



# Patient-Reported Outcomes for Pediatric Adherence and Self-Management: A Systematic Review

Jill M. Plevinsky ,<sup>1</sup> PhD, Ana M. Gutierrez-Colina,<sup>1</sup> PhD, Julia K. Carmody ,<sup>1</sup> PhD, Kevin A. Hommel,<sup>1,2</sup> PhD, Lori E. Crosby,<sup>1,2</sup> PsyD, Meghan E. McGrady,<sup>1,2</sup> PhD, Ahna L.H. Pai,<sup>1,2</sup> PhD, Rachelle R. Ramsey,<sup>1,2</sup> PhD, and Avani C. Modi,<sup>1,2</sup> PhD

<sup>1</sup>Center for Adherence and Self-Management, Division of Behavioral Medicine and Clinical Psychology, Cincinnati Children's Hospital Medical Center and <sup>2</sup>Department of Pediatrics, University of Cincinnati College of Medicine

All correspondence concerning this article should be addressed to: Avani C. Modi, PhD, Center for Adherence and Self-Management, Division of Behavioral Medicine and Clinical Psychology, Cincinnati Children's Hospital Medical Center, 3333 Burnet Avenue (MLC 7039-Adherence Center), Cincinnati, OH 45229, USA. E-mail: avani.modi@cchmc.org

Received April 30, 2019; revisions received August 23, 2019; accepted November 15, 2019

## Abstract

**Objective** Treatment adherence is approximately 50% across pediatric conditions. Patient-reported outcomes (PROs) are the most common method of measuring adherence and self-management across research and clinical contexts. The aim of this systematic review is to evaluate adherence and self-management PROs, including measures of adherence behaviors, adherence barriers, disease management skills, and treatment responsibility. **Methods** Following PRISMA guidelines for systematic reviews, literature searches were performed. Measures meeting inclusion/exclusion criteria were evaluated using Hunsley and Mash's (2018) criteria for evidence-based assessment across several domains (e.g., internal consistency, interrater reliability, test-retest reliability, content validity, construct validity, validity generalization, treatment sensitivity, and clinical utility). Rating categories were adapted for the present study to include the original categories of *adequate*, *good*, and *excellent*, as well as an additional category of *below adequate*. **Results** After screening 172 articles, 50 PROs across a variety of pediatric conditions were reviewed and evaluated. Most measures demonstrated at least *adequate* content validity ( $n = 44$ ), internal consistency ( $n = 34$ ), and validity generalization ( $n = 45$ ). Findings were mixed regarding interrater reliability, test-retest reliability, and treatment sensitivity. Less than half of the measures ( $n = 22$ ) exhibited *adequate*, *good*, or *excellent* construct validity. **Conclusions** Although use of adherence and self-management PROs is widespread across several pediatric conditions, few PROs achieved *good* or *excellent* ratings based on rigorous psychometric standards. Validation and replication studies with larger, more diverse samples are needed. Future research should consider the use of emerging technologies to enhance the feasibility of broad implementation.

**Key words:** adherence barriers; allocation of treatment responsibility; evidence-based assessment; self-management skills; validation.

## Introduction

Over 26% of children and adolescents have chronic conditions in the United States (Van Cleave, Gortmaker, & Perrin, 2010). Advances in medical treatments have extended the life of many children with chronic diseases. However, these medical treatments often require that children and families engage in health behaviors and related processes to manage their condition, termed *self-management*. Behaviors required for self-management may include taking medications, obtaining prescriptions, making clinic appointments, and making lifestyle changes like diet and exercise. According to the Pediatric Self-Management Model (Modi, Pai, et al., 2012), self-management behaviors occur within the context of individual, family, community, and health care systems domains, and thus, can be influenced by factors in any of these domains.

Accounting for the interplay between these factors, adherence refers to the degree to which self-management behaviors coincide with medical or health advice. For example, child understanding of the medication regimen (individual domain), appropriate parental responsibility for disease management (family domain), peer support (community domain), and adequate patient-provider communication (health care system domain) can facilitate self-management behaviors that result in optimal treatment adherence (Modi, Pai, et al., 2012). Conversely, child behavioral difficulties resulting in medication refusal (individual domain), poor parental monitoring (family domain), social stigma (community domain), and barriers to health care access (health care system domain) may negatively impact a patient or family's ability to engage in self-management behaviors, resulting in non-adherence. As a result of these and other factors, approximately 50% of patients and their caregivers have difficulty following treatment recommendations (Rapoff, 2010).

