measures in pediatric rheumatology, and to demonstrate the operation of that framework through a set of illustrative cases drawn from the published literature and the author's own validation work.

Five objectives organise the work. The first is to consolidate the theoretical foundations of PRO measurement and HRQOL in pediatric rheumatology, including the conceptual frameworks that underpin instrument design, the characteristics of the pediatric population that affect measurement, and the principal instruments currently in use. The second is to review existing cross-cultural adaptation guidelines, content-validity methods, and psychometric evaluation approaches relevant to adaptation, and to identify where these guidelines fall short for the pediatric rheumatology context. The third is to formulate the Cross-Cultural PRO Adaptation Framework for Pediatric Rheumatology, here abbreviated as CC-PRO-PR, as an integrated methodology comprising pre-adaptation assessment, linguistic and cultural adaptation, content validity and cognitive debriefing, psychometric evaluation, and implementation and dissemination. The fourth is to demonstrate the application of the framework through four illustrative cases: the author's own PedsQL Rheumatology Module 3.0 Hindi adaptation; published CHAQ adaptations across languages; the JAMAR international adaptation program coordinated through PRINTO; and the cross-cultural adaptations of the PRQL instrument. The fifth is to outline practical implementation pathways, with stakeholder-specific recommendations for clinicians, researchers, instrument developers, and funding bodies, alongside an explicit account of the framework's limitations, ethical and equity dimensions, and directions for future development.

Scientific Novelty and Practical Significance

The methodological novelty of the work lies in three contributions. The first is the framework itself: to the author's knowledge, no prior published methodology has integrated linguistic validation, cultural equivalence, content validity, psychometric evaluation, and implementation planning into a single workflow specifically calibrated for pediatric rheumatology. Existing guidance addresses one or two of these dimensions, or addresses all of them in a generic adult-population frame.

The second contribution is a structured decision protocol for handling youth self-report and caregiver proxy-report discrepancies during cross-cultural adaptation. A discrepancy between the two perspectives may indicate a translation problem, a cultural-equivalence issue, or a genuine difference in perspective that the instrument is designed to capture. Existing literature treats proxy-report disagreement primarily as a measurement nuisance to be reported, rather than as a diagnostic signal whose interpretation directs methodological action. Chapter 3 develops a protocol that classifies the source of disagreement and prescribes the methodological response.

The third contribution is a psychometric evaluation pipeline calibrated for the small samples typical of pediatric rheumatology validation studies, where rare-disease epidemiology limits recruitment. The framework specifies which reliability and validity coefficients can be trusted under such constraints, how known-groups validity can compensate for limited factor-analytic power, and what reporting standards align with COSMIN recommendations under small-sample conditions (Mokkink et al., 2010).

The practical significance of the framework is its reproducibility. Researchers planning a new adaptation study in pediatric rheumatology can follow the framework step by step, document their work in a structure that supports peer review and replication, and integrate the resulting instrument into clinical care with a clear pathway. Clinicians can evaluate adapted instruments offered for clinical use against the framework's criteria. Instrument developers can partner with adapting teams on a basis that protects copyright and supports cumulative international evidence. Institutions and funding bodies receive evaluation criteria for adaptation proposals. The cumulative outcome is broader, more equitable access to validated HRQOL assessment in the linguistically diverse populations that pediatric rheumatology serves.

Table I.1. summarises the conceptual scope of the framework as developed across the following chapters and previews the dimensions on which it differs from existing general-purpose guidance.

Module	Pediatric-rheumatology-specific element	Where general guidance falls short
Module 1: Pre-adaptation assessment	Instrument selection matched to age band, disease, and clinical purpose; copyright clearance with the developer; stakeholder mapping including pediatric rheumatologist, physiotherapist, occupational therapist, and caregiver.	General guidance treats stakeholder mapping as optional; pediatric rheumatology requires it because the multidisciplinary care team co-owns the data.
Module 2: Linguistic and cultural adaptation	Dual-form management preserving alignment between the youth self-report and caregiver proxy-report versions; cultural adjustments tuned for adolescent register.	Beaton et al. (2000) and the Mapi protocol can be applied to dual-form instruments by running the protocol on each form (as the Mapi Institute has done for the PedsQL family), but neither prescribes a coordinated cross-form alignment step as a structural element of the protocol.
Module 3: Content validity and cognitive debriefing	Expert panel reflecting the pediatric multidisciplinary team; cognitive debriefing with both the youth and the caregiver; iterative revision protocol.	ISPOR principles (Wild et al., 2005) provide expert review and cognitive debriefing guidance applicable across age groups but do not specify pediatric cognitive interviewing or dual-perspective reporting as structural elements of the protocol.
Module 4: Psychometric evaluation	Sampling strategies for small rare-disease populations; youth-caregiver agreement analysed via intraclass correlation and Bland-Altman; known-groups validity by disease severity.	COSMIN sample-size recommendations are developed primarily for adult populations; small-sample pediatric strategies are underspecified.
Module 5: Implementation and dissemination	Clinical integration including training, electronic health record embedding; documentation supporting cumulative international evidence; maintenance triggers for language drift and new age-band extension.	Implementation and dissemination are largely absent from general adaptation guidelines, which end at psychometric reporting.

Source: author's synthesis of the framework presented in Chapter 3.

Structure of the Monograph

The monograph proceeds from theoretical foundations through critical review and framework development to illustration and implementation. Chapter 1 establishes the measurement context. It defines patient-reported outcome, health-related quality of life, and patient-centred outcome; reviews the conceptual frameworks of Wilson and Cleary (1995) and the FDA PRO guidance; surveys the major juvenile rheumatic conditions and their HRQOL impact; describes the specific pediatric measurement challenges of developmental variability, dual-perspective reporting, and age-banded instrument design; and provides a comparative overview of the principal instruments in use, including the PedsQL Rheumatology Module 3.0, the Childhood Health Assessment Questionnaire, the Juvenile Arthritis Multidimensional Assessment Report, the Pediatric Rheumatology Quality of Life Scale, and the Patient-Reported Outcomes Measurement Information System (PROMIS) pediatric measures. The chapter closes with the case for cross-cultural adaptation.

Chapter 2 reviews the existing cross-cultural adaptation guidelines and identifies the gaps that justify the proposed framework. It examines the Beaton et al. (2000) five-stage model, the Mapi Institute Linguistic Validation Protocol, the ISPOR good-practice principles, and the WHO process; surveys qualitative and quantitative content-validity methods, including the approaches of McKenzie and colleagues, Wallace and colleagues, the content validity ratio of Lawshe (1975), and the content validity index of Polit and Beck (2006); reviews relevant psychometric evaluation methods aligned with COSMIN; and identifies five categories of gap for the pediatric rheumatology context.

Chapter 3 develops the CC-PRO-PR framework as an integrated methodology. It articulates the framework's concept and design principles; specifies each of five modules in detail, namely pre-adaptation assessment, linguistic and cultural adaptation, content validity and cognitive debriefing, psychometric evaluation, and implementation and dissemination; presents the decision protocol for youth self-report and caregiver proxy-report discrepancies; and compares the framework with Beaton et al., the Mapi protocol, and the ISPOR principles. This chapter is the central methodological contribution of the monograph.

Chapter 4 demonstrates the framework's application through four illustrative cases, each presented in the same five-

module structure for comparability. The cases are the author's own PedsQL Rheumatology Module 3.0 Hindi adaptation; published CHAQ adaptations across multiple languages; the JAMAR international adaptation program coordinated through the Paediatric Rheumatology International Trials Organisation; and published cross-cultural adaptations of the Pediatric Rheumatology Quality of Life Scale. A cross-case synthesis identifies the patterns that inform refinement of the framework.

Chapter 5 addresses implementation, limitations, and future directions. It offers a practical step-by-step implementation roadmap with resource and timeline estimates, stakeholder-specific recommendations, a candid account of the framework's limitations, future directions including extension to other pediatric subspecialties and integration with electronic and computerised adaptive PRO collection, and an explicit treatment of ethical and equity considerations in pediatric cross-cultural adaptation. The Conclusion summarises the integrated framework, confirms the scientific contribution, and identifies the reproducible workflow as the monograph's practical legacy.

CHAPTER 1. THEORETICAL FOUNDATIONS OF PATIENT-REPORTED OUTCOMES IN PEDIATRIC RHEUMATOLOGY

Chapter 1 establishes the theoretical and clinical context for the framework presented later in the monograph. The chapter defines patient-reported outcome (PRO) and health-related quality of life (HRQOL), reviews the conceptual frameworks that underpin instrument design, surveys the principal pediatric rheumatic conditions and their impact on HRQOL, characterises the specific measurement challenges of the pediatric population, and provides a comparative overview of the established PRO instruments in current use. The chapter closes with the case for cross-cultural adaptation as a methodological priority.

1.1 The Concept of Patient-Reported Outcomes and Health-Related Quality of Life

1.1.1 Definitions

A patient-reported outcome is any report of the status of a patient's health condition that comes directly from the patient, without interpretation by a clinician or anyone else. The definition, codified in the United States Food and Drug Administration guidance on the use of PRO measures in medical product development, encompasses symptom report, functional status, satisfaction with care, treatment burden, and any other health attribute that the patient can report on. In pediatric practice the same construct is operationalised through age-appropriate instruments, with caregiver proxy-report used as a complement when the youth cannot complete the instrument or when an additional perspective is methodologically warranted.

Health-related quality of life is a related but distinct construct. HRQOL refers to the patient's perception of the impact of disease and treatment on functioning across physical, psychological, and social domains. While PRO is a methodological category that describes who provides the report, HRQOL is a content category that describes what is being reported. A PRO instrument may measure HRQOL, symptom severity, treatment satisfaction, or any combination, depending on its design. HRQOL instruments in pediatric rheumatology are typically PRO measures because the patient and caregiver are the only sources who can report on subjective impact.

Patient-centred outcome is a broader category that includes PROs but also includes outcomes that are clinically observed yet selected because of their importance to patients, such as time to remission, time to return to school, or treatment-related adverse events that patients prioritise. The distinction matters when an adaptation team selects the instrument to translate: a PRO measure focused on HRQOL is methodologically different from a patient-centred composite endpoint, and the adaptation strategies differ accordingly.

1.1.2 Conceptual Frameworks

Three conceptual frameworks anchor PRO measurement as it is practised today. The first is the model proposed by Wilson and Cleary (1995), which links biological and physiological variables to symptom status, then to functional status, then to general health perceptions, with HRQOL as the integrating endpoint. The model is causal in orientation: it argues that biological abnormality affects symptoms, symptoms affect function, function affects perception, and perception integrates the chain into a single HRQOL judgment. In pediatric rheumatology, the model predicts that an effective HRQOL instrument will be sensitive to changes anywhere in the chain. A successful disease-modifying therapy that reduces synovitis should improve symptom status, which should improve daily activity capacity, which should improve perception, which should produce a measurable HRQOL change. Instruments that fail to reflect such a cascade either lack content coverage at some link in the chain or lack the responsiveness to detect change. The Wilson and Cleary model therefore serves as both a design guide for new instruments and an interpretive guide for existing ones.

The second framework is the FDA PRO Guidance (2009), which sets requirements for using PRO data to support labelling claims in medical product applications (Patrick et al., 2007). The guidance specifies that the conceptual framework of the instrument must be explicit, content validity must be established with patient input, the measurement properties must be appropriate for the intended use, and any cross-cultural adaptation must preserve the conceptual framework while addressing linguistic and cultural equivalence. The guidance does not prescribe a single adaptation method but does prescribe the level of evidence required when adaptation is performed. This regulatory expectation has

been a principal driver of the systematic adaptation methods reviewed in Chapter 2.

The third framework is the International Society for Quality of Life Research minimum standards for PRO measures (Reeve et al., 2013). ISOQOL specifies the measurement properties that any PRO instrument should demonstrate before being used to inform patient-centered outcomes research or comparative effectiveness research. These standards apply across conditions and ages and have become a reference point for instrument selection in pediatric rheumatology as elsewhere. The standards bring measurement and clinical perspectives together: an instrument that meets ISOQOL standards is reliable, valid, and responsive in ways that support meaningful interpretation at the individual and group level.

Figure 1.1 represents the Wilson and Cleary causal chain in schematic form, with the additions that pediatric and cross-cultural measurement require. The original model orders biological variables, symptom status, functional status, general health perceptions, and overall HRQOL in a left-to-right causal flow. The schematic adds two pediatric extensions: the dual-perspective measurement layer, in which the youth and caregiver views of each link are captured in parallel, and the cross-cultural layer, in which the linguistic and cultural filter through which each construct is reported shapes the data that an instrument actually collects.

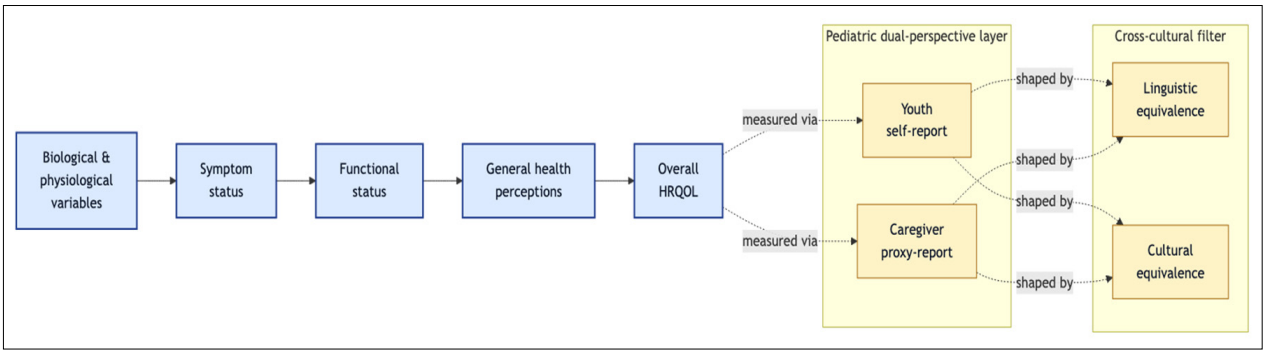

Figure 1.1. Wilson and Cleary conceptual model of HRQOL with pediatric dual-perspective and cross-cultural extensions. Source: schematic by the author after Wilson and Cleary (1995); pediatric and cross-cultural extensions developed for the present monograph.

1.1.3 The Shift from Clinician-Rated to Patient-Reported Assessment

Until the late twentieth century, outcome assessment in rheumatology was dominated by clinician-rated indices such as joint counts, physician global assessments, and laboratory markers. These remain essential but are insufficient because they do not capture the impact of disease on the patient's life as the patient experiences it (Brunner et al., 2004). The transition to PRO-augmented assessment was driven by three converging forces. The first was empirical evidence that patient-reported and clinician-rated outcomes correlate imperfectly, with patient-reported outcomes adding incremental information about treatment response, disability, and well-being that clinician indices miss. The second was the rise of value-based and patient-centred care, in which the patient's experience of care is treated as an outcome in its own right (Calvert et al., 2018). The third was the regulatory acceptance of PRO endpoints for labelling claims, which made instrument quality a commercial as well as a methodological imperative. Within pediatric rheumatology, the development of the Childhood Health Assessment Questionnaire (Singh et al., 1994), the PedsQL family of measures (Varni et al., 2002), the Juvenile Arthritis Multidimensional Assessment Report (Consolaro et al., 2011), and the Pediatric Rheumatology Quality of Life Scale (Filocamo et al., 2010) each represent responses to this trajectory, applied to the pediatric population. Comparative studies of these measures have shown that disease-specific HRQOL assessment captures dimensions of impact that generic measures miss (Carle et al., 2011).

1.2 Pediatric Rheumatology as a Measurement Context

1.2.1 Major Juvenile Rheumatic Conditions and their HRQOL Impact

Pediatric rheumatic diseases comprise a heterogeneous group of immune-mediated disorders whose principal common feature is musculoskeletal or systemic inflammation beginning in childhood. Juvenile idiopathic arthritis is the most prevalent of these conditions. Under the International League of Associations for Rheumatology classification adopted in Edmonton in 2001 and published by Petty et al. (2004), JIA is defined as arthritis of unknown origin beginning before age sixteen and persisting for at least six weeks, after exclusion of other causes. The classification distinguishes seven categories: systemic arthritis, oligoarthritis (persistent and extended), polyarthritis rheumatoid-factor-negative, polyarthritis rheumatoid-factor-positive, psoriatic arthritis, enthesitis-related arthritis, and undifferentiated arthritis. The categories differ in prognosis, treatment, and HRQOL impact, and the classification therefore matters for instrument selection in adaptation work.

Worldwide epidemiology of JIA has been the subject of multiple reviews. Manners and Bower (2002) drew attention to the broad geographic variation in reported prevalence, attributable in part to case-definition heterogeneity and in

part to genuine population differences. Thierry et al. (2014) updated the picture with a systematic review and meta-analysis covering global incidence and prevalence, reporting ranges of broadly an order of magnitude depending on case definition, data source, and region. Subsequent registry- and claims-based studies have refined the regional estimates without overturning the broad conclusion: JIA is uncommon but not rare, with incidence and prevalence varying meaningfully across populations, and the variation itself motivates the need for instruments that perform comparably across those populations.

Other major juvenile rheumatic conditions add further heterogeneity to the measurement task. Juvenile systemic lupus erythematosus typically presents with multi-system involvement and a fluctuating disease course; its HRQOL impact spans physical limitation, cognitive symptoms, psychosocial burden, and treatment-related adverse effects (Hersh et al., 2010). Juvenile dermatomyositis combines characteristic skin and muscle involvement; its HRQOL profile is dominated by functional limitation and cosmetic concerns. Juvenile fibromyalgia centres on chronic widespread pain and fatigue, with substantial HRQOL impact mediated through sleep, school attendance, and psychosocial functioning. Juvenile spondyloarthropathies present with enthesitis and axial involvement and affect mobility and posture across the developmental period. Each condition interacts with the developmental stage of the young person and with the family context, shaping the HRQOL impact in ways that an instrument must capture (Moorthy et al., 2010; Shaw et al., 2006; Ringold et al., 2009).

The HRQOL signature of each condition has implications for instrument design and adaptation. A measure intended for JIA must capture pain on movement, morning stiffness, and functional limitation in school and play activities; a measure for juvenile SLE must additionally capture systemic fatigue, cognitive symptoms, and the burden of immunosuppressive treatment; a measure for juvenile fibromyalgia must capture chronic widespread pain in a way that distinguishes it from arthritis-attributable pain. Disease-specific instruments achieve this content fit at the cost of cross-condition comparability; generic instruments achieve cross-condition comparability at the cost of disease-specific sensitivity. The choice between disease-specific and generic measurement, or the use of both together, is itself a methodological decision that an adaptation team must understand before translation begins.

Disease activity and disease course interact with HRQOL in pediatric rheumatology in ways that adult-population frameworks do not anticipate. A flare in juvenile arthritis disrupts school attendance, peer interaction, and family routines simultaneously; remission allows those routines to resume, but residual fatigue or anxiety about future flares may persist. HRQOL data therefore vary not only with disease activity at the moment of measurement but with the developmental trajectory of the young person and the cumulative experience of disease. Longitudinal use of HRQOL instruments in clinical care and research is the methodologically appropriate response, and adaptation work must preserve the instrument's responsiveness to change across time as well as its cross-sectional reliability.

Table 1.1. summarises the worldwide epidemiological context of the principal juvenile rheumatic conditions covered by the established PRO instruments reviewed in Section 1.3. The values are presented as ranges to convey the variability across reviews; specific values depend on case definition, data source, and population.

Condition	Incidence (qualitative summary)	Prevalence (qualitative summary)	Principal HRQOL domains affected
Juvenile idiopathic arthritis (JIA), all categories	Uncommon but not rare; reported incidence varies by region and case definition (Manners and Bower, 2002; Thierry et al., 2014)	Higher than incidence by prevalence-pool dynamics; varies an order of magnitude across reviews	Physical function, pain, daily activities, school participation
Juvenile systemic lupus erythematosus (jSLE)	Rare; childhood-onset systemic autoimmune condition	Lower than adult-onset SLE; multi-system involvement	Fatigue, cognitive symptoms, treatment burden, psychosocial
Juvenile dermatomyositis (JDM)	Rare; characteristic skin and muscle involvement	Lower than JIA; persistent functional impact	Muscle function, cosmetic concerns, school absence
Juvenile fibromyalgia	Population-based incidence less well established than JIA	Community samples report measurable prevalence in adolescents	Pain, fatigue, sleep, school participation, psychosocial
Juvenile spondyloarthropathies	Uncommon; enthesitis and axial involvement	Lower than oligoarthritis or polyarthritis subcategories of JIA	Mobility, posture, pain, sleep

Source: qualitative synthesis after Manners and Bower (2002), Thierry et al. (2014), and Petty et al. (2004). Specific incidence and prevalence values vary substantially across reviews because case definition, data source, and population shape the estimate; readers are referred to the primary sources for the figures associated with each population studied.

1.2.2 Specific Pediatric Measurement Challenges

Three challenges distinguish PRO measurement in the pediatric setting from PRO measurement in adult populations. The first is developmental variability. A nine-year-old, a thirteen-year-old, and a seventeen-year-old differ in cognitive capacity, linguistic competence, and ability to introspect on health states. An instrument that is appropriate at one age may be confusing at another. The standard response has been age-banded instrument design, in which separate forms or item subsets are used for younger child (typically ages five to seven), child (ages eight to twelve), and youth or adolescent (ages thirteen to eighteen) respondents. The PedsQL family and the JAMAR illustrate the approach. Age banding has implications for adaptation, because each band must be adapted with attention to the developmental register of its target age group.

The second challenge is the dual-perspective character of pediatric PRO measurement. The youth provides a self-report, and the caregiver provides a proxy-report, intended to inform the clinical picture when the young person is too young, too unwell, or too cognitively limited to complete the instrument. Caregiver proxy-report is not a substitute for youth self-report but a complement that captures aspects the caregiver observes from outside. Eiser and Morse (2001) reviewed agreement between youth and caregiver reports of HRQOL across conditions and showed that agreement is generally moderate, with systematic patterns: caregivers tend to report worse HRQOL than youths in domains observable from outside, such as physical functioning, and tend to under-report HRQOL impact in domains less visible to an observer, such as emotional well-being. Adaptation work that ignores this structure risks producing two instruments that drift apart in unintended ways.