Suboptimal treatment adherence can adversely impact outcomes at all levels, including the patient, family, school, hospital, or health care systems levels. For example, suboptimal adherence is associated with a 2.5 times greater risk of relapse in youth with leukemia (Bhatia et al., 2012), a 3.24 times greater likelihood to fail to achieve  $\geq 1$  year of seizure freedom in pediatric epilepsy (Modi, Rausch, & Glauser, 2014), decreased glycemic control in patients with diabetes (Hood, Peterson, Rohan, & Drotar, 2009), and a greater risk of rejection in both kidney (Pizzo et al., 2016) and liver transplant populations (Annunziato et al., 2018). Caregivers are also impacted; parents of youth with suboptimal treatment adherence report greater emotional distress (Fredericks, Lopez, Magee, Shieck, & Opiari-Arrigan, 2007) and poorer quality of life (Ducharme et al., 2011). Treatment nonadherence can

also lead to unnecessary treatment changes (e.g., dose-escalation) in clinical practice (Goodhand et al., 2013; Modi, Wu, Guilfoyle, & Glauser, 2012) as well as greater health care utilization (e.g., hospitalizations) and increased health care costs/charges across pediatric populations (Hommel et al., 2017).

In order to both identify patients who are at-risk for suboptimal adherence and self-management behaviors and evaluate the efficacy of interventions to improve adherence and self-management (Graves, Roberts, Rapoff, & Boyer, 2010; Kahana, Drotar, & Frazier, 2008; Pai & McGrady, 2014; Wu & Pai, 2014), robust assessment strategies are necessary. Multiple adherence and self-management measures have been used, including electronic monitoring, pharmacy refill, diary methods, and *patient-reported outcomes* (PROs; i.e., health outcomes directly reported by the patient who experiences it; McGrady et al., 2018; Quittner, Modi, Lemanek, Ievers-Landis, & Rapoff, 2008). While electronic monitoring provides an objective method of assessing adherence behavior and is often considered the gold standard measure of adherence (Landier et al., 2017), the choice of a specific adherence or self-management measurement approach is dependent on the aim of the study or clinical initiative, resources, and patient population. PROs offer multiple advantages. For example, PROs are low cost, easily and quickly administered in busy clinical settings, and can be adapted to address literacy and cultural factors (Stirratt et al., 2015). In addition, PROs can assess disease-specific aspects of the treatment regimen not captured by more objective measures (e.g., avoidance of seizure or asthma triggers, fluid intake). PROs also facilitate patient involvement in treatment decisions, as patients often report that PRO data gives them new insights about their adherence and symptoms (Hilliard, Ramey, Rohan, Drotar, & Cortina, 2011). Since the most recent review of evidence-based assessments of pediatric adherence was published over a decade ago (Quittner et al., 2008), adherence and self-management PROs have gained significant attention with a growing number of newly published measures. As a result, an updated systematic review using contemporary criteria to summarize and evaluate current PROs for adherence and self-management is needed.

Hunsley and Mash have developed criteria to rigorously evaluate evidence-based assessment tools, including PROs (Hunsley & Mash, 2018; Mash & Hunsley, 2005). Specifically, they provide a framework to evaluate measures based on a variety of factors, including internal consistency, interrater reliability, test-retest reliability, content validity, construct validity, validity generalization, treatment sensitivity, and clinical utility. Hunsley and Mash (2018) make the important clarification that these

psychometric properties are “properties of an instrument when used for a specific purpose with a specific sample”. Therefore, these criteria do not define what constitutes an evidence-based assessment but rather were designed to help researchers and clinicians determine the relative quality of a measure on a given criterion depending on a measure’s specific intended use (Hunsley & Mash, 2018). The overarching goal of this systematic review was to utilize an adapted version of Hunsley and Mash’s criteria to evaluate adherence and self-management PROs used in youth with pediatric chronic illness and their families. For purposes of this review, PROs relevant to adherence and self-management are conceptualized as those that assess adherence behaviors directly or constructs theoretically and practically associated with self-management, including adherence barriers, disease management skills (e.g., ability to accurately use treatment), and treatment responsibility (e.g., parent and child responsibilities in disease management tasks).

## Methods

### Search Strategy

A systematic review of the literature was completed in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) reference standards (see Table I and Figure 1). The literature search was conducted using the following databases: PubMed, CINAHL, HaPI, and APA PsycNET. An initial search was conducted in January 2019 with the following search terms/keywords related to adherence (“adherence,” “adherence barriers,” “disease management skills,” and “treatment responsibility”); “surveys and questionnaires,” “psychological tests”; and medical diagnoses and specialty areas (“asthma,” “cancer,” “cystic fibrosis,” “dermatology,” “epilepsy,” “HIV,” “inflammatory bowel diseases,” “sickle cell disease,” “spina bifida,” “transplant,” “type 1 diabetes,” “rheumatic diseases”). Although the search specified pediatric measures, it was not limited by age to ensure that measures used in a wide age range or multiple age groups were not missed. Prior reference lists from identified studies (Quittner et al., 2008) were also reviewed.

### Inclusion/Exclusion Criteria and PRO Selection

Direct measures of adherence behaviors, as well as measures of self-management related constructs (defined for the purposes of this study as adherence barriers, disease management skills, and/or treatment responsibility), were considered for inclusion, in an a priori fashion, if they were published in peer-reviewed journals. Only PRO development articles published in English were evaluated. Unpublished data, abstracts, and dissertations were excluded. To be included in the

systematic review, the original PRO study had to meet the following criteria: (a) include development and preliminary validation data on a youth or caregiver PROs, (b) focus on constructs of interest (e.g., adherence, self-management, and treatment responsibility); and (c) be validated in a pediatric population. Measures developed and utilized across a variety of health care settings with a wide range of medical populations were included. Three of the authors (J. M. Plevinsky, A. M. Gutierrez-Colina, J. K. Carmody) conducted database searches and manually removed duplicates. Titles and abstracts of manuscripts obtained via the original database searches were scanned for inclusion based on the criteria detailed above. Full-text articles were obtained for manuscripts meeting inclusion criteria and if inclusion could not be determined from the title and abstract alone.

### Obtaining Psychometric Data for Selected Measures

Once the final list of measures was determined, independent searches by all coauthors were conducted for studies citing each measure via PubMed and Google Scholar, thus obtaining further validation data. These searches identified a total of 200 articles reporting additional psychometric data for several of the measures included (see Supplementary Additional References for a complete list of the full-text articles reviewed for further validation).

### Data Extraction

For each study meeting inclusion criteria, data describing the study characteristics, and relevance to evidence-based assessment were extracted by one of the authors. Specifically, original reference details, the disease/medical specialty for which the measure was developed, construct(s) assessed by the measure (e.g., adherence behaviors, adherence barriers, disease management skills, and treatment responsibility), number of items/respondent, internal consistency, interrater reliability, test-retest reliability, content validity, construct validity, validity generalization, treatment sensitivity, and clinical utility were collected (see Supplementary Table 1).

### Quality Assessment and Data Synthesis

Measures were first classified by the disease population and then were subclassified by domains of self-management assessed: adherence behaviors, adherence barriers, disease management skills, or treatment responsibility. Measures validated in multiple chronic illness populations were grouped with general measures for ease of interpretation. The authors then utilized Hunsley and Mash’s (2018) criteria for evidence-based assessment to evaluate the validity and psychometric properties of the measures selected for inclusion in the review. These criteria provide a framework with guidelines