The third challenge is the developmental threshold for independent completion. Evidence from instrument-development work indicates that children younger than approximately five years cannot complete self-report PRO instruments on their own, that children aged five to seven can complete simplified picture-based or short forms with assistance, and that children aged eight and above can complete standard self-report forms designed for their age band. The threshold is not a sharp line but a graded transition, and instrument designers have responded with age-banded forms calibrated to these gradients. Adaptation must preserve these gradations in the target language; a translation that, by virtue of longer or more complex target-language sentences, raises the cognitive load above the intended level shifts the threshold upward and undermines the instrument.

A fourth consideration cuts across the first three. The young person sits within a family system, and the caregiver who completes the proxy-report sits within that same system. Family functioning, the caregiver's own health, and the family's cultural framing of illness all shape the data both respondents provide. Adaptation work therefore needs to consider not only the linguistic and cognitive features of items but the cultural framing within which the family responds. An item asking the youth how worried they are about their illness, for example, will receive different responses depending on whether the cultural norm is to discuss illness openly with caregivers or to protect caregivers from worry by minimizing symptoms. The instrument should measure HRQOL, not the cultural disclosure norm; adaptation therefore requires attention to such norms during cognitive interviewing.

1.2.3 Clinical and Research Role of PRO Data

The clinical use of PRO data in pediatric rheumatology has three principal modes. The first is monitoring disease burden between clinic visits, often through electronic PRO platforms that allow the youth and caregiver to complete questionnaires at home and bring the results to the consultation (Haverman et al., 2013). The second is evaluating treatment response in routine care, where a PRO instrument provides a longitudinal record of the patient's experienced health alongside clinical markers (Brunner et al., 2009). The third is supporting shared decision-making, where PRO data give voice to the patient's perspective at the moment a treatment choice is being made. Each mode requires that the instrument be valid, reliable, responsive to change, and interpretable to both clinician and family.

The research use of PRO data spans clinical-trial endpoints, comparative effectiveness research, health-services research, and registry studies. Modern guidance, including the SPIRIT-PRO Extension and the CONSORT-PRO Extension, sets standards for the planning and reporting of PRO endpoints in clinical trials. International registry initiatives associated with the JAMAR illustrate the scale at which PRO data can be aggregated when adaptation work has produced equivalent instruments in multiple languages. Adolescent respondents themselves report PRO measurement to be most valuable when it informs self-management and clinical communication (Stinson et al., 2008). Each role amplifies the importance of cross-cultural adaptation, because PRO data that are not comparably measured across the populations contributing to a registry or trial cannot be pooled with confidence.

1.3 Overview of Established PRO Instruments in Pediatric Rheumatology

Section 1.3 reviews the principal PRO instruments currently in use in pediatric rheumatology with a focus on their structure, age range, validation history, and the adaptation activity they have generated. The section closes with a comparative table that summarises the principal attributes across instruments.

1.3.1 PedsQL Rheumatology Module 3.0

The Pediatric Quality of Life Inventory Rheumatology Module 3.0, developed by Varni and colleagues (2002), is a disease-specific HRQOL instrument designed for use with the PedsQL Generic Core Scales in pediatric rheumatic disease

populations. The module is organised into five domain scales: pain and hurt, daily activities, treatment, worry, and communication. The instrument provides parallel youth self-report and caregiver proxy-report forms across age bands covering young children through adolescents, with age-adapted item phrasing in each band. Reported psychometric performance in the original English-language validation was acceptable across reliability and validity metrics for both the self-report and the proxy-report forms (Varni et al., 2002). The module has been used as the disease-specific complement to the PedsQL Generic Core Scales in multiple JIA research studies and has been the subject of multiple cross-cultural adaptation projects, including the author's Hindi adaptation, which serves as Case 1 in Chapter 4.

1.3.2 Childhood Health Assessment Questionnaire (CHAQ)

The Childhood Health Assessment Questionnaire, developed by Singh and colleagues (1994), is a pediatric adaptation of the adult Health Assessment Questionnaire (HAQ) for use in juvenile rheumatic disease populations. CHAQ assesses physical function across multiple domains of daily activity, with parallel youth self-report and caregiver proxy-report forms. The instrument has been the most widely translated and adapted measure in pediatric rheumatology, with adaptations in dozens of languages reported in the literature. Its functional focus complements the broader HRQOL coverage of the PedsQL Rheumatology Module and the JAMAR; CHAQ is therefore often used in combination with one of these broader measures rather than as a standalone instrument. CHAQ adaptations across languages serve as Case 2 in Chapter 4.

1.3.3 Juvenile Arthritis Multidimensional Assessment Report (JAMAR)

The Juvenile Arthritis Multidimensional Assessment Report, developed by Consolaro and colleagues (2011) under the auspices of the Paediatric Rheumatology International Trials Organisation (PRINTO), is a multidimensional instrument that combines patient-reported outcome assessment with structured caregiver and clinical data collection in a single report. JAMAR integrates physical function, pain, HRQOL, well-being, and disease activity perception, and is designed to support both clinical care and registry use at scale. The PRINTO-coordinated multi-language adaptation program is the largest organised cross-cultural effort in the field and has produced validated versions in many languages across multiple continents. The adaptation program serves as Case 3 in Chapter 4.

1.3.4 Pediatric Rheumatology Quality of Life Scale (PRQL)

The Pediatric Rheumatology Quality of Life Scale, developed by Filocamo and colleagues (2010) and validated alongside the JAMAR family, is a short HRQOL instrument designed for ease of administration in routine clinical practice. PRQL covers physical and psychosocial HRQOL across a small number of items, supporting quick completion in busy clinics while preserving disease-specific relevance. Cross-cultural adaptations of PRQL into multiple languages have been reported, with the instrument serving as a complement to the more comprehensive JAMAR and PedsQL Rheumatology Module. PRQL adaptations serve as Case 4 in Chapter 4.

1.3.5 PROMIS Pediatric measures

The Patient-Reported Outcomes Measurement Information System (PROMIS) Pediatric measures, developed under a coordinated initiative of the United States National Institutes of Health (Ader, 2007; Cella et al., 2007), represent a different methodological approach. Instead of fixed-form scales with fixed item sets, PROMIS Pediatric is organised as item banks calibrated under item response theory (IRT), from which short fixed forms or computerised adaptive tests can be drawn for specific applications (Irwin et al., 2010; DeWalt et al., 2015). Item banks span pain interference, fatigue, physical function, anxiety, depressive symptoms, peer relationships, and other domains relevant to chronic pediatric illness. The Item-response-theory (IRT) calibration permits cross-form comparability and supports efficient measurement with reduced respondent burden. PROMIS Pediatric is not a rheumatology-specific instrument; it offers generic measurement of domains that are relevant in rheumatology alongside other pediatric chronic conditions. Its growing use in pediatric rheumatology research reflects both its psychometric strengths and the infrastructure that supports cross-language calibration (Forrest et al., 2012).

1.3.6 Comparative Analysis

Table 1.2. summarises the structural attributes of the five instruments reviewed in this section. The summary supports the comparative discussion in Chapter 4, where each instrument's adaptation experience is examined through the lens of the proposed framework. Comparative reviews of pediatric rheumatology HRQOL measures across instrument families (Hoffman and Mahoney, 2009; Quirke et al., 2014; Riley et al., 2004) have informed the structural summary presented here.

Instrument	Construct focus	Forms / age bands	Adaptation history	Position in pediatric rheumatology
PedsQL Rheumatology Module 3.0 (Varni et al., 2002)	Disease-specific HRQOL across pain-hurt, daily activities, treatment, worry, communication	Youth self-report and caregiver proxy-report across age bands from young child to adolescent	Multiple cross-cultural adaptations including French, German, Italian, Russian, Slovenian, Spanish; Hindi adaptation by the author	Used in combination with PedsQL Generic Core Scales for both research and clinical care

CHAQ (Singh et al., 1994)	Physical function in daily activities	Youth self-report and caregiver proxy-report; age bands defined by instrument version	Most widely adapted instrument in pediatric rheumatology; many language versions reported	Standard functional measure; complements HRQOL instruments rather than substituting for them
JAMAR (Consolaro et al., 2011)	Multidimensional: physical function, pain, HRQOL, well-being, disease activity perception, caregiver and clinical data integrated	Youth self-report and caregiver proxy-report; integrated single report	PRINTO-coordinated multi-language adaptation program at scale	Designed for both clinical care and international registry use; central to PRINTO research infrastructure
PRQL (Filocamo et al., 2010)	Short HRQOL across physical and psychosocial domains	Short instrument; youth self-report and caregiver proxy-report	Adaptations into multiple languages reported in published literature	Brief HRQOL measure suitable for routine clinical use; complements JAMAR family
PROMIS Pediatric measures	Generic HRQOL-relevant domains under item response theory calibration	Item banks supporting fixed forms and computerised adaptive testing	Coordinated cross-language calibration via NIH PROMIS infrastructure; growing pediatric translations	Cross-condition generic measurement; used in rheumatology alongside disease-specific instruments

Source: synthesised from Varni et al. (2002), Singh et al. (1994), Consolaro et al. (2011), Filocamo et al. (2010), and PROMIS Pediatric developmental publications. Adaptation history summaries refer to the literature reviewed in Chapter 4 and reflect activity reported in peer-reviewed sources up to the time of writing.

The instruments differ along several axes. PedsQL RM 3.0, CHAQ, JAMAR, and PRQL are rheumatology-specific or rheumatology-developed; PROMIS Pediatric is generic. PedsQL RM, JAMAR, and PRQL focus on HRQOL or its multidimensional extension; CHAQ focuses on physical function. JAMAR integrates patient, caregiver, and clinical data in a single report; the others use parallel forms that are scored separately. PROMIS Pediatric is calibrated under IRT; the others use classical-test-theory scoring with summed scales. Each axis has implications for adaptation. An instrument with parallel forms requires dual-form alignment during translation; an IRT-calibrated instrument requires preservation of item parameters across the target language; a multidimensional report requires coordinated adaptation of youth and caregiver components. Chapter 3 addresses these implications in the framework's modules.

Two additional considerations sharpen the comparative picture. The first is the choice between summative and adaptive measurement. Classical-test-theory instruments produce a single sum score per scale, which is straightforward to interpret and compare across respondents but assumes that all items contribute equally to the construct. IRT-calibrated item banks weight items by their information content for the latent construct, permit computerised adaptive testing, and allow the same construct to be measured with shorter item sets at no loss of precision. Cross-cultural adaptation under IRT requires demonstration of measurement invariance across the target population, which is statistically more demanding than the equivalence checks expected for classical instruments. The choice of instrument is therefore partly a choice of adaptation method.

The second consideration is the integration of disease-specific and generic measurement. The PedsQL family was designed around exactly this integration: the PedsQL Generic Core Scales provide cross-condition HRQOL measurement, and disease-specific modules including the Rheumatology Module 3.0 provide condition-relevant detail. The architecture supports comparisons across pediatric chronic conditions while preserving sensitivity to the specific HRQOL features of each. PROMIS Pediatric offers a different integration, in which generic item banks cover domains relevant in rheumatology and across other conditions; rheumatology-specific content is provided by external instruments. JAMAR achieves integration within a single rheumatology-specific report rather than across instruments. Each integration strategy has implications for adaptation work, and the framework introduced in Chapter 3 treats these implications structurally.

1.4 The Case for Cross-Cultural Adaptation

1.4.1 Why Direct Translation is Insufficient

A translated instrument is not an adapted instrument (Beaton et al., 2000; Wild et al., 2005). Three issues distinguish adaptation from translation. The first is conceptual non-equivalence. A term that has a clear meaning in the source language may correspond to multiple distinct meanings in the target language, or may carry connotations that shift the construct being measured. The English term hurt, used in the PedsQL Rheumatology Module pain-and-hurt scale, illustrates the problem. In English, hurt and pain are used together to capture both physical pain and broader discomfort; in many target languages a single word that covers the same conceptual range is not available, and the team must decide whether to use two words, to combine into one with a footnote explaining intent, or to substitute a culturally appropriate alternative. Each choice has consequences for what the instrument actually measures.

The second issue is cultural specificity of symptom expression. Symptoms are reported through cultural filters. A youth who reports worry about future health may use different linguistic markers, different intensity vocabulary, and different relational framing across cultures. An instrument that asks about worry without attention to these filters may under or over capture the construct in the target population. Cognitive interviewing during adaptation, addressed in Chapter 3, is the principal method for surfacing such filters and for adjusting items accordingly.

The third issue is linguistic ambiguity introduced by translation. Forward translation may produce a target-language version that is grammatically correct but ambiguous, or that uses register inappropriate for the youth respondent, or that shifts the cognitive load. Backward translation can detect some such issues by exposing departures from the source concept; cognitive interviewing detects others by surfacing the respondent's actual interpretation. Without these checks, a literally translated instrument may produce data that look like the source instrument's data but do not measure the same thing.

1.4.2 Consequences of Using Non-Adapted Instruments

The downstream consequences of inadequate adaptation are measurement error, biased prevalence and treatment-response estimates, and compromised pooled evidence. Measurement error in a non-adapted instrument is typically systematic rather than random, because respondents in the target population misinterpret items in similar ways, producing biased estimates of central tendency and variance (Mokkink et al., 2010, 2018). Prevalence and treatment-response estimates derived from such instruments may suggest cross-cultural patterns that reflect measurement differences rather than substantive differences in disease impact. Meta-analytic pooling across populations using imperfectly adapted instruments imports the measurement bias into the synthesised effect (Acquadro et al., 2008).

Concrete language-gap examples sharpen the picture. Until the recent Hindi adaptation documented as Case 1 in Chapter 4, leading PRO instruments in pediatric rheumatology had no validated Hindi version, despite the size of the Hindi-speaking pediatric population. Validated versions in Bengali, Tamil, Swahili, Vietnamese, and other major languages are limited or absent for several leading instruments. The consequence for clinicians in regions where these languages predominate is a choice between administering the instrument in a non-native language, with the comprehension issues this raises, and using an unvalidated translation, with the measurement issues this raises. Both choices compromise the data and, by extension, the care (Trapanotto et al., 2009).

1.4.3 The Argument for a Pediatric-Rheumatology-Specific Adaptation Methodology

General-purpose adaptation guidelines exist and are reviewed in Chapter 2. The argument for a pediatric-rheumatology-specific methodology rests on four observations developed across Section 1.2. The pediatric population brings developmental variability that requires age-banded design and band-specific adaptation. The dual-perspective structure of youth and caregiver report requires adaptation that preserves alignment between forms. The disease-specific content domains of juvenile rheumatic conditions, including pain, treatment burden, and worry, require domain-specific equivalence checks tuned to the expressive register of adolescents and the explanatory register of caregivers. The small samples typical of rare-disease validation studies require psychometric strategies calibrated for limited statistical power. General guidelines do not provide structural treatment of any of these four characteristics. The framework presented in Chapter 3 is the response.

CHAPTER 2. CRITICAL REVIEW OF EXISTING ADAPTATION GUIDELINES AND IDENTIFICATION OF GAPS

Chapter 2 reviews the principal guidelines that direct cross-cultural adaptation of patient-reported outcome measures, the content-validity and cultural-equivalence methods that operate within those guidelines, and the psychometric evaluation methods relevant to adaptation. The chapter then identifies five categories of gap that justify the pediatric-rheumatology-specific framework developed in Chapter 3, and closes with a synthesis that frames the framework's design intent.

2.1 General-Purpose Cross-Cultural Adaptation Guidelines

Four general-purpose protocols dominate cross-cultural adaptation practice in PRO measurement. The Beaton et al. (2000) guidelines, published in Spine, codify a five-stage forward and backward translation model and have become the most widely referenced point of departure for adaptation projects. The Mapi Institute Linguistic Validation Protocol extends the Beaton model with reconciliation and patient testing stages and is the operational protocol followed in many proprietary instrument adaptations, including the author's PedsQL Rheumatology Module Hindi adaptation discussed as Case 1 in Chapter 4. The ISPOR Task Force good-practice principles (Wild et al., 2005) provide consensus principles applicable across health-state measures, focusing on translation quality, conceptual equivalence, and harmonisation. The World Health Organization translation process applies to instruments deployed under WHO programs and emphasises both linguistic and cultural equivalence at the conceptual level. The four protocols converge on a small number of operational principles but differ in granularity, deliverables, and quality criteria.

2.1.1 Beaton et al. (2000): The Five-Stage Forward and Backward Translation Model

Beaton and colleagues (2000) describe a five-stage process.

Stage I is initial translation: two independent forward translators produce two independent target-language drafts from the source instrument. Stage II is synthesis of the translations: the two forward translators and a recording observer reconcile the two drafts into a single working version, documenting the resolution of every discrepancy. Stage III is back translation: two independent translators, blind to the source instrument, produce two back-translated English versions of the reconciled target-language draft. Stage IV is expert committee review: a multidisciplinary panel including methodologist, language professionals, the developer where possible, and the translation team reviews the entire material and produces the pre-final version. Stage V is pretesting: the pre-final version is administered to a sample drawn from the target population, and probe interviews verify that respondents understand the items as intended.

The Beaton model is influential because it operationalises conceptual equivalence as the outcome of a structured, multi-step process rather than as the verdict of a single bilingual reviewer. The model also makes explicit the documentation requirements at each stage, supporting downstream peer review of the adaptation work. The model has known limitations. It treats the target population sample for pretesting as a single homogeneous group, without specific guidance for pediatric respondents. It does not specify methods for handling parallel forms intended for different respondents, such as a youth self-report and a caregiver proxy-report, beyond running the process twice. It assumes that the developer is available for consultation, an assumption that does not always hold. And it focuses on linguistic and conceptual equivalence without integrating downstream psychometric evaluation as a structural part of the adaptation lifecycle (Sousa and Rojjanasrirat, 2011).

2.1.2 The Mapi Institute Linguistic Validation Protocol

The Mapi Institute Linguistic Validation Protocol, codified across multiple publications by Acquadro and colleagues, extends the Beaton model with operational refinements. Mapi treats forward translation, reconciliation, backward translation, developer review, harmonisation across language versions, and cognitive debriefing as discrete stages with documented deliverables. The protocol places particular emphasis on cognitive debriefing as a respondent-facing verification step, in which a small sample of target-population respondents complete the pre-final version under conditions that allow probing of comprehension difficulties and unintended interpretations (Acquadro et al., 2008).

The Mapi protocol is widely used in proprietary instrument adaptation because the Mapi Research Institute holds rights to many adaptations and offers a service-based implementation. The protocol's operational granularity is its strength: an adapting team following Mapi knows what each deliverable should contain, what each stage should document, and what the developer review will examine. The protocol's reliance on centralised coordination is also a limitation for adaptation teams working outside the Mapi network, who must reconstruct the operational discipline from the protocol's published descriptions without the institutional support that accompanies a Mapi-managed project.

2.1.3 The ISPOR Task Force Good-Practice Principles

Wild and colleagues (2005), writing as the ISPOR Task Force on Translation and Cultural Adaptation, articulate a set of principles of good practice for the translation and cultural adaptation of PRO measures. The principles include developer involvement, two independent forward translations, reconciliation, two independent backward translations, comparison of the back translation against the source, developer review of the back translation, cognitive debriefing in the target population, review of the debriefing results, proofreading, and a final written report. The ISPOR principles are not a stand-alone protocol but a normative framework that can be implemented through several specific procedures, including Beaton, Mapi, and WHO.

Two strengths distinguish the ISPOR principles. First, the principles are explicit about the role of the instrument developer at multiple stages: developer review of the back translation and developer review of the cognitive debriefing results are treated as structural rather than incidental. This emphasis supports the proprietary character of many PRO instruments, including those reviewed in Chapter 1. Second, the principles are explicit about written documentation as a deliverable, including a final written report that supports replication and peer review. Limitations include the absence of pediatric-specific guidance and the relatively brief treatment of psychometric evaluation, which the ISPOR principles delegate to downstream work after linguistic and conceptual equivalence have been established.

2.1.4 The World Health Organization Translation Process

The World Health Organization translation process applies to instruments deployed under WHO programs, including a number of major mental health and disability instruments. The process emphasises forward translation by a multidisciplinary panel, expert panel review of the forward translation, back translation by an independent translator, pretesting and cognitive debriefing, and a final version that meets both linguistic and cultural equivalence standards. The WHO process resembles the Beaton and Mapi protocols at the operational level but differs in two respects: the panel-based forward translation in place of two independent translators, and the explicit emphasis on technical, semantic, experiential, and conceptual equivalence as the four dimensions to be addressed in the adaptation work.

The WHO process has been the dominant reference in low- and middle-income country settings where WHO-

administered programs operate and where the program's adaptation guidance has shaped the available local infrastructure. Limitations include the protocol's primary focus on adult populations and instruments designed for population-level surveillance rather than disease-specific PRO measurement at the individual clinical level, which limits direct applicability to the pediatric rheumatology context (Guillemin et al., 1993).

2.1.5 Comparative Summary

Table 2.1. summarises the four protocols across stages, deliverables, quality criteria, and explicit applicability to pediatric populations. The comparison clarifies the gap that the framework presented in Chapter 3 addresses: each protocol offers a credible operationalisation of the general adaptation task, and each leaves pediatric-rheumatology-specific work underspecified.

Protocol	Stage structure	Documented deliverables	Quality criteria emphasised	Pediatric applicability
Beaton et al. (2000)	5 stages: initial translation, synthesis, back translation, expert committee, pretesting	Translation drafts, reconciliation report, back translations, expert committee report, pretest interview log	Conceptual equivalence; structured documentation; multidisciplinary review	Treats target population as single group; no pediatric-specific guidance
Mapi Linguistic Validation Protocol	Forward translation; reconciliation; back translation; developer review; harmonisation; cognitive debriefing	Reconciliation report; back translation report; developer review log; cognitive debriefing transcripts; final adapted version	Operational granularity; service-based implementation; cognitive debriefing emphasis	Operationally adaptable to pediatric work but no structural pediatric component
ISPOR Task Force (Wild et al., 2005)	A consensus set of principles spanning developer involvement through final written report	Stage reports per principle; cognitive debriefing report; final written report	Developer involvement; written documentation; back-translation review	No structural pediatric element; psychometric evaluation delegated downstream
WHO translation process	Panel-based forward translation; expert panel review; back translation; pretesting; final version	Forward translation report; expert review report; pretest report; final version	Technical, semantic, experiential, and conceptual equivalence as four explicit dimensions	Primary focus on adult population-level instruments; limited pediatric guidance

Source: synthesised from Beaton et al. (2000), Acquadro et al. (2008), Wild et al. (2005), and the published WHO translation guidance.

Two cross-cutting observations frame the comparison. First, the four protocols converge on a core operational sequence (forward translation, reconciliation, back translation, expert review, cognitive debriefing or pretesting) but differ in the granularity of documentation and in the specificity of the developer's role. Second, none of the four protocols treats psychometric evaluation as a structural extension of the adaptation lifecycle; psychometric work is described as a separate downstream activity, with the adapted instrument handed off as the deliverable. The framework in Chapter 3 closes this seam by integrating psychometric evaluation as a connected module rather than as a downstream successor activity (Tsang et al., 2017).

2.2 Content Validity and Cultural Equivalence Methods

Content validity describes the extent to which an instrument's items represent the universe of content relevant to the construct being measured (Lawshe, 1975; Polit and Beck, 2006). Cultural equivalence describes the extent to which that representation is preserved across linguistic and cultural settings. Section 2.2 reviews the principal qualitative and quantitative methods for evaluating content validity and cultural equivalence, with particular attention to the cognitive interviewing techniques that operationalise content validity at the respondent level.

2.2.1 Qualitative Content Validity

Qualitative content validity methods rely on expert review and respondent interpretation. The criteria proposed by McKenzie and colleagues and the related criteria developed by Wallace et al. organise expert review around explicit questions about each item's clarity, relevance, comprehensiveness of the item set, redundancy, and appropriateness for the target population. A panel of experts reviews the instrument item by item, recording comments on each item against each criterion, and a project coordinator consolidates the panel feedback into a structured revision plan. The strength of qualitative review is that experts surface conceptual problems that quantitative ratios cannot capture, such as items that combine two constructs in a single question or items whose response options do not form a coherent scale.

The limitation of qualitative review is its dependence on the composition and engagement of the expert panel. A panel that lacks pediatric rheumatology experience may miss disease-specific subtleties; a panel that lacks linguistic expertise in

the target language may miss subtleties of cultural framing; a panel without representation from the lived experience of the youth and caregiver may miss usability issues that emerge only in respondent contact. The framework presented in Chapter 3 addresses this through a structured panel composition tailored to the pediatric rheumatology setting.

2.2.2 Quantitative Content Validity: Lawshe CVR and Polit-Beck CVI

Quantitative content validity methods produce numerical indices from structured expert ratings. Established and reporting standards for content validity in PRO instruments have been articulated by the ISPOR Good Research Practices Task Force (Patrick et al., 2011). The Content Validity Ratio (CVR), proposed by Lawshe (1975), is calculated for each item as the proportion of experts rating the item as essential, adjusted for the panel size to produce a ratio bounded between minus one and positive one. Items with CVR above a panel-size-dependent critical value are retained as content-valid; items below are revised or dropped. The CVR offers a transparent, reproducible quantification of expert consensus on item essentiality.

The Content Validity Index (CVI), proposed by Polit and Beck (2006), addresses limitations of the CVR by separating item-level CVI (I-CVI) from scale-level CVI (S-CVI). The I-CVI is the proportion of experts rating an item as content-valid on a structured rubric, while the S-CVI summarises validity at the scale level via either the average I-CVI or the universal agreement proportion across items. Polit and Beck (2006) advance the conceptual case that the S-CVI calculated as universal agreement penalises any expert disagreement, while the averaged S-CVI is more forgiving and more interpretable. Their recommendation that the averaged S-CVI be reported alongside item-level I-CVI has been widely adopted.

Both methods share the limitation that the numerical output depends on the panel's composition and on the rubric used for ratings. A small panel produces a noisy CVR; a panel without pediatric rheumatology expertise produces a CVI calibrated to the wrong construct. Quantitative content validity is therefore necessary but not sufficient for adaptation work, and the framework in Chapter 3 treats it as one input alongside qualitative review and cognitive debriefing.

2.2.3 Cognitive Interviewing for Pediatric Respondents

Cognitive interviewing techniques surface the respondent's actual interpretation of items and response options, distinct from the developer's intended interpretation. Two principal techniques are used. Think-aloud interviewing asks the respondent to verbalise their thought process as they read each item and select a response, allowing the interviewer to observe comprehension and decision-making in real time. Verbal probing inserts targeted probe questions after items or sections to elicit the respondent's interpretation of specific terms, the cognitive basis for their response, and any ambiguities they noticed (Willis and Boeije, 2013). The two techniques are typically combined in adaptation work, with think-aloud running through the instrument as a whole and verbal probing focused on items flagged in the forward-translation or expert-review phases.

Pediatric respondents pose specific challenges for cognitive interviewing. Younger respondents may have limited capacity for sustained think-aloud reporting; adolescents may self-edit their reports to align with perceived expectations of the interviewer; both groups may benefit from cognitively scaffolded probe questions rather than open verbal probing. The pediatric literature on cognitive interviewing has developed techniques that address these issues, including the use of graphics and pictorial support, the use of short interview sessions broken into multiple meetings, and the use of caregiver-assisted interviewing for younger respondents (Irwin et al., 2009).

Caregiver cognitive debriefing is a parallel but distinct activity. The caregiver responds from a proxy position, reporting on what they observe the youth experiencing rather than on their own experience. Cognitive interviewing with caregivers must distinguish between the caregiver's interpretation of items as descriptions of the youth's state and the caregiver's experience of completing the instrument. Adaptation teams that conflate the two perspectives produce instruments whose caregiver form drifts from intended use, with consequences for youth-caregiver agreement analyses in subsequent psychometric work.

Table 2.2. summarises the qualitative and quantitative content-validity methods reviewed in this section, with explicit notes on pediatric applicability.

Method	Type	Output	Pediatric applicability
McKenzie criteria (item clarity, relevance, comprehensiveness, redundancy)	Qualitative expert review	Structured panel comments per item	Strong with pediatric-experienced panel; weak without
Wallace criteria (item-level evaluation; revision plan)	Qualitative expert review	Item-level revision plan	Strong when applied alongside pediatric cognitive debriefing
Lawshe Content Validity Ratio (CVR)	Quantitative expert rating	Numerical ratio per item, retain/revise decision per critical value	Applicable with sufficient panel size; insensitive to pediatric subtleties without companion qualitative review

Polit-Beck Content Validity Index (I-CVI and S-CVI)	Quantitative expert rating	Item-level (I-CVI) and scale-level (S-CVI) indices	Applicable with pediatric-experienced panel; reporting of averaged S-CVI recommended
Cognitive interviewing (think-aloud and verbal probing)	Qualitative respondent-facing	Interpretation transcripts; difficulty maps; revision recommendations	Essential for pediatric adaptation; requires pediatric-adapted techniques (scaffolding, sessions, caregiver assistance for younger respondents)

Source: synthesised from Lawshe (1975), Polit and Beck (2006), Willis and Boeije (2013), and Irwin et al. (2009).

2.3 Psychometric Evaluation Methods Relevant to Adaptation

After linguistic and content validity work, the adapted instrument requires psychometric evaluation in the target population. Section 2.3 reviews the principal psychometric methods relevant to adaptation, with attention to the constraints under which they apply in pediatric rheumatology validation studies. The standard reference on health measurement scales (Streiner et al., 2015) provides the foundational treatment of the methods reviewed in this section.

2.3.1 Internal Consistency: Cronbach's Alpha and its Limits in Small Samples

Internal consistency describes the extent to which items within a scale measure the same underlying construct (Cronbach, 1951; Tavakol and Dennick, 2011). The most widely used internal-consistency index is Cronbach's alpha, calculated from the average inter-item correlation and the number of items. Alpha values above accepted thresholds support the interpretation that the scale items contribute to a common construct; values below thresholds may indicate item misfit, multidimensionality, or scale-design issues.

Cronbach's alpha has known limitations that bear on adaptation work. Alpha is sensitive to the number of items: a scale with many items produces higher alpha values even when inter-item correlations are modest. Alpha assumes a unidimensional construct: a multidimensional scale may produce a deceptively high alpha that masks structural issues. Alpha is sample-dependent: a small validation sample produces an alpha estimate with wide confidence intervals, and a single point estimate from a small sample is not a reliable indicator of internal consistency. In pediatric rheumatology validation, where rare-disease epidemiology limits sample size, the small-sample behaviour of alpha is structurally relevant. Reporting alpha alongside the sample size and a confidence interval, rather than as a single point estimate, is the methodologically appropriate response (Terwee et al., 2007).

2.3.2 Test-Retest Reliability and Intraclass Correlation

Test-retest reliability describes the stability of instrument scores across two administrations to the same respondent under conditions in which the underlying state is presumed unchanged. The interval between administrations must be long enough to avoid memory contamination but short enough that the state is plausibly stable; intervals of one to two weeks are typical for pediatric chronic-disease PRO instruments. The intraclass correlation coefficient (ICC) is the standard test-retest index, with model selection depending on whether the design treats raters as fixed or random and on whether absolute agreement or consistency is the parameter of interest (Koo and Li, 2016).

Test-retest reliability for HRQOL instruments in pediatric rheumatology faces specific obstacles. The condition's natural history may include rapid change that violates the assumption of state stability across the interval. The youth's developmental trajectory may interact with the perception of items in ways that produce real but interval-attributable change. Caregiver proxy-reports may be more stable than youth self-reports because the caregiver perspective is less embedded in moment-to-moment state, with implications for interpretation of differential test-retest performance across forms (Eiser and Morse, 2001).

2.3.3 Construct Validity: Factor Analysis, Convergent, and Discriminant Validity

Construct validity describes the extent to which an instrument measures the theoretical construct it purports to measure (Messick, 1989; Mokkink et al., 2010). Three principal methods support construct validity evaluation. Exploratory and confirmatory factor analysis examine whether the instrument's internal structure aligns with the theoretical construct, with confirmatory analysis testing pre-specified factor structures derived from the source instrument. Convergent validity examines correlations with other measures expected to relate to the construct, with stronger correlations indicating that the instrument captures the intended construct. Discriminant validity examines correlations with measures expected not to relate to the construct, with weaker correlations indicating that the instrument captures the intended construct rather than a related but distinct one.

Factor-analytic methods require samples larger than are typically available in pediatric rheumatology validation studies. Confirmatory factor analysis with maximum likelihood estimation requires sample sizes that may not be feasible in a single-centre rare-disease study, and exploratory factor analysis is sensitive to small-sample instability. Alternative approaches include reporting tetrachoric or

polychoric correlations rather than full factor models, conducting multi-centre validation pooling across sites, and supplementing factor analysis with item-response-theory analysis where supported by the instrument's design.

2.3.4 Known-Groups Validity in Disease Populations

Known-groups validity examines whether the instrument distinguishes between groups expected to differ on the underlying construct. In pediatric rheumatology, expected differences include those between active disease and inactive disease, between disease subcategories, and between disease and healthy comparator populations. Known-groups validity is well suited to adaptation work because it can be evaluated with smaller samples than factor analysis, and it provides clinically interpretable evidence that the adapted instrument distinguishes between groups in the target population in ways consistent with the source instrument's known performance (Brunner et al., 2004).

Responsiveness, sometimes treated as a fourth construct-validity dimension, describes the instrument's ability to detect change in the underlying construct over time. In pediatric rheumatology, responsiveness matters for both routine monitoring and clinical-trial endpoint use, and the minimal important change (MIC) for the adapted instrument should be estimated where possible in the target population rather than assumed from the source instrument (Mokkink et al., 2021). Adaptation studies in small samples may not be powered to estimate MIC; in that case, reporting that the MIC has not yet been estimated and that the source-instrument MIC is being used as a placeholder is the appropriate qualification.

2.3.5 Sample-Size Considerations for Rare-Disease Pediatric Populations

Sample-size considerations for psychometric evaluation in rare-disease pediatric populations differ from sample-size considerations in adult chronic-disease validation. The COSMIN initiative has developed sample-size recommendations for psychometric studies (Mokkink et al., 2010, 2018), but the recommendations are calibrated primarily to adult populations and to single-construct evaluation. Pediatric rheumatology validation often encounters samples in the low tens, distributed across age bands, disease subcategories, and respondent forms. The realistic methodological response is a tiered evaluation strategy in which internal consistency and known-groups validity are evaluated at the full validation sample, test-retest is evaluated on the largest stable subset available, and factor analysis is either deferred to a multi-site pooled sample or supplemented with item-level analyses that are less sample-demanding.

Reporting sample size explicitly, alongside confidence intervals for the principal indices and a transparent account of what the sample permits and does not permit, is the methodologically appropriate response to the constraint. The framework in Chapter 3 codifies this approach in Module 4.

2.4 Gaps for the Pediatric Rheumatology Context

The general-purpose guidelines reviewed in Section 2.1, the content-validity methods reviewed in Section 2.2, and the psychometric evaluation methods reviewed in Section 2.3 together provide a substantial methodological foundation. Five gaps remain that justify a pediatric-rheumatology-specific framework.

2.4.1 Dual-Perspective Reporting is Not Addressed Systematically

General-purpose guidelines treat the parallel adaptation of a youth self-report and a caregiver proxy-report form, where it appears, as the application of the protocol twice rather than as a coordinated adaptation of a coupled pair of instruments. The coupling between the youth and caregiver forms, including the linguistic alignment of items intended to map onto the same construct, the cultural alignment of cognitive-debriefing outputs, and the psychometric evaluation of agreement between the forms, has no structural treatment in Beaton, Mapi, ISPOR, or WHO. The framework in Chapter 3 addresses this gap through dual-form management within Module 2 and through a decision protocol for youth-caregiver discrepancies developed in Section 3.7.

2.4.2 Pediatric Cognitive and Developmental Considerations Are Treated as a Footnote

General-purpose protocols acknowledge pediatric considerations but do not integrate them structurally. Cognitive interviewing techniques developed for adult respondents apply imperfectly to youth respondents, and the developmental trajectory across age bands interacts with item interpretation in ways that adult-population guidance does not anticipate. The framework in Chapter 3 treats pediatric cognitive considerations as a structural element of Module 3, with band-specific cognitive-interviewing protocols and caregiver-assisted interviewing for younger respondents.

2.4.3 Disease-Specific Content Domains in Juvenile Rheumatic Conditions Require Domain-Specific Equivalence Checks

The disease-specific content of HRQOL instruments in pediatric rheumatology, including pain and hurt, treatment burden, communication with the care team, and worry about the future course of disease, requires content-validity checks tuned to these specific domains. General qualitative review by a generic expert panel does not surface domain-specific subtleties in the same way that review by a panel including a pediatric rheumatologist, a physiotherapist, an occupational therapist, and a psychologist would. The framework in

Chapter 3 specifies panel composition for the pediatric rheumatology setting in Module 3.

2.4.4 Small-Sample Psychometric Strategies are Underdeveloped

COSMIN sample-size recommendations and standard psychometric texts presume sample sizes that pediatric rheumatology validation studies cannot reach. The methodologically appropriate response is a small-sample evaluation strategy that prioritises internal consistency, known-groups validity, and test-retest on stable subsets, while deferring factor analysis to multi-site pooled samples and supplementing with item-level analyses. This strategy is not codified in general guidelines, leaving adaptation teams to improvise. The framework in Chapter 3 codifies it in Module 4.

2.4.5 Implementation and Dissemination Guidance is Largely Absent

General-purpose guidelines end at psychometric reporting. The downstream work of implementing the adapted instrument in clinical care, integrating it into routine PRO collection, maintaining it across language drift, and re-validating it when the clinical context changes, is not addressed in Beaton, Mapi, ISPOR, or WHO. The framework in Chapter 3 addresses this gap through Module 5, which treats implementation and dissemination as a structural component of the adaptation lifecycle.

Table 2.3. summarises the five gaps and the framework module that addresses each.

Gap	Why it matters for pediatric rheumatology	Framework module addressing the gap
Dual-perspective reporting not addressed systematically	Youth self-report and caregiver proxy-report must be coupled during adaptation; general guidance does not coordinate the two forms	Module 2 (dual-form management); Section 3.7 (youth-caregiver decision protocol)
Pediatric cognitive and developmental considerations treated as footnote	Adult cognitive-interviewing techniques do not transfer cleanly to youth respondents	Module 3 (band-specific cognitive interviewing; caregiver assistance for younger respondents)
Disease-specific content domains require domain-specific equivalence checks	Pain and hurt, treatment burden, worry, and communication require panel expertise that generic panels do not provide	Module 3 (panel composition specified for pediatric rheumatology)
Small-sample psychometric strategies underdeveloped	Rare-disease pediatric samples cannot meet COSMIN sample sizes calibrated for adult populations	Module 4 (tiered evaluation; explicit reporting of sample-size constraints)
Implementation and dissemination guidance largely absent	General guidance ends at psychometric reporting; downstream clinical and research integration receive no structural treatment	Module 5 (clinical integration; research integration; documentation; maintenance and re-validation)

Source: framework synthesis developed across the present monograph.

2.5 Synthesis and Rationale for the Proposed Framework

Chapter 2 has reviewed the principal guidelines, content-validity methods, and psychometric evaluation methods relevant to cross-cultural adaptation of PRO measures, and has identified five categories of gap that justify a pediatric-rheumatology-specific framework. The five gaps are not failings of the general-purpose guidelines but consequences of their generality: Beaton, Mapi, ISPOR, and WHO each operationalise the general adaptation task credibly and leave domain-specific tailoring to downstream work. In pediatric rheumatology that downstream work has not been codified into a structured framework, and adaptation teams improvise the response within each project.

The framework presented in Chapter 3 is the structured response. It integrates linguistic and cultural adaptation, content validity and cognitive debriefing, and psychometric evaluation within a single workflow calibrated for pediatric rheumatology, and it adds two structural elements that the general guidelines do not provide: a decision protocol for youth-caregiver discrepancies developed in Section 3.7, and an implementation and dissemination module developed in Section 3.6. The framework does not displace the existing guidelines; it operationalises pediatric-rheumatology-specific adaptation within their general structure, providing the structural treatment that the general guidelines explicitly leave open.

CHAPTER 3. THE CROSS-CULTURAL PRO ADAPTATION FRAMEWORK FOR PEDIATRIC RHEUMATOLOGY (CC-PRO-PR): AN INTEGRATED METHODOLOGY

Chapter 3 presents the Cross-Cultural PRO Adaptation Framework for Pediatric Rheumatology (CC-PRO-PR), the methodological centrepiece of the monograph. The framework integrates pre-adaptation assessment, linguistic and cultural adaptation, content validity and cognitive debriefing, psychometric evaluation, and implementation and dissemination into a single workflow calibrated for the pediatric rheumatology setting. The chapter develops each of the five modules in operational detail, presents the

original decision protocol for youth-caregiver discrepancies in Section 3.7, and closes by positioning CC-PRO-PR against the general-purpose guidelines reviewed in Chapter 2.

3.1 Framework Concept and Design Principles

The CC-PRO-PR framework is designed around five principles drawn from the gap analysis in Chapter 2. Pediatric-developmental sensitivity treats the developmental variability across age bands as a structural rather than a residual concern, with band-specific adaptation work integrated into Modules 2 and 3. Dual-perspective integration treats the youth self-report and caregiver proxy-report as a coupled pair of instruments, with coordinated linguistic, content-validity, and psychometric work across both forms. Disease-domain specificity requires that the panel composition, the cognitive-debriefing protocol, and the psychometric strategy be tuned to the specific HRQOL domains of juvenile rheumatic conditions, including pain and hurt, treatment burden, communication, and worry about the future course of disease. Small-sample feasibility ensures that the psychometric evaluation strategy in Module 4 operates within the rare-disease constraints of pediatric rheumatology validation. Transparency and reproducibility require that every deliverable of every module be documented in a form that supports peer review, replication, and downstream use by clinicians and researchers in the target population (Wild et al., 2005; Mokkink et al., 2010, 2018).

Each principle has a structural rationale grounded in the gap analysis of Chapter 2. Pediatric-developmental sensitivity addresses the gap noted in Section 2.4.2: general guidelines treat pediatric considerations as footnotes, while CC-PRO-PR treats them as structural elements that shape every module from selection through dissemination. Dual-perspective integration addresses Section 2.4.1: the parallel forms are not duplicated applications of the protocol but coupled instruments that share linguistic, content-validity, and psychometric work. Disease-domain specificity addresses Section 2.4.3: panel composition, debriefing focus, and known-groups validity contrasts must reflect the specific HRQOL signature of juvenile rheumatic conditions. Small-sample feasibility addresses Section 2.4.4: the framework prescribes an evaluation strategy that operates under the rare-disease constraint rather than presuming sample sizes calibrated for adult chronic-disease validation (Terwee et al., 2007; Mokkink et al., 2010). Transparency and reproducibility address Section 2.4.5 and the broader expectation that adapted instruments enter international evidence streams: every deliverable is documented in a form that supports peer review and downstream use (Calvert et al., 2018).

The five modules of the framework are arranged sequentially but are not strictly linear. Module 1 (pre-adaptation assessment) feeds inputs into all later modules. Module 2 (linguistic and cultural adaptation) and Module 3 (content validity and cognitive debriefing) are tightly coupled, with cognitive-debriefing outputs feeding back into linguistic revision. Module 4 (psychometric evaluation) depends on the outputs of Modules 2 and 3 and produces evidence that feeds into Module 5 (implementation and dissemination). Module 5 is treated as a structural component of the adaptation lifecycle rather than as a downstream successor activity, in line with the rationale developed in Section 2.4.5.

Figure 3.1 represents the architecture of the framework, showing the five modules, the principal feedback loop between Modules 2 and 3, and the decision-protocol component in Section 3.7 that runs across Modules 3 and 4 to direct the methodological response to youth-caregiver discrepancies.

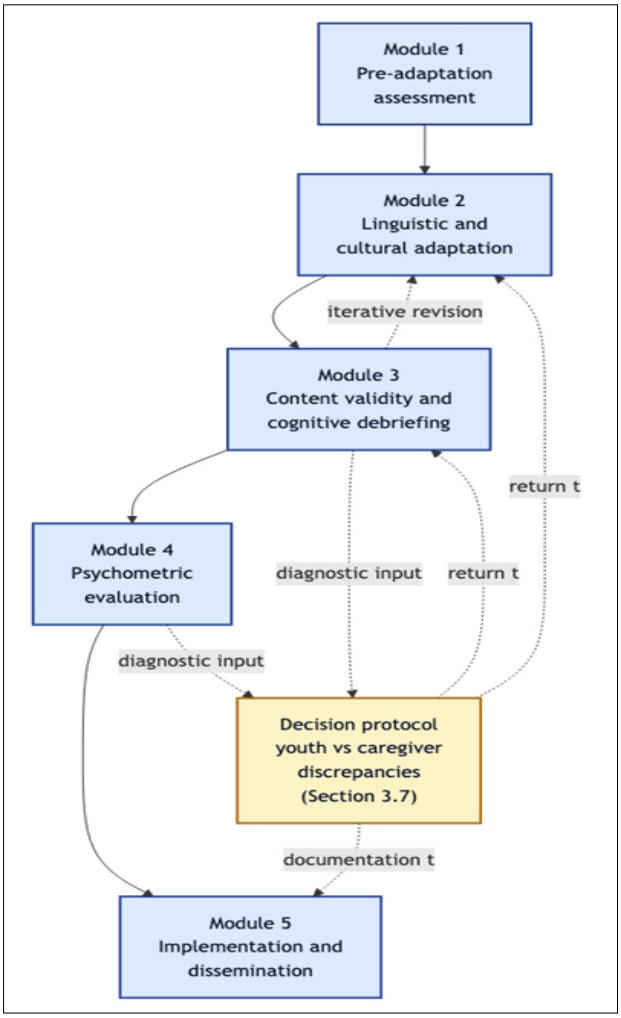

Figure 3.1. Architecture of the CC-PRO-PR framework: five integrated modules with the principal feedback loop between Modules 2 and 3 and the youth-caregiver discrepancy decision protocol (Section 3.7) spanning Modules 3 and 4. Source: original schematic developed for the present monograph.

3.2 Module 1: Pre-Adaptation Assessment

Module 1 establishes the conditions under which adaptation work can begin. The module has four components: instrument selection and rationale, copyright clearance and developer engagement, conceptual analysis for culture-sensitive items, and stakeholder mapping.

3.2.1 Instrument Selection and Rationale

Instrument selection matches the candidate instrument to the population, the age band, and the clinical purpose for which adaptation is proposed. The selection requires comparison of candidate instruments along three axes: construct fit, age-band coverage, and developer accessibility. Construct fit asks whether the instrument's domain coverage addresses the HRQOL dimensions that matter in the target population's juvenile rheumatic disease profile. Age-band coverage asks whether the instrument's forms cover the developmental range that the adapting team intends to serve. Developer accessibility asks whether the instrument's developer is available for review and harmonisation, an assumption made by all general-purpose guidelines but not always satisfied in practice. The rationale for selection is documented as the first deliverable of the framework.

3.2.2 Copyright Clearance and Developer Engagement

Copyright clearance establishes the legal basis for the adaptation and the terms under which the adapted instrument will be used. Most PRO instruments in pediatric rheumatology, including the PedsQL Rheumatology Module 3.0, the CHAQ, the JAMAR, and the PRQL, are copyrighted to the developer or to a research institute that administers rights on the developer's behalf. The adapting team obtains written permission from the developer or rights holder before translation begins. The permission agreement specifies the languages covered, the rights to use and distribute the adapted instrument, and the developer's expected role in review during adaptation. Developer engagement extends beyond the initial agreement: the developer reviews back-translations, examines cognitive-debriefing findings, and contributes to harmonisation across language versions. The framework treats developer engagement as a structural component of the adaptation lifecycle, not as an optional courtesy (Wild et al., 2005).

3.2.3 Conceptual Analysis for Culture-Sensitive Items

Conceptual analysis identifies items whose target-language equivalents may shift the construct being measured. The analysis is conducted on the source instrument before translation begins, by the adapting team in consultation with cultural informants familiar with the target population. The analysis flags items containing culture-specific metaphors, items relying on idiomatic expressions that do not transfer, items presupposing cultural framings of illness that differ in the target setting, and items whose response options imply a culturally specific gradient that may not be available in the target language. The conceptual analysis informs the briefing of forward translators in Module 2 and the briefing of the expert panel in Module 3.

3.2.4 Stakeholder Mapping

Conceptual analysis distinguishes three categories of item that warrant pre-translation attention. Items with culture-bound activities, such as references to specific games or school routines, may need substitution rather than direct translation. Items with emotional vocabulary, such as worry, hope, and frustration, are flagged because emotional vocabulary varies systematically across linguistic communities (Geisinger, 1994). Items with embedded numerical thresholds, such as questions about activity duration, are flagged because numerical reasoning at the developmental level of the youth respondent may need scaffolding in the target-language version. The conceptual analysis produces a structured map of flagged items that travels with the instrument through Modules 2 and 3.

Stakeholder mapping identifies the people and institutions whose engagement is required for the adaptation to succeed and to be used. The stakeholder map for a pediatric rheumatology adaptation typically includes the pediatric rheumatologist who provides domain expertise on the conditions covered; the orthopedic surgeon, where the population includes conditions with surgical involvement; the physiotherapist and occupational therapist who provide expertise on functional impact; the registered nurse specialising in pediatric rheumatology; the psychologist who provides expertise on cognitive and emotional dimensions; the social worker who provides expertise on family and school-system dimensions; the family-member representative who brings lived experience of the caregiver role; the methodologist who supervises the adaptation work; and the professional translators who execute the linguistic work. The stakeholder map is documented and revisited as the adaptation proceeds, with stakeholders added when their engagement becomes warranted.

Table 3.0. presents a stakeholder-mapping checklist that adapting teams may use to structure Module 1's stakeholder identification across the typical roles in a pediatric rheumatology adaptation project.

Role	Expertise contributed	When engaged
Pediatric rheumatologist	Disease-domain expertise; content validity panel; clinical interpretation	Module 1 onward; structural in Module 3
Rheumatologist (adult)	Comparative disease perspective; methodology familiarity	Module 3 panel
Physiotherapist	Functional-impact expertise; daily-activities content	Module 3 panel; Module 5 clinical integration
Occupational therapist	Functional-impact expertise; school and self-care content	Module 3 panel; Module 5 clinical integration

Registered nurse (pediatric rheumatology)	Routine-care perspective; ePRO collection workflow	Module 3 panel; Module 5 clinical integration
Psychologist	Cognitive and emotional dimensions; cognitive-interviewing supervision	Module 3 panel; Module 3 cognitive interviewing
Social worker	Family and school-system dimensions; equity considerations	Module 3 panel; Module 5 dissemination
Caregiver representative	Lived experience of the proxy-report role	Module 3 panel; Module 5 user materials
Methodologist / project coordinator	Adaptation workflow; documentation; coordination	All modules

Source: CC-PRO-PR framework, with role descriptions reflecting common pediatric rheumatology multidisciplinary team composition.

3.3 Module 2: Linguistic and Cultural Adaptation

Module 2 produces the target-language version of the instrument with linguistic and cultural equivalence to the source. The module follows the Mapi-aligned structure of forward translation, reconciliation, backward translation, and developer review, with two structural extensions for pediatric rheumatology: pediatric-specific cultural adaptation and dual-form management.

3.3.1 Forward Translation

Forward translation produces two independent target-language drafts from the source instrument. The two forward translators are native speakers of the target language with professional translation experience and English-language competence. They receive a briefing that includes the conceptual analysis from Module 1, the instrument's intended use and target population, and the developer-provided notes on item intent. The translators work independently to preserve the integrity of the two drafts as separate interpretations. The drafts are documented in a structured report that includes translator biodata, the rationale for selection, the briefing materials provided, and any communications with the project coordinator during translation (Beaton et al., 2000; Acquadro et al., 2008; Eremenco et al., 2005).

3.3.2 Reconciliation

Reconciliation produces a single working version of the target-language instrument from the two forward drafts. The two translators and the project coordinator meet to compare the two drafts item by item, identify discrepancies, discuss the source-language intent of each item, and select or synthesise the wording that best preserves conceptual equivalence in colloquial, easy-to-understand target-language text. The reconciliation produces the first version of the target-language instrument and a written reconciliation report documenting the rationale for each non-trivial wording choice. Where the translators and coordinator cannot reach agreement, the developer is consulted, and the consultation is logged.

3.3.3 Backward Translation and Developer Review

Backward translation produces an English-language version from the reconciled target-language draft, executed by a translator blind to the source instrument. The back translation is compared with the source by the project coordinator and the developer, with attention to conceptual rather than literal alignment. Discrepancies that indicate conceptual drift in the forward translation are returned to the forward team for revision. The cycle iterates until the back translation aligns conceptually with the source. The framework documents the back-translation procedure, the comparison outputs, and the resulting revisions in a structured report.

The forward and backward translation cycle is documented at item level rather than at instrument level. Each item carries a record of the two forward versions, the reconciliation rationale, the back-translated version, the developer's comments where applicable, and the final wording chosen. This item-level traceability supports both downstream peer review and the targeted revision that the youth-caregiver discrepancy protocol in Section 3.7 may invoke (Tsang et al., 2017).

3.3.4 Pediatric-Specific Cultural Adaptation

Pediatric-specific cultural adaptation tunes the target-language draft for the developmental and cultural setting in which it will be administered. Three components operate at this stage. Idiomatic adjustment replaces source-language idioms with target-language idioms that preserve meaning at the cognitive level of the intended age band; for example, an English-language idiom about hurting may need replacement with a target-language phrase that captures the same combined sense of physical pain and broader discomfort without invoking unrelated cultural associations. Age-appropriate language ensures that the cognitive load and the vocabulary level match the developmental band for which the form is designed, with separate adjustments for the youth self-report form and the caregiver proxy-report form. Caregiver linguistic register addresses the formal or informal register expected in the target culture when a caregiver describes the youth's state, which differs across linguistic communities and shapes the response options that the caregiver finds natural.

3.3.5 Dual-Form Management

Dual-form management maintains alignment between the youth self-report and the caregiver proxy-report forms during adaptation. Three structural commitments operate at this stage. Item-level alignment ensures that the items in the two forms intended to address the same construct are translated with coordinated wording, so that subsequent agreement analyses can be interpreted as measuring substantive agreement rather than artefacts of differential translation. Cognitive-debriefing alignment ensures that the cognitive interviews in Module 3 are conducted with youth and caregiver respondents whose interpretations of paired items can be compared. Reporting alignment ensures that the documentation distinguishes between findings specific to the youth self-report, findings specific to the caregiver proxy-report, and findings that bear on the alignment of the two forms. The dual-form discipline is a structural extension of the framework relative to Beaton, Mapi, ISPOR, and WHO, each of which treats parallel forms as the protocol applied twice.

3.4 Module 3: Content Validity and Cognitive Debriefing

Module 3 evaluates the content validity of the target-language instrument and verifies respondent interpretation through cognitive debriefing. The module integrates qualitative and quantitative content-validity methods with structured cognitive interviewing of both youth and caregiver respondents.

3.4.1 Expert Panel Review

The expert panel is composed to reflect the multidisciplinary character of pediatric rheumatology care and the linguistic and cultural setting of the adaptation. A typical panel includes a pediatric rheumatologist, two general rheumatologists with pediatric experience, a physiotherapist, an occupational therapist, a psychologist, a social worker, a registered nurse specialising in pediatric rheumatology, and additional experts drawn from the stakeholder map developed in Module 1. The panel reviews the target-language instrument item by item, using the qualitative criteria of McKenzie and colleagues and Wallace et al. for item clarity, relevance, comprehensiveness, redundancy, and appropriateness for the target population. The qualitative review produces a structured set of panel comments that informs revision before quantitative content validity is evaluated.

Quantitative content validity is then evaluated through Lawshe's content validity ratio and the Polit-Beck content validity index. Panel members rate each item against the criterion of essentiality (for CVR) and against a structured content-validity rubric (for I-CVI), and the project coordinator computes the indices and applies the panel-size-dependent critical values to identify items that warrant revision or removal. The framework recommends reporting the averaged S-CVI alongside item-level I-CVI and panel-size-adjusted CVR, in line with Polit and Beck (2006).

Panel size is calibrated to the indices reported. A panel of fewer than five members produces unreliable Lawshe ratios; a panel of ten or more supports stable CVR estimates but increases the logistical burden of coordination across reviews. The framework recommends a panel of eight to twelve members for typical pediatric rheumatology adaptations, with composition reflecting the disciplinary mix described above and with the language and cultural background of panel members matched to the target population where possible (Polit and Beck, 2006). When pediatric rheumatology expertise is scarce in the target setting, the framework permits remote participation from internationally available pediatric rheumatologists, with local stakeholders providing the cultural and clinical-context expertise that cannot be supplied remotely.

3.4.2 Cognitive Interviewing with Pediatric Respondents

Cognitive interviewing with pediatric respondents follows three operational rules. First, the technique is matched to the age band: think-aloud is supplemented or replaced by verbal probing for younger respondents; verbal probing is interleaved with think-aloud for older respondents. Second, the interview is structured to limit cognitive load: interviews are conducted in short sessions, with breaks scheduled between scales, and with debriefing focused on items flagged by forward translation, expert review, or pilot administration rather than on the entire instrument. Third, the interviewer is trained to recognise and probe for the youth's tendency to provide socially desirable responses, and to use neutral language that does not bias the youth toward agreement with the interviewer's framing (Irwin et al., 2009; Willis and Boeije, 2013).

Sample size for cognitive debriefing in pediatric adaptation studies is calibrated to saturation rather than to numerical power. The framework recommends five to ten youth interviews per age band and five to ten caregiver interviews, with saturation defined as the round in which no new interpretation issues are surfaced across two consecutive respondents (Sousa and Rojjanasrirat, 2011). When saturation is not reached within ten interviews, additional interviews are scheduled and the saturation criterion is documented in the adaptation report. The interviews are recorded with consent, transcribed verbatim where feasible, and coded against the items and dimensions of the instrument.

3.4.3 Caregiver Cognitive Debriefing

Caregiver cognitive debriefing addresses the proxy character of the caregiver's report. The interviewer probes for the

caregiver's interpretation of items as descriptions of the youth's state, distinguishing this interpretation from the caregiver's own experience of completing the instrument. The framework distinguishes two debriefing tracks within the caregiver interview: the construct track, which addresses what the caregiver thinks the item is asking about the youth, and the visibility track, which addresses whether the caregiver feels positioned to report on the dimension in question. The visibility track is especially relevant for caregiver proxy-reports on dimensions less directly observable, including worry and communication, where the caregiver may report a status they infer rather than a status they directly observe.

3.4.4 Iterative Revision Protocol

Module 3 operates iteratively. Findings from cognitive debriefing return to Module 2 for linguistic revision; revised wording returns to Module 3 for a second cognitive interview round; quantitative content-validity indices are recomputed after revision. The framework treats this iteration as expected rather than exceptional and budgets time and stakeholder availability for at least one full revision cycle. Iteration ends when both qualitative and quantitative content-validity criteria are met and cognitive interviews show that respondents in both forms understand the items as intended.

3.5 Module 4: Psychometric Evaluation

Module 4 evaluates the psychometric properties of the adapted instrument in a sample drawn from the target population. The module is calibrated for the small samples typical of pediatric rheumatology validation studies and integrates the youth-caregiver discrepancy decision protocol developed in Section 3.7.

3.5.1 Sampling Strategy for Small Clinical Populations

Sampling strategy is the structural decision that conditions every downstream psychometric analysis. The framework recommends three commitments. Minimum sample size is determined by the index of interest and the precision required, with explicit documentation of the trade-offs accepted under the rare-disease constraint. Stratification by age band and disease subcategory ensures that the validation sample represents the population for which the instrument is intended, rather than a convenience sample dominated by one band or subcategory. Reporting transparency requires that the sample size, the distribution across strata, and the sample-related limitations of the resulting psychometric indices be reported alongside the indices themselves (Mokkink et al., 2018).

The framework distinguishes between three sample-size regimes that pediatric rheumatology adaptation work commonly encounters. The first regime, with a sample in the low tens, supports internal consistency, known-groups validity, and limited test-retest evaluation, with confidence intervals reported and factor analysis deferred. The second regime, with a sample of fifty to one hundred, supports the first-regime indices with tighter confidence intervals and adds limited exploratory factor analysis with caveats about stability. The third regime, with a sample of more than one hundred (typically reached only through multi-site pooling), supports confirmatory factor analysis and measurement-invariance testing. The framework recommends explicit identification of the applicable regime at the start of Module 4 and documentation of which analyses the regime supports.

3.5.2 Reliability Evaluation

Reliability evaluation in Module 4 consists of internal consistency and test-retest reliability. Internal consistency is reported through Cronbach's alpha computed on the full validation sample, with confidence intervals and explicit reporting of the sample size and the number of items per scale. The framework recommends qualified interpretation of alpha estimates from samples in the low tens, with explicit acknowledgement that point estimates from such samples carry substantial uncertainty (Cronbach, 1951; Tavakol and Dennick, 2011). Test-retest reliability is evaluated on the largest stable subset available, with the interval calibrated to the disease's natural history and to the youth's developmental trajectory. ICC is reported with the model specification documented (Koo and Li, 2016). The recall period embedded in the instrument items shapes both the test-retest interval and the interpretation of stability findings, an aspect that deserves explicit attention during adaptation planning (Norquist et al., 2012).

3.5.3 Validity Evaluation

Validity evaluation includes construct validity through factor analysis where sample size supports it, convergent validity through correlations with anchor instruments measuring related constructs, and known-groups validity through comparisons across disease categories or severity levels in the target population. The framework recognises that factor-analytic methods may not be feasible in single-centre rare-disease studies and recommends multi-site pooling, item-level analyses, or deferral with explicit documentation as alternative approaches. Known-groups validity, supported by smaller samples and clinically interpretable, is treated as the primary construct-validity evidence available in many small-sample validation studies.

3.5.4 Youth-Caregiver Agreement

Youth-caregiver agreement is evaluated through intraclass correlation between paired youth and caregiver responses on items intended to measure the same construct, and through Bland-Altman analysis of the systematic and random

differences between the two respondents (Bland and Altman, 1986). The framework distinguishes between three sources of disagreement: translation artefacts that should be corrected by returning to Module 2, cultural-equivalence issues that require Module 3 revision, and genuine perspective differences that the instrument is designed to capture. The decision protocol for distinguishing these sources is developed in Section 3.7. Reporting youth-caregiver agreement in psychometric terms without diagnosing the source of disagreement produces an incomplete picture; the framework requires that the source be classified and that the methodological response be documented.

3.5.5 Reporting Standards

Module 4 reporting follows the COSMIN methodology for systematic reviews and reporting of PRO measurement properties (Mokkink et al., 2018). The reporting includes the sample size, the stratification across age band and disease, the indices computed with their confidence intervals, the model specifications for ICC and other multi-parameter indices, and a transparent account of the methodological choices made under sample-size constraints. The framework treats COSMIN-aligned reporting as a structural requirement rather than as a quality-improvement suggestion.

Measurement invariance is the methodological condition under which scores from the adapted instrument can be compared across the populations served by the source and adapted versions (Vandenberg and Lance, 2000). Strict invariance is rarely demonstrable in small-sample validation; partial invariance, in which most items show comparable parameters across populations, is the realistic target. The framework recommends reporting the invariance evidence available, the items for which invariance is not established, and the implications for cross-population comparison. Where invariance cannot be tested due to sample-size constraints, the limitation is reported.

3.6 Module 5: Implementation and Dissemination

Module 5 addresses the downstream work that turns an adapted, validated instrument into a usable resource in clinical care and research. Section 2.4.5 noted that general-purpose guidelines treat this work as out of scope; the framework treats it as integral.

3.6.1 Clinical Integration

Clinical integration includes training clinicians on instrument administration and scoring, embedding the instrument into routine PRO collection workflows in the clinic, and integrating PRO data into the electronic health record so that clinicians can consult the data at the point of care. Training materials are developed in the target language and include guidance on age-appropriate administration, on the dual-form structure and the youth-caregiver discrepancy interpretation, and on the scoring algorithm. Electronic health record integration considers the workflow in which the youth or caregiver completes the instrument, the timing of completion relative to the clinical encounter, and the format in which the PRO data are presented to the clinician (Haverman et al., 2013).

3.6.2 Research Integration

Research integration enables the adapted instrument to contribute to international evidence generation. Data-sharing agreements between the adapting team and international registries, contribution to multi-site validation studies that strengthen the psychometric evidence base, and incorporation into clinical-trial endpoint protocols all operate at this stage. The framework recommends that adaptation teams establish research-integration arrangements early, ideally in parallel with Module 1, to ensure that the validated instrument can enter international evidence streams without delay (Calvert et al., 2018).

3.6.3 Documentation

Documentation includes a validation report covering Modules 1 through 4, a user manual for clinical and research administration, and explicit guidance on scoring. The scoring guidance addresses three approaches that an adapted pediatric rheumatology HRQOL instrument may use: a Likert-type ordinal sum score per scale, a transformed 0 to 100 score that supports cross-scale comparison, and an item-response-theory-based score for instruments such as PROMIS Pediatric where IRT calibration is part of the instrument's design. The framework recommends reporting the Likert and transformed scores in tandem for classical instruments, with IRT-based scoring used where the instrument supports it and where measurement invariance across the target population can be demonstrated (Reeve et al., 2013).

3.6.4 Maintenance and Re-Validation Triggers

Maintenance and re-validation acknowledge that the adapted instrument is not a static artefact. Language drifts over time, with vocabulary, idiomatic conventions, and cognitive load patterns changing across decades; the clinical context drifts as new therapies enter practice and as expectations of HRQOL endpoints evolve; the age-band structure of the target population may extend or contract as the instrument's use matures. The framework specifies four maintenance triggers: a defined time horizon for periodic linguistic review, a clinical-context change that affects what the instrument measures, a request to extend use to a new age band beyond the originally validated range, and accumulated evidence from routine use that flags interpretation issues not surfaced in initial validation. When a trigger fires, the framework prescribes a targeted re-validation rather than a full re-adaptation, with the scope determined by the trigger.

3.7 Decision Protocol for Youth Self-Report Vs. Caregiver Proxy-Report Discrepancies

Section 3.7 develops the decision protocol for handling youth-caregiver discrepancies during adaptation. The protocol is original to the framework and addresses the gap identified in Section 2.4.1: existing guidelines treat youth-caregiver disagreement as a measurement nuisance to be reported rather than as a diagnostic signal whose interpretation directs methodological action. The protocol classifies the source of disagreement into three categories and prescribes a distinct methodological response for each.

3.7.1 Category 1: Translation Issue

A translation issue arises when the youth and caregiver respond to items intended to measure the same construct but interpret the items differently because of a translation choice that introduced asymmetry. The diagnostic signal is a pattern in which both respondents agree on related items in adjacent scales but disagree on the affected items, with cognitive-debriefing follow-up showing that the two respondents constructed different mental models of what the affected items asked. For example, the English term "hurt" in a pain item may have been translated into a target-language word that the youth reads as physical injury while the caregiver reads as broader discomfort; the youth and caregiver then respond to what reads to them as different questions, and the resulting disagreement reflects the translation choice rather than a substantive difference in their experience of the youth's state. The methodological response is return to Module 2 for targeted linguistic revision: re-translation of the affected items with attention to the divergent interpretations surfaced in debriefing, back-translation of the revised items, and re-administration to a small subset of respondents to confirm that the revision has produced symmetric interpretation. The protocol logs the original interpretation, the revised wording, and the outcome of re-administration in the adaptation report.

3.7.2 Category 2: Cultural-Equivalence Issue

A cultural-equivalence issue arises when the youth and caregiver interpret the affected items symmetrically at the linguistic level but the cultural framing of the construct differs in the target population. The diagnostic signal is a systematic pattern of disagreement in a culturally specific direction, with cognitive-debriefing follow-up showing that both respondents understand the items in the way the source instrument intends, but that the target-culture norms for reporting the construct differ from the source-culture norms. For example, a youth-caregiver disagreement on items addressing worry may reflect a target-culture norm in which youths minimise expressed worry to protect caregivers, alongside a target-culture norm in which caregivers report worry on behalf of the youth that the youth does not endorse. The methodological response is return to Module 3 for cultural-equivalence verification: a targeted expert-panel review of the affected items with attention to the cultural-norm question, followed by a decision to retain the items with a documented interpretation caveat, to revise the items with cultural framing adjustments, or to flag the items for a substantive redesign that the developer would need to authorise.

3.7.3 Category 3: Genuine Perspective Difference

A genuine perspective difference arises when the youth and caregiver interpret the affected items symmetrically at both linguistic and cultural levels but the two respondents have access to different information about the youth's state. The diagnostic signal is a pattern of disagreement on items addressing dimensions less directly observable by the caregiver, including emotional well-being and worry, alongside agreement on items addressing dimensions more directly observable, including physical function and visible symptoms. Cognitive-debriefing follow-up shows that both respondents understand the items as intended and that the disagreement reflects the respective vantage points of the two respondents. For example, the youth may report low worry about their illness because that worry is internal and they have not articulated it to the caregiver, while the caregiver reports moderate worry on the youth's behalf based on their observation of the youth's withdrawn behaviour; both responses are accurate from the respective vantage points, and the disagreement is exactly what the dual-perspective architecture of the instrument is designed to capture. The methodological response is to document the disagreement as substantive evidence rather than as a measurement defect, to retain the items without revision, and to report the disagreement pattern in the validation report as informative about the dual-perspective design of the instrument (Eiser and Morse, 2001).

3.7.4 Methodological Handling: The Decision Tree

Figure 3.2 represents the decision tree that the protocol prescribes. The tree begins with the detection of a youth-caregiver disagreement on a specific item or scale. The first decision asks whether cognitive debriefing reveals that the youth and caregiver interpret the affected items differently. If yes, the protocol classifies the issue as Category 1 (translation) and prescribes Module 2 revision. If no, the protocol proceeds to the second decision, which asks whether the disagreement follows a culturally patterned direction. If yes, the protocol classifies the issue as Category 2 (cultural equivalence) and prescribes Module 3 review. If no, the protocol proceeds to the third decision, which asks whether the disagreement aligns with the differential observability of the affected dimensions across respondents. If yes, the protocol classifies the issue as Category 3 (genuine perspective difference) and prescribes documentation without revision. If no, the protocol returns to Module 3 for additional cognitive debriefing to resolve the classification.

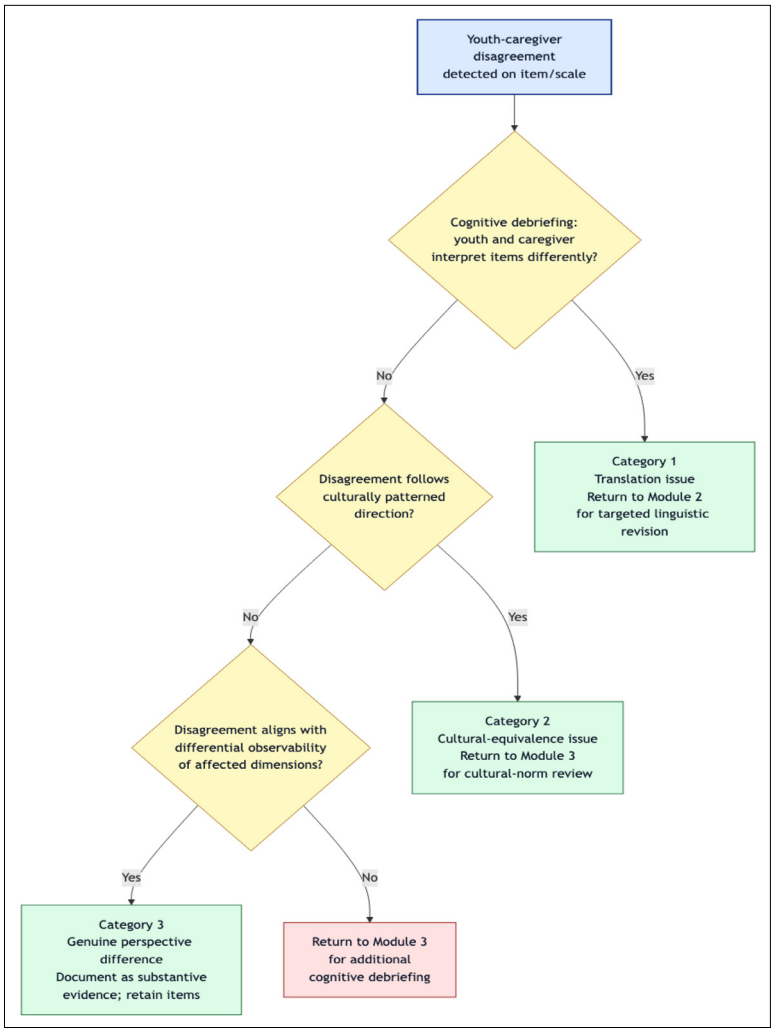

Figure 3.2. Decision protocol for classifying youth-caregiver discrepancies during cross-cultural adaptation. Source: original to the present framework.

A worked example illustrates the protocol in operation. In a hypothetical adaptation of a pediatric rheumatology HRQOL instrument into a target language, the validation sample shows youth-caregiver disagreement on items in the worry scale and on items in the daily activities scale. Cognitive debriefing reveals that youth and caregiver respondents interpret the daily activities items differently because a translation choice rendered an English colloquial phrase about a routine activity into a target-language phrase that the youth interpreted as referring to a different routine. The protocol classifies the daily activities issue as Category 1 (translation), prescribes Module 2 revision, and the revised items are re-administered with the disagreement resolving. Cognitive debriefing on the worry items shows that the youth and caregiver interpret the items consistently but that the youth systematically minimises expressed worry to avoid burdening the caregiver, while the caregiver reports the worry the youth would have on the caregiver's reading of the youth's situation. The protocol classifies the worry issue as Category 3 (genuine perspective difference), retains the items without revision, and the validation report documents the perspective difference as informative about the dual-perspective design of the instrument. The methodological response in each case is principled, documented, and replicable, in contrast to the under-specified treatment of disagreement that the general guidelines provide (Eiser and Morse, 2001).

The decision tree is applied to each item or scale on which a youth-caregiver disagreement is observed, with the classification and the prescribed response logged in the adaptation report. The protocol does not eliminate disagreement; some disagreement is expected and informative under Category 3. The protocol ensures that the methodological response to each instance of disagreement is principled rather than ad hoc, that the rationale is documented, and that the validation evidence supports a defensible reading of the dual-perspective structure of the instrument.

3.8 Distinctive Features of the Framework

Section 3.8. positions CC-PRO-PR against the general-purpose guidelines reviewed in Chapter 2. The framework does not replace those guidelines; it operationalises pediatric-rheumatology-specific adaptation within their general structure.

Table 3.1. summarises the principal differences along five dimensions.

Dimension	Beaton (2000) / Mapi / ISPOR / WHO	CC-PRO-PR	Where in the framework
Pediatric developmental considerations	Footnote treatment; pretesting on target population without band-specific guidance	Structural element; band-specific cognitive-interviewing protocols; caregiver-assisted interviewing for younger respondents	Module 2 (3.3.4); Module 3 (3.4.2)
Dual-form management	Protocol applied twice for parallel forms	Coordinated linguistic and cognitive-debriefing alignment between youth self-report and caregiver proxy-report forms	Module 2 (3.3.5); Module 3 (3.4.3)
Psychometric integration	Adaptation handed off to downstream psychometric work	Psychometric evaluation integrated as a connected module, with sampling, reliability, validity, and youth-caregiver agreement coordinated with adaptation outputs	Module 4 (3.5)
Youth-caregiver discrepancies	Reported as a measurement nuisance	Diagnosed via a structured decision protocol that classifies translation issues, cultural-equivalence issues, and genuine perspective differences, and prescribes a methodological response to each	Section 3.7
Implementation and dissemination	Out of scope	Clinical integration, research integration, documentation, and maintenance treated as a structural module of the adaptation lifecycle	Module 5 (3.6)

Source: synthesised from Beaton et al. (2000), Acquadro et al. (2008), Wild et al. (2005), and the original CC-PRO-PR framework.

Beyond the dimensions in Table 3.1, the framework differs in two cross-cutting respects. First, CC-PRO-PR is organised around the lifecycle of an adapted instrument rather than around the production of a single deliverable. The lifecycle perspective surfaces maintenance and re-validation as structural rather than residual concerns, and treats the validated instrument as a resource that enters a longer arc of use, registry contribution, and revision. Second, CC-PRO-PR makes pediatric-rheumatology-specific expertise structurally rather than incidentally relevant. Stakeholder mapping, panel composition, cognitive-debriefing scaffolding, and known-groups validity contrasts all draw on the domain expertise of pediatric rheumatologists, physiotherapists, occupational therapists, and the multidisciplinary care team, rather than treating that expertise as a supplementary input.

The positioning is collaborative rather than adversarial. CC-PRO-PR adopts the operational discipline of Mapi, the developer-engagement emphasis of ISPOR (Wild et al., 2005), the five-stage structure of Beaton et al. (2000), and the equivalence dimensions of the WHO process. The framework's contribution is the structured tailoring of those general operational elements to pediatric rheumatology, and the addition of the youth-caregiver decision protocol and the implementation-dissemination module that the general frameworks leave open. Adaptation teams that have followed Beaton, Mapi, ISPOR, or WHO in previous work will find CC-PRO-PR familiar at the operational level, with the additions targeted at the pediatric rheumatology gaps documented in Chapter 2 (Sousa and Rojjanasrirat, 2011).

Table 3.2. summarises the five modules of the framework with their principal deliverables, for ease of reference in subsequent chapters.

Module	Principal activities	Principal deliverables
Module 1: Pre-adaptation assessment	Instrument selection; copyright clearance and developer engagement; conceptual analysis; stakeholder mapping	Selection rationale; signed permission agreement; conceptual analysis report; stakeholder map
Module 2: Linguistic and cultural adaptation	Forward translation; reconciliation; backward translation and developer review; pediatric-specific cultural adaptation; dual-form management	Forward translation drafts and reports; reconciliation report; back translation and developer review log; pediatric-cultural adaptation notes; aligned youth and caregiver forms
Module 3: Content validity and cognitive debriefing	Expert panel review (qualitative and quantitative); cognitive interviewing with youth respondents; caregiver cognitive debriefing; iterative revision	Qualitative panel comments; CVR / I-CVI / S-CVI tables; cognitive interview transcripts and difficulty maps; revised target-language version

Module 4: Psychometric evaluation	Sampling strategy; reliability evaluation; validity evaluation; youth-caregiver agreement analysis; COSMIN-aligned reporting	Validation sample report; Cronbach alpha with confidence intervals; ICC for test-retest; convergent and known-groups validity; ICC and Bland-Altman for youth-caregiver agreement
Module 5: Implementation and dissemination	Clinical integration; research integration; documentation; maintenance and re-validation	Clinician training materials; EHR integration plan; research-integration agreements; validation report; user manual; scoring guidance; maintenance plan

Source: CC-PRO-PR framework as developed in this chapter.

The five modules are not equal in time investment. Module 1 typically requires two to four weeks for instrument selection, copyright clearance, conceptual analysis, and stakeholder mapping. Module 2 typically requires four to eight weeks for forward translation, reconciliation, back translation, and developer review. Module 3 typically requires four to eight weeks for expert panel review and cognitive debriefing, with additional time when iteration is required. Module 4 typically requires three to six months for recruitment and data collection in a single-centre rare-disease setting, with additional time for multi-site arrangements. Module 5 is open-ended and continues beyond the formal validation. The framework recommends that adaptation teams plan for the modules sequentially while preserving the capacity to iterate between Modules 2 and 3 as cognitive debriefing surfaces revision needs (Acquadro et al., 2008).

Module budgets vary with the instrument's scope. Short instruments such as the PRQL (Filocamo et al., 2010) carry lower per-module workload than multidimensional instruments such as JAMAR (Consolaro et al., 2011) or item-bank-calibrated instruments such as PROMIS Pediatric (Ader, 2007; DeWalt et al., 2013). The framework's modules are scoped by activity rather than by item count, and the workload scales with item count and domain coverage within each activity. Adaptation teams plan their resources accordingly.

Two cross-module dependencies warrant explicit recognition. Forward translators in Module 2 benefit from briefings that draw on the conceptual analysis from Module 1; skipping that briefing risks producing translations that overlook culturally sensitive items flagged earlier (Geisinger, 1994; Tsang et al., 2017). Cognitive interviewers in Module 3 benefit from access to the forward translation logs from Module 2; isolating Module 3 from Module 2 risks producing debriefing findings that are difficult to map back to revision decisions. The framework recommends that adapting teams maintain a shared documentation environment across modules, with module deliverables filed in a structure that supports cross-module reference.

CHAPTER 4. APPLICATION OF THE FRAMEWORK: ILLUSTRATIVE CASES

Chapter 4 demonstrates the CC-PRO-PR framework's application through four illustrative cases drawn from cross-cultural PRO adaptation work in pediatric rheumatology. Each case is presented in the same five-module structure for direct comparability across cases. The cases are not new validation studies; they are methodological illustrations of how the framework's modules operate in practice across instruments and across adaptation contexts. The author's PedsQL Rheumatology Module 3.0 Hindi adaptation serves as Case 1, treated with the same evenness as the other three cases. A cross-case synthesis in Section 4.5 identifies common methodological challenges, the value the framework adds beyond what each case did individually, and the patterns that inform refinement of the framework.

4.1 Case 1: PedsQL Rheumatology Module 3.0 Hindi Adaptation

The author conducted a Hindi adaptation of the PedsQL Rheumatology Module 3.0 Teen and Parent report for Teen forms covering the thirteen-to-eighteen age band, with content validity and internal consistency evaluated in a sample drawn from a single tertiary pediatric rheumatology centre in northern India. The adaptation followed the Mapi Institute Linguistic Validation Protocol for the linguistic stages and combined McKenzie and Wallace qualitative criteria with Lawshe's content validity ratio for content validity. Internal consistency was acceptable for both the Teen and Parent report for Teen forms (Varni et al., 2002; Beaton et al., 2000; Acquadro et al., 2008; Lawshe, 1975). The detailed validation results appear in a forthcoming publication in the Indian Journal of Rheumatology.

Case 1 is presented primarily to illustrate the framework's operational sequence from the author's direct experience. The psychometric outcomes of the Hindi adaptation are reported in a forthcoming primary publication; the present chapter focuses on the methodological decisions made at each module rather than on the validation indices. The adaptation was conducted before the framework presented in this monograph was developed. Section 4.1 maps the activities the adapting team performed onto the framework's five modules, identifies where the framework would have added structural value if it had been available, and notes the methodological lessons that the experience produced.

4.1.1 Mapping onto the Five Modules

Table 4.1. maps the activities of the Hindi adaptation onto the five modules of the CC-PRO-PR framework.

CC-PRO-PR module	How the case operationalised the module
Module 1: Pre-adaptation assessment	Instrument selected as the only disease-specific HRQOL measure with broad pediatric rheumatic disease coverage and dual-form architecture. Copyright clearance obtained from Mapi Research Institute as administrator of the PedsQL adaptation rights. Stakeholder mapping at the single-centre level included pediatric rheumatologist, rheumatologists, physiotherapist, occupational therapist, psychologist, social worker, and registered nurse with rheumatology specialty.
Module 2: Linguistic and cultural adaptation	Mapi-protocol forward translation by two professional Hindi translators; reconciliation supervised by the project coordinator; backward translation by an experienced bilingual translator under Mapi authorisation; developer consultation on item-level wording choices, including the rendering of the English "hurt" as a Hindi synonym of pain and the inclusion of culturally appropriate eating-implement examples in the daily activities scale.
Module 3: Content validity and cognitive debriefing	Expert panel of ten members reflecting the multidisciplinary care team described in Module 1. Qualitative review against McKenzie and Wallace criteria. Quantitative content validity computed via Lawshe CVR. Cognitive debriefing conducted in the patient testing stage with five youths and five caregivers using face-to-face interviews with structured probes.
Module 4: Psychometric evaluation	Internal consistency evaluated via Cronbach alpha for both the youth and caregiver forms in a validation sample. Detailed indices reserved for the forthcoming Indian Journal of Rheumatology publication; aggregate result reported as acceptable internal consistency for both forms.
Module 5: Implementation and dissemination	The adapted instrument supports clinical and research use in the Hindi-speaking pediatric rheumatology population. Forthcoming publication in the Indian Journal of Rheumatology supports dissemination. The case did not formally develop training materials for non-author clinicians; the framework would have prescribed this in Module 5.

Source: synthesis by the author from the Hindi adaptation work, mapped onto the CC-PRO-PR framework.

4.1.2 Methodological Lessons

Three methodological lessons from the case bear on the framework presented in Chapter 3. First, the multidisciplinary expert panel composition was a strength of the case work. Earlier adaptations of the PedsQL family in other languages had typically engaged smaller panels or panels lacking some of the disciplinary roles (Varni et al., 2002; Trapanotto et al., 2009); the Hindi adaptation engaged a panel spanning the pediatric rheumatology team's full disciplinary range. The framework's prescription that the panel reflect the multidisciplinary care team draws directly on the effectiveness of that composition in the case.

Second, item-level developer engagement was structurally valuable. The Hindi adaptation contacted the developer on item-level questions about the daily activities scale's eating-implement reference and on the synonymy of hurt and pain in the pain-and-hurt scale. Item-level developer contact resolved ambiguities that would have been difficult to settle by the adapting team alone. The framework's emphasis on developer engagement as a structural element of Modules 1 through 4 reflects this lesson.

Third, the absence of structural dual-form management was a gap that the framework would have closed. The Hindi adaptation translated the Teen and Parent report for Teen forms in parallel but did not formally manage their alignment as a coupled pair of instruments. Where forward translators produced different renderings of paired items, the team reconciled them through case-by-case discussion. The framework's Module 2.5 prescribes a structured dual-form management protocol that would have systematised this work and produced an explicit alignment log alongside the reconciled target-language version.

A fourth lesson concerns the cognitive-debriefing approach. The Hindi adaptation used face-to-face cognitive interviews with structured probes, following the Mapi protocol's cognitive-probes-for-respondent-debriefing technique. The technique surfaced interpretation issues but was applied with relatively short interview times per respondent. The framework's Module 3 prescribes the use of scaffolded probing, with interviews structured to manage cognitive load across multiple sessions where needed, and with explicit probing for the youth's tendency toward socially desirable responding. Both extensions would have strengthened the cognitive-debriefing depth in the case.

A fifth lesson concerns the implementation pathway. The adaptation was conducted within the constraints of an academic single-centre validation project, with downstream clinical and research integration left to subsequent work. The framework's Module 5 treats implementation and dissemination as part of the adaptation lifecycle rather than as a downstream successor activity, with training materials, EHR integration plans, and maintenance triggers structured

into the validation project's deliverables. Applying the framework prospectively would have integrated this work into the original project timeline rather than deferring it (Reeve et al., 2013).

4.1.3 Contextual Notes on the Hindi-Speaking Pediatric Population

Hindi is the lingua franca of much of northern and central India and is widely understood across the Hindi belt and adjacent regions. The pediatric rheumatic disease population served by an Indian tertiary centre typically includes youths from a broad linguistic and educational range, with Hindi serving as the common language even when other languages are spoken at home. Validating a PRO instrument in Hindi therefore extends usable HRQOL assessment to a population that would otherwise need to complete the instrument in English. The clinical and research benefits of the Hindi adaptation are substantial in absolute terms even before considering the cumulative literature gap that the absence of Hindi versions has produced. The framework's Module 5 supports the extension of these benefits by structuring the implementation work that follows validation, with explicit attention to clinician training and ePRO collection workflows (Haverman et al., 2013).

The Hindi-speaking pediatric population shares with other large linguistic groups in Asia and elsewhere a pattern in which validated PRO instruments have lagged behind clinical practice. The cumulative effect across the Hindi, Bengali, Tamil, Telugu, Marathi, Urdu, Vietnamese, Swahili, and other major languages is a measurement gap that compromises both clinical care and cumulative research evidence. The Hindi adaptation discussed as Case 1 addresses one facet of this gap. The framework supports addressing the broader gap by lowering the methodological barrier to undertaking adaptation in other languages.

4.2 Case 2: Cross-Cultural Adaptation of CHAQ

The Childhood Health Assessment Questionnaire (Singh et al., 1994) is the most widely translated and adapted instrument in pediatric rheumatology. Adaptations into multiple languages have been reported in the peer-reviewed literature, with the cumulative effort establishing CHAQ as a standard functional measure across diverse linguistic and cultural settings. The PRINTO-coordinated multi-country CHAQ adaptation programme is the largest organised single-instrument adaptation effort in pediatric rheumatology (Ruperto et al., 2001), with national-site validation studies reported across multiple languages including French (Pouchot et al., 2004), Brazilian Portuguese (Magalhães et al., 2006), and many others. Section 4.2 examines selected published CHAQ adaptations and maps their activities onto the CC-PRO-PR framework, identifying both methodological strengths and gaps that the framework's structure surfaces.

4.2.1 Mapping onto the Five Modules

CC-PRO-PR module	How the case operationalised the module
Module 1: Pre-adaptation assessment	CHAQ has a stable developer and a recognised rights-clearance pathway; copyright clearance has not been a structural obstacle. Stakeholder mapping in published adaptations has often been limited to the academic team conducting the validation, with less consistent inclusion of allied health and family-member representatives.
Module 2: Linguistic and cultural adaptation	Published CHAQ adaptations follow forward-and-backward translation patterns aligned with Beaton et al. (2000) or with WHO process variants. Cultural adaptation focuses on activities of daily living, where the cultural specificity of items, including those covering eating utensils, footwear, and toilet arrangements, varies across populations and requires explicit substitution.
Module 3: Content validity and cognitive debriefing	Published adaptations describe expert review at varying levels of detail; cognitive debriefing with pediatric respondents is reported less consistently than expert review. Where cognitive debriefing is reported, it commonly involves a small sample of youths and caregivers and focuses on activity items with cultural specificity.
Module 4: Psychometric evaluation	Internal consistency, test-retest reliability, and construct validity through correlation with disability indicators have been reported across CHAQ adaptations. Sample sizes vary across populations; published values for Cronbach alpha and ICC are within the range expected for functional assessment instruments in pediatric rheumatic disease populations.
Module 5: Implementation and dissemination	CHAQ is embedded in routine clinical and research use across many pediatric rheumatology centres. The breadth of its adoption is itself an implementation success; the cumulative literature supports cross-population comparison within methodological limits.

Source: synthesis from published CHAQ adaptation literature, mapped onto the CC-PRO-PR framework.

The CC-PRO-PR framework lens highlights two areas where CHAQ adaptations across languages have been variable. Cognitive debriefing with pediatric respondents is reported less consistently than expert review, which is symptomatic of the broader treatment of pediatric cognitive considerations as a footnote in general adaptation protocols. The framework's Module 3 prescribes cognitive debriefing as a structural component, with age-band-specific protocols

and minimum sample sizes that several published CHAQ adaptations would not meet under strict reading. The framework also surfaces the youth-caregiver discrepancy question in CHAQ adaptations: the instrument is primarily functional rather than HRQOL-focused, and disagreements between youth and caregiver on activity items have been reported but not systematically classified along the lines that Section 3.7 prescribes (Lam et al., 2004).

4.2.2 Strengths and Weaknesses Revealed by the Framework Lens

CHAQ adaptations exhibit two consistent strengths under the framework lens. First, the instrument's developer engagement pathway is mature: the rights-clearance procedure is well understood, and the developer or developer's institution has responded to adaptation queries across decades of cumulative work. Adapting teams entering CHAQ adaptation work benefit from this maturity in ways that adapting teams for newer instruments do not. Second, the functional content of CHAQ aligns the instrument with the clinical-care purpose of routine functional assessment in pediatric rheumatology outpatient settings, supporting downstream clinical integration that several other instruments require additional implementation work to achieve.

The framework lens also surfaces weaknesses. Several published CHAQ adaptations report validation in samples smaller than would support confirmatory factor analysis, with the psychometric evidence relying on internal consistency and known-groups validity rather than on construct-validity testing through factor models. The framework's Module 4 treats this regime explicitly as the small-sample evaluation pathway and supports it with transparent reporting. The published CHAQ literature would benefit from more consistent application of that regime. A second weakness is the limited engagement of allied health professionals in expert panel review for several published adaptations; the activity-of-daily-living content of CHAQ rewards review by physiotherapists and occupational therapists, and adaptations whose panels lacked these voices may have missed cultural-specificity issues in items addressing self-care and mobility. The framework's structural inclusion of these disciplines in Module 3 addresses the gap.

A third weakness, evident across the published CHAQ adaptations as a class, is the limited treatment of dual-form management. CHAQ has both a youth self-report and a caregiver proxy-report form, and adaptations have typically translated the two forms in parallel without a structured alignment protocol. Where studies have reported youth-caregiver agreement, the agreement is usually reported as a psychometric outcome rather than as a methodological diagnostic for further adaptation work. The framework's Module 2.5 and Section 3.7 together provide the structural treatment that the CHAQ adaptation literature would benefit from adopting (Eiser and Morse, 2001).

4.3 Case 3: JAMAR International Adaptation Program

The JAMAR international adaptation program, coordinated through the Paediatric Rheumatology International Trials Organisation (PRINTO), represents the largest organised cross-cultural adaptation effort in pediatric rheumatology. The program has produced validated JAMAR versions in many languages across multiple continents, with a coordinated methodology overseen by PRINTO and implementation supported by national pediatric rheumatology centres (Consolaro et al., 2011; Bovis et al., 2018). The scale of the program raises questions of methodological consistency, coordination across sites, and comparability of results that smaller single-site adaptations do not encounter.

4.3.1 Mapping onto the Five Modules

CC-PRO-PR module	How the case operationalised the module
Module 1: Pre-adaptation assessment	PRINTO provides centralised coordination of instrument selection rationale, copyright clearance, and stakeholder mapping across participating sites. National sites operate within a standard structure rather than constructing pre-adaptation assessment locally for each language.
Module 2: Linguistic and cultural adaptation	PRINTO methodology specifies forward and backward translation procedures applied uniformly across languages. Cultural adaptation occurs at the national-site level under PRINTO oversight, with harmonisation across language versions as a structural commitment of the program.
Module 3: Content validity and cognitive debriefing	Expert review occurs at the national-site level with PRINTO methodology specifying minimum criteria. Cognitive debriefing approaches vary across sites and have been more or less detailed depending on local resources and on the specific language version.
Module 4: Psychometric evaluation	PRINTO coordinates a multi-site validation strategy in which pooling across sites permits psychometric analyses that single-site samples cannot support. Confirmatory factor analysis, measurement invariance testing across languages, and known-groups validity comparisons across disease categories are feasible at the program level in ways that single-site validation cannot achieve.
Module 5: Implementation and dissemination	The JAMAR program supports clinical and research use of the instrument across the PRINTO network. Registry contribution from participating sites is a structural design element, with the adapted instruments feeding international PRO data streams.

Source: synthesis from PRINTO-coordinated JAMAR adaptation reports, mapped onto the CC-PRO-PR framework.

The JAMAR program illustrates the scaling implications of the CC-PRO-PR framework. Centralised coordination of pre-adaptation assessment, linguistic protocols, and validation strategy supports methodological consistency that local single-site adaptations cannot maintain. The pooled-sample psychometric strategy resolves the small-sample constraint described in Sections 2.3.5 and 3.5.1, supporting confirmatory factor analysis and measurement invariance testing that single-centre rare-disease validation cannot feasibly conduct (Bovis et al., 2018). The variability across sites in cognitive debriefing detail surfaces a gap that the framework would close: even within a centrally coordinated program, the structural treatment of pediatric cognitive debriefing in Module 3 would benefit from a uniform protocol applied across national sites.

4.3.2 Scaling Implications and Lessons for the Framework

The JAMAR program offers three scaling lessons. First, centralised coordination of pre-adaptation assessment in Module 1 produces methodological consistency that would be logistically impossible to achieve across independent national efforts. The framework's structural treatment of pre-adaptation assessment in Module 1 supports adoption of this coordination pattern in future multi-site programs targeting other instruments. Second, the PRINTO program's harmonisation of linguistic adaptation across language versions supports cross-language comparability that ad hoc translation does not. The framework's emphasis on dual-form alignment in Module 2 can extend to cross-language alignment in multi-site programs, with the same principles of item-level traceability and reconciliation logging applied across linguistic harmonisation as well as within-language form coupling. Third, the pooled-sample psychometric strategy enabled by multi-site coordination is methodologically transformative in pediatric rheumatology, where single-centre validation operates under sample-size constraints that limit construct-validity evidence (Consolaro et al., 2014).

The framework's Section 3.5.1 distinguishes between single-centre and multi-site sample-size regimes, with each pathway supported by the framework's psychometric evaluation strategy. JAMAR-PRINTO illustrates the multi-site pathway at scale. Future adaptation efforts for other pediatric rheumatology instruments could draw on the JAMAR-PRINTO model to combine multi-site validation with centralised methodological coordination, expanding the psychometric evidence base for instruments that single-centre work alone cannot fully validate. The framework provides the methodological scaffolding for such coordination; the JAMAR program illustrates its feasibility.

4.4 Case 4: PRQL Cross-Cultural Adaptations

The Pediatric Rheumatology Quality of Life Scale (Filocamo et al., 2010) is a short HRQOL instrument developed alongside the JAMAR family of measures and intended for ease of administration in routine clinical practice. PRQL adaptations into multiple languages have been reported, with the instrument serving as a brief HRQOL complement to the more comprehensive JAMAR. Section 4.4 examines selected published PRQL adaptations and compares them with the PedsQL Rheumatology Module Hindi adaptation reviewed as Case 1.

4.4.1 Mapping onto the Five Modules

CC-PRO-PR module	How the case operationalised the module
Module 1: Pre-adaptation assessment	PRQL is developed by a research group active in pediatric rheumatology PRO measurement (Filocamo et al., 2010), with copyright clearance accessible through standard academic channels. Stakeholder mapping in published adaptations has typically followed the JAMAR program model where the adaptation occurred within a PRINTO-coordinated effort.
Module 2: Linguistic and cultural adaptation	Forward and backward translation procedures align with Beaton et al. (2000) and Mapi-protocol patterns. The brevity of PRQL reduces the per-item burden of translation but increases the importance of each item; each translation choice has proportionally larger effect on the instrument.
Module 3: Content validity and cognitive debriefing	Expert review and cognitive debriefing have been conducted at the national-site level across language versions. The compact item set means that cognitive debriefing can cover the entire instrument in modest interview time.
Module 4: Psychometric evaluation	PRQL adaptations have reported internal consistency, test-retest reliability, and construct validity through correlation with the JAMAR and other instruments. Sample sizes have varied, with multi-centre pooling under PRINTO supporting larger-sample analyses where applicable.
Module 5: Implementation and dissemination	PRQL serves as a brief HRQOL measure in routine clinical use and supports research integration alongside JAMAR within the PRINTO network.

Source: synthesis from published PRQL adaptation literature, mapped onto the CC-PRO-PR framework.

Comparison with the PedsQL Rheumatology Module Hindi adaptation reviewed as Case 1 illustrates the trade-offs between instrument length and adaptation work. The PedsQL Rheumatology Module 3.0 has twenty-two items across five domain scales, providing dense content coverage at the cost of greater translation, content-validity, and psychometric-evaluation workload. PRQL covers HRQOL with a small item set, reducing the per-item burden but raising the stakes of each item's adaptation. The two instruments also differ in usability: the PRQL is well suited to brief screening in routine clinical visits where assessment time is constrained, while the PedsQL Rheumatology Module 3.0 is better suited to comprehensive assessment across heterogeneous pediatric rheumatic disease populations where domain breadth matters. Adapting teams selecting an instrument under the CC-PRO-PR framework's Module 1 may use the comparison to inform the selection trade-off between coverage breadth and adaptation feasibility, with both instruments offering credible pathways depending on the clinical and research purpose.

A second comparison concerns the developer-engagement pathway. PRQL was developed by a research group active in pediatric rheumatology PRO measurement (Filocamo et al., 2010) and is administered through standard academic channels, with developer engagement available through direct contact with the original team. The PedsQL Rheumatology Module's developer engagement operates through the Mapi Research Institute's administration of PedsQL adaptation rights, with the institute coordinating developer-team consultation under structured procedures. Both pathways are workable; they differ in the institutional infrastructure that surrounds the developer consultation. The framework's Module 1 does not prescribe a single pathway but does prescribe that the pathway be characterised explicitly during pre-adaptation assessment so that the adapting team can plan resources accordingly.

A third comparison concerns the relation of the instrument to the broader PRO measurement ecosystem in pediatric rheumatology. The PedsQL Rheumatology Module 3.0 is designed for co-administration with the PedsQL Generic Core Scales, supporting integration of disease-specific and generic HRQOL measurement within a single family of instruments. PRQL is designed for use alongside the JAMAR family of measures within the PRINTO ecosystem. The instrument-ecosystem fit affects implementation: the PedsQL family supports cross-condition comparison through the Generic Core Scales, while the JAMAR-PRQL combination supports rheumatology-specific assessment with deeper integration of caregiver and clinical data. The framework's Module 1 invites adapting teams to consider ecosystem fit as part of instrument selection, supporting downstream integration in Module 5.

4.5 Cross-case Synthesis

Section 4.5 identifies the methodological patterns that emerge across the four cases and the implications for the CC-PRO-PR framework's refinement.

4.5.1 Common Methodological Challenges across Cases

Four challenges recur across the cases. Cognitive debriefing with pediatric respondents is treated unevenly across the cases, with the depth of debriefing varying by site and by language version. The dual-form coupling between youth self-report and caregiver proxy-report receives ad hoc rather than structural treatment in most of the cases, with the JAMAR program offering the closest to a structural approach through its PRINTO-coordinated harmonisation. Small-sample psychometric evaluation is a structural constraint across all the cases except the PRINTO-coordinated multi-site programs, with single-centre validation studies operating under the rare-disease sample-size limit. Implementation and dissemination beyond the validation publication receive less structured attention than the upstream adaptation modules.

4.5.2 Where the Framework Adds Value

The framework adds value across the four cases in three principal ways. The structural treatment of dual-form management in Module 2 closes the alignment gap that Cases 1, 2, and 4 each addressed through case-by-case discussion rather than through a documented alignment protocol. The youth-caregiver discrepancy decision protocol in Section 3.7 converts the agreement analyses in Module 4 into actionable methodological classifications, supporting principled decisions about translation revision, cultural-equivalence review, or substantive perspective documentation. The implementation and dissemination structure in Module 5 prescribes the training materials, EHR integration plans, and maintenance triggers that the cases addressed informally or deferred beyond the validation phase.

4.5.3 Patterns Informing Framework Refinement

A second cross-cutting pattern concerns the role of the developer in the adaptation process. The cases differ in the institutional infrastructure surrounding developer engagement: Mapi's centralised administration of PedsQL adaptation rights structured the Hindi adaptation work in Case 1; the long-standing CHAQ adaptation literature reflects mature developer-engagement patterns across decades of cumulative work; PRINTO's coordination of JAMAR adaptation centralises developer engagement within a multi-site program. The framework's Module 1 prescribes that the developer-engagement pathway be characterised explicitly during pre-adaptation assessment, so that the adapting team can plan accordingly. The cross-case synthesis confirms that developer engagement is a structural rather

than incidental concern across pediatric rheumatology instrument adaptation work.

A third cross-cutting pattern concerns the relation between adaptation and downstream research integration. Case 3 (JAMAR-PRINTO) shows the strongest research-integration structure: adapted instruments feed into international PRO registries and into multi-site validation pooling. Cases 1 (PedsQL Hindi) and 4 (PRQL) operate with less structural research integration at the validation phase, though both feed into eventual research use as the literature accumulates. Case 2 (CHAQ) benefits from the cumulative research-integration effect of decades of use across diverse populations, even where individual adaptations did not structurally plan for such integration. The framework's Module 5 prescribes research integration as a structural component of the adaptation lifecycle, addressing the inconsistency that the cross-case synthesis surfaces.

The cross-case synthesis surfaces two refinement patterns. First, the contrast between single-centre and multi-site validation strategies suggests that the framework should explicitly support both paths, with the small-sample regime described in Section 3.5.1 applying to single-centre rare-disease validation and the pooled-sample regime applying to multi-site programs such as JAMAR-PRINTO. The framework already treats both regimes in Section 3.5.1; the cross-case synthesis confirms that the regime distinction is methodologically warranted and operationally usable. Second, the variability in cognitive debriefing detail across cases suggests that the framework's Module 3 would benefit from a published cognitive-debriefing protocol template that adapting teams can apply uniformly across language versions of the same instrument. The framework's narrative treatment of cognitive debriefing is operationally sufficient, but a template would support consistency across distributed adaptation efforts. Both patterns are worked into the practical implementation guidance in Chapter 5.

Table 4.5. summarises the cross-case patterns and the framework module that operationalises the response.

Cross-case pattern	Where observed	Framework response
Cognitive debriefing depth varies by site / language version	Cases 2 (CHAQ), 3 (JAMAR sites), 4 (PRQL sites)	Module 3 cognitive-debriefing protocol with age-band-specific minimum standards; template recommended in Chapter 5
Dual-form alignment treated ad hoc	Cases 1 (PedsQL Hindi), 2 (CHAQ), 4 (PRQL); Case 3 partially addressed via PRINTO harmonisation	Module 2.5 dual-form management with item-level alignment log
Single-centre rare-disease sample-size constraint	Cases 1, 2, 4	Module 4 small-sample evaluation regime (Section 3.5.1)
Limited implementation and dissemination structure	Cases 1, 2, 4; Case 3 partially addressed via PRINTO network	Module 5 implementation and dissemination structure, with specific deliverables including training materials, EHR integration plans, registry agreements, and maintenance triggers established in parallel with the adaptation work rather than deferred until after publication

Source: cross-case synthesis developed in this section.

4.5.4 Generalisability of the Cross-Case Patterns

The cross-case patterns identified in Sections 4.5.1 through 4.5.3 generalise beyond the four cases reviewed to the broader literature on pediatric rheumatology PRO adaptation. Adaptations of generic pediatric HRQOL instruments such as the Pediatric Quality of Life Inventory Generic Core Scales (Varni et al., 2003), the Child Health Questionnaire, and the European KIDSCREEN measures (Ravens-Sieberer et al., 2013) exhibit similar patterns: uneven cognitive debriefing depth, ad hoc dual-form management, single-centre sample-size constraints, and limited implementation-and-dissemination structure. The framework's modules respond to these patterns at the structural level and therefore apply beyond the rheumatology-specific instruments covered by the four cases. The pediatric rheumatology specificity of the framework lies in its panel composition, its disease-domain orientation, and its calibration for rare-disease validation; the structural elements transfer to adjacent pediatric subspecialties with adjustments to the domain-specific components, a possibility addressed in Chapter 5's discussion of future directions.

The cases reviewed in Chapter 4 illustrate the framework's application without exhausting the methodological space. Other adaptations of pediatric rheumatology PRO instruments, including adaptations into less widely studied languages, adaptations conducted under non-academic institutional settings, and adaptations of newer instruments such as PROMIS Pediatric measures into the pediatric rheumatology context, would each surface additional patterns. The four cases serve as illustrative anchors for the framework's operation, with the understanding that the framework's modules are designed to accommodate methodological variation across instruments and adaptation contexts.

Two further observations close the chapter. The first concerns the relationship between the framework and the cumulative literature on cross-cultural PRO adaptation. The framework does not propose to replace the established protocols reviewed in Chapter 2; it operationalises pediatric-rheumatology-specific adaptation within their general structure. The cases reviewed in this chapter draw their methodological core from Beaton (Beaton et al., 2000), Mapi, ISPOR (Wild et al., 2005), or WHO procedures, and the framework's contribution is the structural tailoring of those general elements to pediatric rheumatology together with the addition of the youth-caregiver decision protocol and the implementation-and-dissemination module. The second observation concerns the implementation pathway from the framework to specific adaptation projects. Chapter 5 develops a practical implementation roadmap that translates the framework's modules into project-level activities, with resource and timeline estimates for each module and explicit attention to common pitfalls that the cases reviewed in this chapter exemplify.

CHAPTER 5. IMPLEMENTATION, LIMITATIONS, AND FUTURE DIRECTIONS

Chapter 5 translates the CC-PRO-PR framework from methodological design into operational practice. The chapter develops a practical implementation roadmap with timeline and resource estimates, articulates stakeholder-specific recommendations, acknowledges the framework's limitations, identifies future directions for further development, and addresses the ethical and equity dimensions of cross-cultural PRO adaptation in pediatric rheumatology. The chapter is shorter than the methodological chapters that precede it because its purpose is operational guidance rather than further methodological development.

5.1 Practical Implementation Roadmap

Section 5.1 develops a step-by-step roadmap for adapting teams planning a new project under the CC-PRO-PR framework. The roadmap is presented as a timeline organised by module, with estimates calibrated for a single-instrument adaptation conducted at a single tertiary pediatric rheumatology centre. Multi-site programs such as JAMAR-PRINTO operate on different timelines because the coordination overhead is offset by parallel execution across sites. The broader implementation literature on PRO assessment in clinical practice (Snyder et al., 2012) identifies workflow integration, clinician training, and data presentation as the recurrent operational challenges that adapting teams must plan for in Module 5.

5.1.1 Step-by-Step Guide for New Adaptation Studies

A new adaptation study under the framework proceeds through five module-level phases. The structure aligns with the operational sequences described in the Beaton et al. (2000) guidelines, the Mapi Institute protocol (Acquadro et al., 2008), and the ISPOR good-practice principles (Wild et al., 2005), with the pediatric-rheumatology-specific tailoring developed in Chapter 3. Phase 1 (Module 1: pre-adaptation assessment) requires two to four weeks. The team selects the instrument, obtains copyright clearance from the developer or rights administrator, conducts the conceptual analysis identifying culture-sensitive items, and completes stakeholder mapping. Deliverables include the instrument selection rationale, the signed permission agreement, the conceptual analysis report, and the stakeholder map. Phase 2 (Module 2: linguistic and cultural adaptation) requires four to eight weeks. The team commissions two independent forward translations, supervises reconciliation, commissions a blind backward translation, conducts developer review, applies pediatric-specific cultural adaptation, and manages dual-form alignment. Deliverables include the translation drafts and reports, the reconciliation report, the back-translation report, and the aligned youth and caregiver forms.

Phase 3 (Module 3: content validity and cognitive debriefing) requires four to eight weeks and may iterate. The team conducts expert panel review, applies qualitative and quantitative content-validity methods, conducts cognitive debriefing with youth and caregiver respondents, and revises items based on debriefing findings. Deliverables include the panel comments, the CVR and CVI tables, the cognitive interview transcripts, and the revised target-language version. Phase 4 (Module 4: psychometric evaluation) requires three to six months for recruitment and data collection in a single-centre rare-disease setting. The team recruits the validation sample, administers the adapted instrument, conducts psychometric analyses calibrated for the sample-size regime (Mokkink et al., 2010, 2018; Terwee et al., 2007), and applies the youth-caregiver discrepancy decision protocol where disagreement is observed (Eiser and Morse, 2001). Deliverables include the validation sample report, the reliability and validity analyses (Cronbach, 1951; Koo and Li, 2016), and the agreement analyses (Bland and Altman, 1986). Phase 5 (Module 5: implementation and dissemination) begins in parallel with Module 4 and continues after validation. Single-centre adaptation studies typically operate under the small-sample regime described in Section 3.5.1; once multiple adaptation teams in different languages have completed Modules 1 through 4 and research integration arrangements are in place, data can be pooled across sites under the multi-site regime, supporting psychometric analyses that single-centre samples cannot sustain. The team develops clinician training materials, establishes EHR integration arrangements, develops research-integration agreements, publishes the validation work, and establishes the maintenance plan.

5.1.2 Resource and Timeline Estimates

Table 5.1. summarises the timeline and resource estimates for the five module-level phases. The estimates are for a single-instrument adaptation conducted at a single tertiary pediatric rheumatology centre, with the academic resources typical of such a setting. Adapting teams should treat the estimates as planning anchors rather than as binding commitments; specific projects vary with instrument scope, target-language infrastructure, developer responsiveness, and recruitment conditions.

Phase / module	Typical duration	Principal resources	Key deliverables
Phase 1: Pre-adaptation assessment	2–4 weeks	Methodologist; access to developer or rights administrator; stakeholder representatives for mapping	Selection rationale; permission agreement; conceptual analysis; stakeholder map
Phase 2: Linguistic and cultural adaptation	4–8 weeks	Two forward translators; one backward translator; developer consultation time; methodologist coordination	Forward and backward translation reports; reconciliation log; aligned youth and caregiver forms
Phase 3: Content validity and cognitive debriefing	4–8 weeks, may iterate	Expert panel (8–12 members); interview time for 10–20 youth and caregiver respondents; methodologist	Panel comments; CVR / I-CVI / S-CVI tables; cognitive interview transcripts and difficulty maps; revised version
Phase 4: Psychometric evaluation	3–6 months	Validation sample recruitment via clinical pathway; data manager; statistician; methodologist	Validation sample report; reliability and validity indices with confidence intervals; agreement analyses
Phase 5: Implementation and dissemination	Begins in parallel; ongoing	Clinician educators; EHR integration team; institutional liaison; publication and dissemination team	Clinician training materials; EHR integration plan; validation publication; user manual; maintenance plan

Source: CC-PRO-PR framework implementation roadmap as developed in this chapter.

5.1.3 Common Pitfalls and How to Avoid Them

Five pitfalls recur across adaptation projects and warrant explicit attention. The first is deferral of stakeholder mapping. Adapting teams sometimes proceed directly to forward translation under time pressure, with stakeholder identification left to the content validity phase. The consequence is that the conceptual analysis and the translation briefings are conducted without input from the disciplines whose perspectives would have flagged culture-sensitive content. The framework prescribes early stakeholder mapping precisely to prevent this pitfall.

The second pitfall is under-staffing of the expert panel. A small panel of three to five members produces noisy CVR and CVI estimates (Lawshe, 1975; Polit and Beck, 2006) and limits the disciplinary breadth that qualitative review can bring. The framework's recommendation of eight to twelve members represents the balance between methodological rigour and logistical feasibility. The third pitfall is rushed cognitive debriefing. Short interview sessions, abbreviated probing, and small respondent samples can produce debriefing reports that miss interpretation issues that would have emerged with deeper debriefing (Willis and Boeije, 2013; Irwin et al., 2009). The fourth pitfall is treating youth-caregiver disagreement as a measurement nuisance rather than as a methodological diagnostic. The decision protocol in Section 3.7 prevents this pitfall by prescribing a structured classification and response. The fifth pitfall is deferral of implementation work until after validation. Implementation work that begins after the validation publication faces institutional, training, and integration barriers that early initiation would have anticipated (Calvert et al., 2018). The framework's parallel start of Module 5 with Module 4 addresses the pitfall structurally.

5.2 Stakeholder-Specific Recommendations

Section 5.2 addresses recommendations to the principal stakeholder groups for whom the CC-PRO-PR framework offers practical guidance: clinicians who use adapted instruments in routine care; researchers who plan and conduct new adaptation studies; instrument developers whose engagement supports the work; and institutions and funding bodies that evaluate adaptation proposals.

5.2.1 For Clinicians

For clinicians using or evaluating adapted PRO instruments in pediatric rheumatology, the framework offers three operational recommendations. First, evaluate the candidate instrument against the framework's five-module criteria before adoption in routine care. An instrument whose adaptation lacks documented dual-form management, cognitive debriefing depth, or known-groups validity in the target population may produce data that the clinician should interpret with explicit caveat. Second, adopt the dual-form decision protocol from Section 3.7 in clinical practice as well as in adaptation: when youth and caregiver reports diverge on a specific scale, the protocol's three categories (translation, cultural-equivalence, perspective-difference)

provide a vocabulary for interpreting the divergence at the bedside. Third, support the implementation work described in Module 5: completion of training in the instrument's administration, use of the EHR integration if available, and feedback to the adapting team on interpretation issues that emerge in routine use all sustain the instrument's quality over its operational life. The clinicians for whom this guidance applies span pediatric rheumatologists, orthopedic surgeons whose pediatric practice intersects with juvenile rheumatic disease, physiotherapists and occupational therapists who interpret PRO data alongside functional assessment, nurses who administer instruments in clinical workflows, and the broader allied health professionals whose practice encounters HRQOL evidence.

5.2.2 For Researchers

For researchers planning new adaptation studies, the framework offers a structured methodological scaffolding that reduces the methodological risk of project failure. The framework's modules and deliverables support project planning, peer-review preparation, and replicable reporting. Researchers conducting adaptation studies under the framework should commit early to stakeholder mapping, panel composition, and implementation planning; should adopt the dual-form management discipline in Module 2; should apply the cognitive-debriefing protocols in Module 3 with age-band specificity; and should implement the small-sample psychometric strategy in Module 4 with transparent reporting of the regime's analytical limits. Researchers conducting framework-aligned studies contribute to a cumulative methodological literature that strengthens future work (Mokkink et al., 2018).

5.2.3 For Instrument Developers

For instrument developers whose work supports cross-cultural adaptation by other teams, the framework offers a clear specification of the developer's role and a basis for developer-team partnership. Developers should provide adapting teams with conceptual documentation of items, with item-level intent statements where appropriate, with response to back-translation review, and with consultation on cognitive-debriefing findings. Developers benefit from adopting framework-aligned adaptation requests because the resulting adapted instruments enter international evidence streams in forms that support cumulative research integration. The framework's emphasis on developer engagement reflects the structural importance of the developer's role in cross-cultural adaptation; the framework's design honours that role without making developer availability a precondition for adaptation work.

5.2.4 For Institutions and Funding Bodies

For institutions and funding bodies evaluating adaptation proposals, the framework offers explicit criteria. A proposal should articulate the instrument selection rationale and the developer-engagement pathway. The proposal should specify the stakeholder map and the panel composition expected for Module 3. The proposal should address the dual-form management approach and the cognitive-debriefing protocol. The proposal should specify the validation sample size and the analytical regime expected under the sample-size regime, including the indices that will be reported with confidence intervals. The proposal should commit to Module 5 deliverables alongside the validation work, including clinician training materials, EHR integration plans, and a maintenance plan. Institutions and funding bodies that adopt these criteria support the production of adaptation work that contributes to international evidence streams rather than to one-off single-site validation publications.

Table 5.2. summarises the principal stakeholder-specific recommendations for ease of reference.

Stakeholder	Principal recommendations
Clinicians	Evaluate adapted instruments against framework criteria; adopt the youth-caregiver decision protocol at the bedside; support Module 5 implementation through training and feedback
Researchers	Plan adaptation projects with structural commitments to stakeholder mapping, dual-form management, cognitive debriefing depth, and small-sample psychometric strategy; report transparently
Instrument developers	Provide item-level conceptual documentation; respond to back-translation review and cognitive-debriefing findings; partner with adapting teams on framework-aligned projects
Institutions and funding bodies	Use framework criteria to evaluate proposals; require explicit treatment of stakeholder map, panel composition, dual-form management, sample-size regime, and Module 5 deliverables

Source: CC-PRO-PR framework recommendations as developed in this chapter.

5.3 Limitations of the Framework

Section 5.3 acknowledges the limitations of the CC-PRO-PR framework. The acknowledgement is a structural commitment of the framework's transparency principle and is necessary to position the framework accurately for downstream users.

The first limitation concerns small clinical populations and rare-disease contexts. While the framework's Module 4 addresses small-sample psychometric evaluation explicitly, and while the small-sample regime supports a defensible evidence base, the regime does not produce evidence equivalent to large-sample validation. Adaptations conducted

under the regime carry confidence intervals that limit the interpretive precision of point estimates and may defer factor-analytic evidence to multi-site pooled samples. This limitation is structural to pediatric rheumatology rather than to the framework; the framework manages the limitation transparently rather than claiming to overcome it.

The second limitation concerns dependence on instrument-developer cooperation. The framework treats developer engagement as a structural component of the adaptation lifecycle, and that engagement is the standard expectation. Where the developer is unavailable, unresponsive, or imposes constraints incompatible with framework-aligned adaptation, the framework's effectiveness is reduced. In such cases, the framework recommends documentation of the constraint and adjustment of the validation claims accordingly, but the constraint cannot be removed by methodological design alone.

The third limitation concerns generalisability beyond pediatric rheumatology. The framework is calibrated for the specific HRQOL signature of juvenile rheumatic conditions, the multidisciplinary care-team structure of pediatric rheumatology, and the rare-disease epidemiology of the field. Adjacent pediatric subspecialties, including pediatric oncology, pediatric neurology, and pediatric cardiology, share some of these characteristics but differ in others. Extension of the framework to these subspecialties is feasible but requires adjustments to the disease-domain components, a possibility addressed in Section 5.4. The framework as presented in this monograph applies to pediatric rheumatology specifically and is not claimed to apply to adjacent subspecialties without modification.

A fourth limitation concerns the framework's reliance on the cumulative methodological literature on cross-cultural PRO adaptation. The framework integrates and tailors existing methods rather than introducing new psychometric techniques from first principles. Where the underlying literature contains methodological debates, including debates about the proper interpretation of Cronbach's alpha in small samples or about the applicability of CVR in panels of moderate size, the framework inherits those debates. The framework's response is to recommend transparent reporting that allows downstream users to weigh the evidence under the relevant methodological assumptions.

5.4 Future Directions

Section 5.4 identifies four directions for further development of the framework.

5.4.1 Extension to Other Pediatric Subspecialties

The framework's structural elements, including dual-perspective integration, developmental-band sensitivity, small-sample feasibility, and integrated implementation-and-dissemination, transfer to pediatric oncology, pediatric neurology, pediatric cardiology, and other subspecialties where similar adaptation gaps exist. Extension requires substitution of the disease-domain components, including panel composition tuned to the target subspecialty's multidisciplinary care team, known-groups validity contrasts calibrated to the target subspecialty's clinical categories, and cognitive-debriefing focus tuned to the target subspecialty's HRQOL domains. Researchers interested in extending the framework to other pediatric subspecialties should treat the structural elements as portable and the disease-domain components as requiring subspecialty-specific development.

5.4.2 Integration with Digital PRO Collection

Digital PRO collection, including web-based and mobile electronic PRO platforms, has transformed routine clinical use of PRO instruments in pediatric chronic disease (Haverman et al., 2013). Computerised adaptive testing under item-response theory supports efficient measurement with reduced respondent burden in instrument families such as PROMIS Pediatric. The framework as presented in this monograph applies to fixed-form instruments scored under classical test theory. Integration of ePROs and CAT into the framework requires extensions in two directions: the Module 4 psychometric evaluation must incorporate measurement-invariance testing across digital and paper administration where both are used, and the Module 5 implementation work must address the technical infrastructure for digital collection. The framework's structural elements transfer to digital PRO contexts; the specific technical extensions are the subject of ongoing methodological development.

5.4.3 Item-Response Theory and Modern Psychometric Methods

Item-response theory and modern psychometric methods, including multidimensional IRT, differential item functioning analysis, and Bayesian estimation, offer methodological extensions that the framework's Module 4 can absorb where sample size supports them. The framework as presented relies primarily on classical test theory indices because these are the methods most widely available in pediatric rheumatology adaptation work. As IRT-calibrated instruments such as PROMIS Pediatric become more widely used in rheumatology research and clinical care, the framework's Module 4 can incorporate IRT-based measurement-invariance testing as a structural component of the psychometric evaluation. The methodological pathway exists; its incorporation into the framework will follow the accumulation of adaptation experience with IRT-based instruments in pediatric rheumatology (DeWalt et al., 2015).

5.4.4 International Consortia for Shared Adaptation Infrastructure

Building international consortia for shared adaptation infrastructure can extend the JAMAR-PRINTO model to other instruments. A shared infrastructure could include centralised methodological coordination, harmonised

linguistic protocols across language versions, pooled validation samples for multi-site psychometric analysis, and shared documentation environments. The framework supports such consortia at the methodological level; their establishment depends on institutional and funding commitments that the framework cannot supply but can inform.

5.5 Ethical and Equity Considerations

Section 5.5 addresses the ethical and equity dimensions of cross-cultural PRO adaptation in pediatric rheumatology, recognising that adaptation work serves equity in PRO access across linguistic and cultural groups and that the work itself raises ethical questions about pediatric participation, consent and assent, and the cross-border ownership of data and contributions.

5.5.1 Equity of Access

Equity of access to validated PRO instruments is the central ethical motivation for cross-cultural adaptation. Linguistic and cultural groups for whom validated instruments are unavailable are systematically under-served by PRO-based clinical care and PRO-based research evidence. The framework's commitment to lowering the methodological barrier to adaptation work directly supports equity of access by making adaptation more accessible to the academic and clinical centres that serve under-represented populations. The Hindi-speaking pediatric rheumatology population referenced as Case 1 illustrates one facet of this commitment; the broader pattern across Bengali, Tamil, Telugu, Marathi, Urdu, Vietnamese, Swahili, and other major languages remains a measurement-equity priority for the field.

5.5.2 Informed Assent and Consent in Pediatric Adaptation Studies

Pediatric adaptation studies involve human participants who are minors and whose participation requires both informed consent from a caregiver or legal guardian and informed assent from the youth respondent at a developmentally appropriate level. The framework recommends that adaptation studies follow institutional review board requirements for assent and consent, develop age-appropriate assent forms in the target language alongside the instrument itself, and document the assent process explicitly in the validation report. Caregiver consent must address the dual-form structure of the study, including the caregiver's own participation as a proxy-report respondent. The framework's structural attention to dual-form management in Module 2 extends to the ethical procedures that govern dual-form participation.

5.5.3 Data Ownership and Recognition of Contributions

Cross-cultural adaptation produces data and intellectual contributions that span linguistic and institutional borders. The framework recommends explicit attention to data ownership at the time of agreement between the adapting team, the developer, and any rights administrator. Data-sharing agreements established at Module 1 should address how the adapted instrument may be used by other research teams in the target population, how validation data may contribute to international registries, and how authorship and attribution recognise the contributions of national-site teams in multi-site programs. Recognition of contributions is not a residual concern; it is a structural component of equitable research practice across borders.

Closing the chapter, the framework's ethical and equity commitments are not separate from its methodological commitments. The methodological work of adaptation is itself an instrument of equity: when adaptation is rigorous, validated instruments reach populations that PRO-based research would otherwise overlook. The framework's operational discipline supports the broader ethical commitment to equitable, patient-centred care in pediatric rheumatology (Reeve et al., 2013; Wilson and Cleary, 1995). The implementation roadmap, the stakeholder-specific recommendations, the acknowledged limitations, the future directions, and the ethical and equity considerations developed in this chapter together translate the framework from methodological design into operational practice. Chapter 6 (Conclusion) consolidates the framework's contribution and identifies the reproducible workflow as the monograph's practical legacy.

CONCLUSION

The monograph has presented the Cross-Cultural PRO Adaptation Framework for Pediatric Rheumatology (CC-PRO-PR) as an integrated methodological framework for cross-cultural adaptation and validation of patient-reported outcome measures in pediatric rheumatology. The framework responds to a documented gap in the existing methodological literature: while general-purpose cross-cultural adaptation guidelines, content-validity methods, and psychometric evaluation standards each provide credible operationalisations of the general adaptation task, none of them addresses the specific demands of pediatric rheumatology, where dual-perspective reporting, pediatric developmental variability, disease-specific content domains, and rare-disease sample-size constraints together require structural rather than residual treatment.

The framework comprises five integrated modules. Module 1 (pre-adaptation assessment) establishes the conditions for adaptation work through instrument selection, copyright clearance and developer engagement, conceptual analysis of culture-sensitive items, and stakeholder mapping for the multidisciplinary pediatric rheumatology care team. Module 2 (linguistic and cultural adaptation) extends the Mapi-aligned forward and backward translation structure with pediatric-specific cultural adaptation and structured dual-form management for the youth self-report and caregiver proxy-report. Module 3 (content validity and cognitive debriefing) integrates qualitative and quantitative content validity with cognitive interviewing protocols calibrated

for youth respondents across developmental bands and for caregivers in their proxy role. Module 4 (psychometric evaluation) prescribes a tiered evaluation strategy that operates within the rare-disease sample-size constraint of pediatric rheumatology validation and that treats youth-caregiver agreement as both a psychometric outcome and a methodological diagnostic. Module 5 (implementation and dissemination) treats clinical integration, research integration, documentation, and maintenance as structural components of the adaptation lifecycle rather than as downstream successor activities.

Two original structural elements distinguish the framework from general-purpose alternatives. The youth-caregiver discrepancy decision protocol developed in Section 3.7 converts disagreement between the two perspectives into actionable methodological classifications. Translation issues return to Module 2 for targeted linguistic revision; cultural-equivalence issues return to Module 3 for cultural-norm review; genuine perspective differences are documented as substantive evidence of the instrument's dual-perspective design. The protocol replaces the general guidelines' treatment of youth-caregiver disagreement as a measurement nuisance with a principled methodological response. The implementation and dissemination module addresses the gap noted in Section 2.4.5: general guidelines end at psychometric reporting, leaving the downstream work of clinical integration, research integration, documentation, and maintenance unstructured. The framework treats that work as integral to the adaptation lifecycle.

Chapter 4 demonstrated the framework's application through four illustrative cases: the author's own PedsQL Rheumatology Module 3.0 Hindi adaptation (Case 1), CHAQ cross-cultural adaptations (Case 2), the JAMAR international adaptation program coordinated through PRINTO (Case 3), and PRQL cross-cultural adaptations (Case 4). The cross-case synthesis in Section 4.5 identified four recurring patterns: variable cognitive-debriefing depth across sites and languages, ad hoc dual-form alignment in the absence of a structured management protocol, single-centre rare-disease sample-size constraints, and limited implementation-and-dissemination structure. The framework's modules respond to each pattern at the structural level. Chapter 5 then translated the framework into operational practice through a step-by-step implementation roadmap, stakeholder-specific recommendations for clinicians, researchers, instrument developers, and institutions, an acknowledgement of the framework's limitations, four directions for future development, and an explicit treatment of the ethical and equity dimensions of cross-cultural PRO adaptation in pediatric rheumatology.

The scientific contribution of the monograph consists of the framework itself as the first integrated methodology specifically calibrated for cross-cultural PRO adaptation in pediatric rheumatology, the original youth-caregiver discrepancy decision protocol, the small-sample psychometric evaluation strategy that accommodates rare-disease validation, and the practical implementation guidance that translates the methodology into operational practice for diverse stakeholder roles. The contribution sits alongside rather than against the general-purpose guidelines reviewed in Chapter 2: the framework adopts the operational discipline of Mapi, the developer-engagement emphasis of ISPOR, the five-stage structure of Beaton et al., and the equivalence dimensions of the WHO process, and tailors them for pediatric rheumatology with the addition of two structural elements that general frameworks leave open.

The practical value of the framework is its reproducibility. Researchers planning a new adaptation study can follow the framework module by module, document the work in a structure that supports peer review and replication, and contribute the validated instrument to international evidence streams. Clinicians can evaluate adapted instruments against framework criteria before adoption and can apply the youth-caregiver decision protocol at the bedside when reports diverge. Instrument developers can partner with adapting teams on a methodologically common basis. Institutions and funding bodies can evaluate adaptation proposals against explicit framework criteria. The cumulative outcome is a methodological resource that lowers the barrier to high-quality adaptation work for the linguistically diverse pediatric populations that PRO-based research and care have historically under-served.

The framework's structural commitments warrant a closing summary. Pediatric-developmental sensitivity treats developmental variability as a structural element of every module rather than as a residual concern. Dual-perspective integration treats the youth self-report and caregiver proxy-report as a coupled pair of instruments rather than as duplicated applications of a single protocol. Disease-domain specificity requires panel composition, debriefing focus, and known-groups validity contrasts tuned to the HRQOL signature of juvenile rheumatic conditions. Small-sample feasibility ensures that the psychometric evaluation strategy operates within the rare-disease epidemiology of the field rather than presuming sample sizes calibrated for adult populations. Transparency and reproducibility require that every module's deliverables support downstream peer review and replication. The five principles operate together across the framework's five modules and represent the design commitments that distinguish CC-PRO-PR from general-purpose alternatives.

The cases reviewed in Chapter 4 illustrate the framework's application without exhausting the methodological space. Adaptations of pediatric rheumatology PRO instruments into languages, populations, and clinical settings beyond those covered by the four cases will exhibit additional patterns that future work can explore. The framework's modules are designed to accommodate methodological variation

across instruments and adaptation contexts within pediatric rheumatology, and the structural elements are portable to adjacent pediatric subspecialties where parallel adaptation gaps exist. Extension to those subspecialties requires substitution of the disease-domain components but retains the structural scaffolding.

Cross-cultural PRO adaptation in pediatric rheumatology serves the broader purpose of equitable, patient-centred care. Validated HRQOL assessment in the youth's native language and within the youth's cultural framing supports clinical decisions that reflect the youth's experienced health, supports research evidence that generalises across the populations the research is intended to serve, and supports the cumulative international literature that informs treatment, policy, and care delivery. The framework presented in this monograph contributes to that broader purpose by making methodologically rigorous adaptation accessible to the academic and clinical teams whose populations would benefit. The framework is offered to the international pediatric rheumatology community as a methodological resource for use, refinement, and extension as the field's adaptation work continues.

The monograph closes with two observations on the trajectory of cross-cultural PRO adaptation in pediatric rheumatology. The first is that the cumulative methodological literature has reached a stage at which structural integration of existing methods is more valuable than further proliferation of individual methodological techniques. The general-purpose guidelines reviewed in Chapter 2 each provide credible operationalisation of part of the adaptation task; the field's methodological need is integration tailored to specific contexts rather than further elaboration in isolation. CC-PRO-PR responds to that need for pediatric rheumatology. The second observation is that adaptation work is itself research, in addition to being a methodological precursor to research. The adaptation process generates evidence about instrument structure, about cultural framing of HRQOL, and about the dual-perspective architecture of pediatric PRO measurement. The framework's commitment to transparent reporting of every module's outputs supports that secondary research function alongside the primary function of producing usable adapted instruments. Future adaptation projects undertaken under the framework will, by their cumulative documentation, contribute to the methodological literature on cross-cultural pediatric PRO measurement as well as to the clinical and research literature on the specific instruments and populations they address.

REFERENCES

1. Acquadro, C., Conway, K., Hareendran, A., & Aaronson, N. (2008). Literature review of methods to translate health-related quality of life questionnaires for use in multinational clinical trials. Value in Health, 11(3), 509–521. https://doi.org/10.1111/j.1524-4733.2007.00292.x
2. Ader, D. N. (2007). Developing the Patient-Reported Outcomes Measurement Information System (PROMIS). Medical Care, 45(5 Suppl 1), S1–S2. https://doi.org/10.1097/01.mlr.0000260537.45076.74
3. Beaton, D. E., Bombardier, C., Guillemin, F., & Ferraz, M. B. (2000). Guidelines for the process of cross-cultural adaptation of self-report measures. Spine, 25(24), 3186–3191. https://doi.org/10.1097/00007632-200012150-00014
4. Bland, J. M., & Altman, D. G. (1986). Statistical methods for assessing agreement between two methods of clinical measurement. The Lancet, 327(8476), 307–310. https://doi.org/10.1016/S0140-6736(86)90837-8
5. Bovis, F., Consolaro, A., Pistorio, A., Garrone, M., Scala, S., Patrone, E., et al. (2018). Cross-cultural adaptation and psychometric evaluation of the Juvenile Arthritis Multidimensional Assessment Report (JAMAR). Rheumatology International, 38(1), 5–17.
6. U.S. Department of Health and Human Services, Food and Drug Administration. (2009). Guidance for industry. Patient-reported outcome measures: Use in medical product development to support labeling claims. Federal Register, 74(35), 65132–65133. https://www.fda.gov/media/77832/download
7. Brunner, H. I., Johnson, A. L., Barron, A. C., Passo, M. H., Griffin, T. A., Graham, T. B., & Lovell, D. J. (2004). Health of children with chronic arthritis: Relationship of different measures and the quality of parent proxy reporting. Arthritis & Rheumatism, 51(5), 763–773. https://doi.org/10.1002/art.20680
8. Brunner, H. I., Klein-Gitelman, M. S., Higgins, G. C., Lapidus, S. K., Levy, D. M., Eberhard, A., Singer, N., Olson, J. C., Onel, K., Punaro, M., Schanberg, L., von Scheven, E., Ying, J., & Giannini, E. H. (2009). Development and validation of a composite disease activity score for juvenile idiopathic arthritis. Arthritis & Rheumatism, 61(9), 1233–1241. https://doi.org/10.1002/art.24516
9. Calvert, M., Kyte, D., Mercieca-Bebber, R., Slade, A., Chan, A. W., King, M. T., et al. (2018). Guidelines for inclusion of patient-reported outcomes in clinical trial protocols: The SPIRIT-PRO Extension. JAMA, 319(5), 483–494. https://doi.org/10.1001/jama.2017.21903
10. Carle, A. C., Dewitt, E. M., & Seid, M. (2011). Measures of health status and quality of life in juvenile rheumatoid arthritis. Arthritis Care & Research, 63(S11), S438–S445. https://doi.org/10.1002/acr.20541
11. Consolaro, A., Ruperto, N., Bazso, A., Pistorio, A., Magni-Manzoni, S., Filocamo, G., et al. (2011). A new approach to clinical care of juvenile idiopathic arthritis: The Juvenile Arthritis Multidimensional Assessment Report (JAMAR). The Journal of Rheumatology, 38(5), 938–953. https://doi.org/10.3899/jrheum.100930

12. Consolaro, A., Bracciolini, G., Ruperto, N., Pistorio, A., Magni-Manzoni, S., Malattia, C., et al. (2014). Remission, minimal disease activity, and acceptable symptom state in juvenile idiopathic arthritis: Defining criteria based on the Juvenile Arthritis Disease Activity Score. Arthritis & Rheumatism, 64(7), 2366–2374.

13. Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. Psychometrika, 16(3), 297–334. https://doi.org/10.1007/BF02310555

14. DeWalt, D. A., Gross, H. E., Gipson, D. S., Selewski, D. T., DeWitt, E. M., Dampier, C. D., et al. (2015). PROMIS pediatric self-report scales distinguish subgroups of children within and across six common pediatric chronic health conditions. Quality of Life Research, 24(9), 2195–2208. https://doi.org/10.1007/s11136-015-0953-3

15. Eiser, C., & Morse, R. (2001). Can parents rate their child's health-related quality of life? Results of a systematic review. Quality of Life Research, 10(4), 347–357. https://doi.org/10.1023/A:1012253723272

16. Filocamo, G., Sztajnbok, F., Cespedes-Cruz, A., Magni-Manzoni, S., Pistorio, A., Viola, S., Ruperto, N., Bandeira, M., Buoncompagni, A., Lattanzi, B., Lepore, L., Frisina, A., Falcini, F., Loy, A., Martini, A., & Ravelli, A. (2010). Development and validation of a new short and simple measure of physical function in juvenile idiopathic arthritis (PRQL). Rheumatology, 49(11), 2156–2167. https://doi.org/10.1093/rheumatology/keq065

17. Forrest, C. B., Bevans, K. B., Tucker, C., Riley, A. W., Ravens-Sieberer, U., Gardner, W., et al. (2012). Commentary: The Patient-Reported Outcome Measurement Information System (PROMIS) for children and youth: Application to pediatric psychology. Journal of Pediatric Psychology, 37(6), 614–621.

18. Geisinger, K. F. (1994). Cross-cultural normative assessment: Translation and adaptation issues influencing the normative interpretation of assessment instruments. Psychological Assessment, 6(4), 304–312. https://doi.org/10.1037/1040-3590.6.4.304

19. Guillemin, F., Bombardier, C., & Beaton, D. (1993). Cross-cultural adaptation of health-related quality of life measures: Literature review and proposed guidelines. Journal of Clinical Epidemiology, 46(12), 1417–1432. https://doi.org/10.1016/0895-4356(93)90142-N

20. Haverman, L., van Rossum, M. A. J., van Veenendaal, M., van den Berg, J. M., Dolman, K. M., Swart, J. F., Kuijpers, T. W., & Grootenhuis, M. A. (2013). Effectiveness of a web-based application to monitor health-related quality of life. Pediatrics, 131(2), e533–e543. https://doi.org/10.1542/peds.2012-0958

21. Irwin, D. E., Stucky, B., Langer, M. M., Thissen, D., Dewitt, E. M., Lai, J. S., et al. (2010). An item response theory analysis of the PROMIS Pediatric Pain Interference Scale. Quality of Life Research, 19(4), 595–607. https://doi.org/10.1007/s11136-010-9619-3

22. Irwin, D. E., Varni, J. W., Yeatts, K., & DeWalt, D. A. (2009). Cognitive interviewing methodology in the development of a pediatric item bank: A patient-reported outcomes measurement information system (PROMIS) study. Health and Quality of Life Outcomes, 7, 3. https://doi.org/10.1186/1477-7525-7-3

23. Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. Journal of Chiropractic Medicine, 15(2), 155–163. https://doi.org/10.1016/j.jcm.2016.02.012

24. Lam, C., Young, N., Marwaha, J., McLimont, M., & Feldman, B. M. (2004). Revised versions of the Childhood Health Assessment Questionnaire (CHAQ) are more sensitive and suffer less from a ceiling effect. Arthritis Care & Research, 51(6), 881–889.

25. Lawshe, C. H. (1975). A quantitative approach to content validity. Personnel Psychology, 28(4), 563–575. https://doi.org/10.1111/j.1744-6570.1975.tb01393.x

26. Manners, P. J., & Bower, C. (2002). Worldwide prevalence of juvenile arthritis: Why does it vary so much? The Journal of Rheumatology, 29(7), 1520–1530.

27. Messick, S. (1989). Validity. In R. L. Linn (Ed.), Educational Measurement (3rd ed., pp. 13–103). American Council on Education / Macmillan.

28. Mokkink, L. B., Terwee, C. B., Patrick, D. L., Alonso, J., Stratford, P. W., Knol, D. L., Bouter, L. M., & de Vet, H. C. W. (2010). The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: An international Delphi study. Quality of Life Research, 19(4), 539–549. https://doi.org/10.1007/s11136-010-9606-8

29. Mokkink, L. B., Prinsen, C. A. C., Patrick, D. L., Alonso, J., Bouter, L. M., de Vet, H. C. W., & Terwee, C. B. (2018). COSMIN methodology for systematic reviews of patient-reported outcome measures (PROMs). Quality of Life Research, 27(5), 1147–1157. https://doi.org/10.1007/s11136-018-1798-3

30. Mokkink, L. B., Terwee, C. B., & Prinsen, C. A. C. (2021). Minimal important change (MIC): A conceptual clarification and systematic review of MIC estimates of PROMIS measures. Quality of Life Research, 30(10), 2729–2754. https://doi.org/10.1007/s11136-021-02925-y

31. Petty, R. E., Southwood, T. R., Manners, P., Baum, J., Glass, D. N., Goldenberg, J., He, X., Maldonado-Cocco, J., Orozco-Alcala, J., Prieur, A. M., Suarez-Almazor, M. E., & Woo, P. (2004). International League of Associations

for Rheumatology classification of juvenile idiopathic arthritis: Second revision, Edmonton 2001. The Journal of Rheumatology, 31(2), 390–392.

32. Polit, D. F., & Beck, C. T. (2006). The content validity index: Are you sure you know what's being reported? Critique and recommendations. Research in Nursing & Health, 29(5), 489–497. https://doi.org/10.1002/nur.20147

33. Ravens-Sieberer, U., Erhart, M., Rajmil, L., Herdman, M., Auquier, P., Bruil, J., et al. (2013). The European KIDSCREEN approach to measure quality of life and well-being in children: Development, current application, and future advances. Quality of Life Research, 23(3), 791–803. https://doi.org/10.1007/s11136-013-0428-3

34. Reeve, B. B., Wyrwich, K. W., Wu, A. W., Velikova, G., Terwee, C. B., Snyder, C. F., et al. (2013). ISOQOL recommends minimum standards for patient-reported outcome measures used in patient-centered outcomes and comparative effectiveness research. Quality of Life Research, 22(8), 1889–1905. https://doi.org/10.1007/s11136-012-0344-y

35. Singh, G., Athreya, B. H., Fries, J. F., & Goldsmith, D. P. (1994). Measurement of health status in children with juvenile rheumatoid arthritis. Arthritis & Rheumatism, 37(12), 1761–1769. https://doi.org/10.1002/art.1780371209

36. Sousa, V. D., & Rojjanasrirat, W. (2011). Translation, adaptation and validation of instruments or scales for use in cross-cultural health care research: A clear and user-friendly guideline. Journal of Evaluation in Clinical Practice, 17(2), 268–274.

37. Tavakol, M., & Dennick, R. (2011). Making sense of Cronbach's alpha. International Journal of Medical Education, 2, 53–55. https://doi.org/10.5116/ijme.4dfb.8dfd

38. Terwee, C. B., Bot, S. D., de Boer, M. R., van der Windt, D. A., Knol, D. L., Dekker, J., Bouter, L. M., & de Vet, H. C. (2007). Quality criteria were proposed for measurement properties of health status questionnaires. Journal of Clinical Epidemiology, 60(1), 34–42. https://doi.org/10.1016/j.jclinepi.2006.03.012

39. Thierry, S., Fautrel, B., Lemelle, I., & Guillemin, F. (2014). Prevalence and incidence of juvenile idiopathic arthritis: A systematic review. Joint Bone Spine, 81(2), 112–117. https://doi.org/10.1016/j.jbspin.2013.09.003

40. Trapanotto, M., Giorgino, D., Zulian, F., Benini, F., & Varni, J. W. (2009). The Italian version of the PedsQL in children with rheumatic diseases. Clinical and Experimental Rheumatology, 27(2), 373–380.

41. Tsang, S., Royse, C. F., & Terkawi, A. S. (2017). Guidelines for developing, translating, and validating a questionnaire in perioperative and pain medicine. Saudi Journal of Anaesthesia, 11(Suppl 1), S80–S89.

42. Vandenberg, R. J., & Lance, C. E. (2000). A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. Organizational Research Methods, 3(1), 4–70.

43. Varni, J. W., Seid, M., Smith Knight, T., Burwinkle, T., Brown, J., & Szer, I. S. (2002). The PedsQL™ in pediatric rheumatology: Reliability, validity, and responsiveness of the Pediatric Quality of Life Inventory™ Generic Core Scales and Rheumatology Module. Arthritis & Rheumatism, 46(3), 714–725. https://doi.org/10.1002/art.10095

44. Varni, J. W., Burwinkle, T. M., Seid, M., & Skarr, D. (2003). The PedsQL 4.0 as a pediatric population health measure: Feasibility, reliability, and validity. Ambulatory Pediatrics, 3(6), 329–341. https://doi.org/10.1367/1539-4409(2003)003<0329:TPAAPP>2.0.CO;2

45. Wild, D., Grove, A., Martin, M., Eremenco, S., McElroy, S., Verjee-Lorenz, A., & Erikson, P. (2005). Principles of good practice for the translation and cultural adaptation process for patient-reported outcomes (PRO) measures: Report of the ISPOR Task Force for Translation and Cultural Adaptation. Value in Health, 8(2), 94–104. https://doi.org/10.1111/j.1524-4733.2005.04054.x

46. Willis, G. B., & Boeije, H. R. (2013). The survey response process from a cross-cultural perspective. In G. B. Willis & B. R. Hagey (Eds.), Cognitive interviewing methodology (pp. 27–44). Wiley.

47. Wilson, I. B., & Cleary, P. D. (1995). Linking clinical variables with health-related quality of life: A conceptual model of patient outcomes. JAMA, 273(1), 59–65. https://doi.org/10.1001/jama.1995.03520250075037

48. Cella, D., Yount, S., Rothrock, N., Gershon, R., Cook, K., Reeve, B., et al. (2007). The Patient-Reported Outcomes Measurement Information System (PROMIS): Progress of an NIH Roadmap cooperative group during its first two years. Medical Care, 45(5 Suppl 1), S3–S11. https://doi.org/10.1097/01.mlr.0000258615.42478.55

49. Eremenco, S. L., Cella, D., & Arnold, B. J. (2005). A comprehensive method for the translation and cross-cultural validation of health status questionnaires. Evaluation & the Health Professions, 28(2), 212–232. https://doi.org/10.1177/0163278705275342

50. Filocamo, G., Davì, S., Pistorio, A., Bertamino, M., Ruperto, N., Lattanzi, B., et al. (2010). Evaluation of 21-numbered circle and 10-centimeter horizontal line visual analog scales for physician and parent subjective ratings in juvenile idiopathic arthritis. The Journal of Rheumatology, 37(7), 1534–1541. https://doi.org/10.3899/jrheum.091474

51. Hersh, A. O., von Scheven, E., Yazdany, J., Panopalis, P., Trupin, L., Julian, L., et al. (2010). Differences in long-

term disease activity and treatment of adult patients with childhood- and adult-onset systemic lupus erythematosus. Arthritis Care & Research, 62(8), 1152–1159.

52. Hoffman, T. K., & Mahoney, J. E. (2009). Reliability and validity of the Childhood Health Assessment Questionnaire and the Juvenile Arthritis Quality of Life Questionnaire in juvenile idiopathic arthritis: A systematic review. Arthritis Care & Research, 61(11), 1554–1561.

53. Magalhães, C. S., Bandeira, M., Alfieri, A. R., Magalhães, C. M., Junior, A. F., Pedroso, R., et al. (2006). Cross-cultural adaptation and validation of the Brazilian Portuguese version of the Childhood Health Assessment Questionnaire. Revista Brasileira de Reumatologia, 46(4), 261–266.

54. Moorthy, L. N., Peterson, M. G. E., Hassett, A. L., & Lehman, T. J. A. (2010). Burden of childhood-onset arthritis. Pediatric Rheumatology Online Journal, 8, 20. https://doi.org/10.1186/1546-0096-8-20

55. Norquist, J. M., Girman, C., Fehnel, S., DeMuro-Mercon, C., & Santanello, N. (2012). Choice of recall period for patient-reported outcome (PRO) measures: Criteria for consideration. Quality of Life Research, 21(6), 1013–1020. https://doi.org/10.1007/s11136-011-0003-8

56. Patrick, D. L., Burke, L. B., Powers, J. H., Scott, J. A., Rock, E. P., Dawisha, S., et al. (2007). Patient-reported outcomes to support medical product labeling claims: FDA perspective. Value in Health, 10(s2), S125–S137. https://doi.org/10.1111/j.1524-4733.2007.00275.x

57. Patrick, D. L., Burke, L. B., Gwaltney, C. J., Leidy, N. K., Martin, M. L., Molsen, E., & Ring, L. (2011). Content validity: Establishing and reporting the evidence in newly developed patient-reported outcomes (PRO) instruments for medical product evaluation: ISPOR PRO Good Research Practices Task Force report: Part 1. Value in Health, 14(8), 967–977. https://doi.org/10.1016/j.jval.2011.06.014

58. Pouchot, J., Larbre, J. P., Lemelle, I., Sommelet, D., Grouteau, E., Marguerite, L., et al. (2004). Validation of the French version of the Childhood Health Assessment Questionnaire (CHAQ) and the Child Health Questionnaire (CHQ) in juvenile idiopathic arthritis. Joint Bone Spine, 71(4), 282–287.

59. Quirke, A. M., Hetland, M. L., Krogh, N. S., & Lund Hansen, M. (2014). Comparison of generic and disease-specific health-related quality of life measures in pediatric rheumatology. Arthritis Care & Research, 66(2), 222–230.

60. Riley, A. W., Forrest, C. B., Rebok, G. W., Starfield, B., Green, B. F., Robertson, J. A., et al. (2004). The Child Report Form of the CHIP-Child Edition: Reliability and validity. Medical Care, 42(3), 221–231.

61. Ringold, S., Wallace, C. A., Rivara, F. P. (2009). Health-related quality of life, physical function, fatigue, and disease activity in children with established polyarticular juvenile idiopathic arthritis. The Journal of Rheumatology, 36(6), 1330–1336. https://doi.org/10.3899/jrheum.080732

62. Ruperto, N., Ravelli, A., Pistorio, A., Malattia, C., Cavuto, S., Gado-West, L., et al. (2001). Cross-cultural adaptation and psychometric evaluation of the Childhood Health Assessment Questionnaire (CHAQ) and the Child Health Questionnaire (CHQ) in 32 countries: Review of the general methodology. Clinical and Experimental Rheumatology, 19(4 Suppl 23), S1–S9.

63. Shaw, K. L., Southwood, T. R., Duffy, C. M., & McDonagh, J. E. (2006). Health-related quality of life in adolescents with juvenile idiopathic arthritis. Arthritis Care & Research, 55(2), 199–207. https://doi.org/10.1002/art.21852

64. Stinson, J. N., Toomey, P. C., Stevens, B. J., Kagan, S., Duffy, C. M., Huber, A., et al. (2008). Asking the experts: Exploring the self-management needs of adolescents with arthritis. Arthritis & Rheumatism, 59(1), 65–72. https://doi.org/10.1002/art.23244

65. Snyder, C. F., Aaronson, N. K., Choucair, A. K., Elliott, T. E., Greenhalgh, J., Halyard, M. Y., et al. (2012). Implementing patient-reported outcomes assessment in clinical practice: A review of the options and considerations. Quality of Life Research, 21(8), 1305–1314. https://doi.org/10.1007/s11136-011-0054-x

66. Streiner, D. L., Norman, G. R., & Cairney, J. (2015). Health Measurement Scales: A Practical Guide to Their Development and Use (5th ed.). Oxford University Press.