**Table I.** Search Strategy

PubMed	(((((“Surveys and Questionnaires”[Mesh] OR “Psychological Tests”[Mesh]) AND ((asthma[tiab] OR epilepsy[tiab] OR transplant[tiab] OR HIV[tiab] OR sickle cell[tiab] OR Cancer[tiab] OR neoplasm[tiab] OR Dermatology[tiab] OR “Skin diseases”[tiab] OR “type 1 diabetes”[tiab] OR “inflammatory bowel diseases”[tiab] OR “cystic fibrosis”[tiab] OR “spina bifida”[tiab] OR “Spinal Dysraphism”[tiab] OR “rheumatic diseases”[tiab]) OR (“Diabetes Mellitus, Type 1”[Mesh] OR “Asthma”[Mesh] OR “Epilepsy”[Mesh] OR “Transplants”[Mesh] OR “Transplantation”[Mesh] OR “HIV”[Mesh] OR “Anemia, Sickle Cell”[Mesh] OR “Neoplasms”[Mesh] OR “Skin Diseases”[Mesh] OR “Inflammatory Bowel Diseases”[Mesh] OR “Cystic Fibrosis”[Mesh] OR “Spinal Dysraphism”[Mesh] OR “Rheumatic Diseases”[Mesh]))) AND ((“Guideline Adherence”[Mesh] OR “Self-Management”[Mesh]) AND (adherence[tiab] OR “self management”[tiab]))) AND ((“pediatrics”[MeSH Terms] OR “pediatrics”[All Fields] OR “pediatric”[All Fields]) OR (“infant”[MeSH Terms] OR “infant”[All Fields]) OR (“child”[MeSH Terms] OR “child”[All Fields]) OR (“adolescent”[MeSH Terms] OR “adolescent”[All Fields])) AND english[filter] AND (“2008/01/01”[PDAT] : “3000/12/31”[PDAT])
CINAHL	((MH “Asthma+”) OR (MH “Epilepsy+”) OR (MH “Transplantation+”) OR (MH “Human Immunodeficiency Virus+”) OR (MH “Anemia, Sickle Cell+”) OR (MH “Neoplasms+”) OR (MH “Skin Diseases+”) OR (MH “Diabetes Mellitus, Type 1+”) OR (MH “Inflammatory Bowel Diseases+”) OR (MH “Cystic Fibrosis”) OR (MH “Spina Bifida”) OR (MH “Rheumatic Diseases+”) ) AND ((MH “Patient Compliance+”) OR (MH “Medication Compliance”) OR (MH “Self Care+”) OR “self management” ) AND (MM “Questionnaires+”) OR (MM “Psychological Tests+”) OR (MM “Research Instruments+”) Limiters - Published Date: 20080101-; English Language; Age Groups: All Child
HaPI	(asthma OR epilepsy OR transplant OR HIV OR sickle cell OR Cancer OR neoplasm OR Dermatology OR “Skin diseases” OR “type 1 diabetes” OR “inflammatory bowel diseases” OR “cystic fibrosis” OR “spina bifida” OR “Spinal Dysraphism” OR “rheumatic diseases” ) AND (adherence OR “self management” ) AND (infant OR child OR adolescent OR pediatric) Limiters—Publication Date: 2008-2016 Narrow by Language—English
APA PsycNet	(Any Field: asthma OR Any Field: epilepsy OR Any Field: transplant OR Any Field: HIV OR Any Field: sickle cell OR Any Field: Cancer OR Any Field: neoplasm OR Any Field: Dermatology OR Any Field: “Skin diseases” OR Any Field: “type 1 diabetes” OR Any Field: “inflammatory bowel diseases” OR Any Field: “cystic fibrosis” OR Any Field: “spina bifida” OR Any Field: “Spinal Dysraphism” OR Any Field: “rheumatic diseases”) AND (Any Field: adherence OR Any Field: “self management”) AND (Any Field: infant OR Any Field: child OR Any Field: adolescent OR Any Field: pediatric) AND Age Group: Childhood (birth-12 yrs) OR Neonatal (birth-1 mo) OR Infancy (2-23 mo) OR Preschool Age (2-5 yrs) OR School Age (6-12 yrs) OR Adolescence (13-17 yrs) AND Year: 2008 To 2019

evaluating the scientific rigor of assessments as either *adequate*, *good*, or *excellent* across the following domains: norms, internal consistency, interrater reliability, test-retest reliability, content validity, construct validity, validity generalization, treatment sensitivity, and clinical utility (Hunsley & Mash, 2018).

Due to the nature of PROs focused on adherence and self-management-related constructs, modifications were made to the original criteria (see Table II). For example, we excluded the criteria assessing norms since assessment of self-management in pediatric psychology rarely have norms or cut-off scores from large pediatric samples (Holmbeck et al., 2008). We also added the rating of *below adequate* to distinguish between measures with published data that did not meet criteria for an *adequate*, *good*, or *excellent* rating and those without published data for a particular criterion. Table II provides details regarding the rating categories for each criterion. The authors engaged in group discussion when there was lack of clarity for the rating.

## Results

### Studies Included

After removal of duplicates, 132 unique measures were identified. Seventy-one measures were excluded

based on inclusion and exclusion criteria, with an additional 11 removed after full-text review. Fifty fulfilled the inclusion criteria and were included in this review based on author consensus (Figure 1).

### Study Characteristics

#### Selected Measures

A comprehensive overview of the 50 included measures is presented in Table III. Measures were designed for and validated in youth or caregivers of youth with diabetes ( $n = 15$ ), asthma ( $n = 10$ ), spina bifida ( $n = 5$ ), solid organ and bone marrow transplant ( $n = 4$ ), cystic fibrosis ( $n = 2$ ), epilepsy ( $n = 2$ ), inflammatory bowel diseases ( $n = 2$ ), food allergy ( $n = 1$ ), HIV ( $n = 1$ ), or juvenile idiopathic arthritis ( $n = 1$ ). Two PROs were developed across pediatric conditions and an additional five were validated in more than one chronic illness population, all of which were grouped together (General Measures and Measures Adapted for Multiple Chronic Illness Populations;  $n = 7$ ).

Eighteen of the measures reviewed assessed disease management skills. Thirteen assessed adherence behaviors, eight assessed adherence barriers, and seven assessed treatment responsibility. Two measures assessed both adherence behaviors and adherence

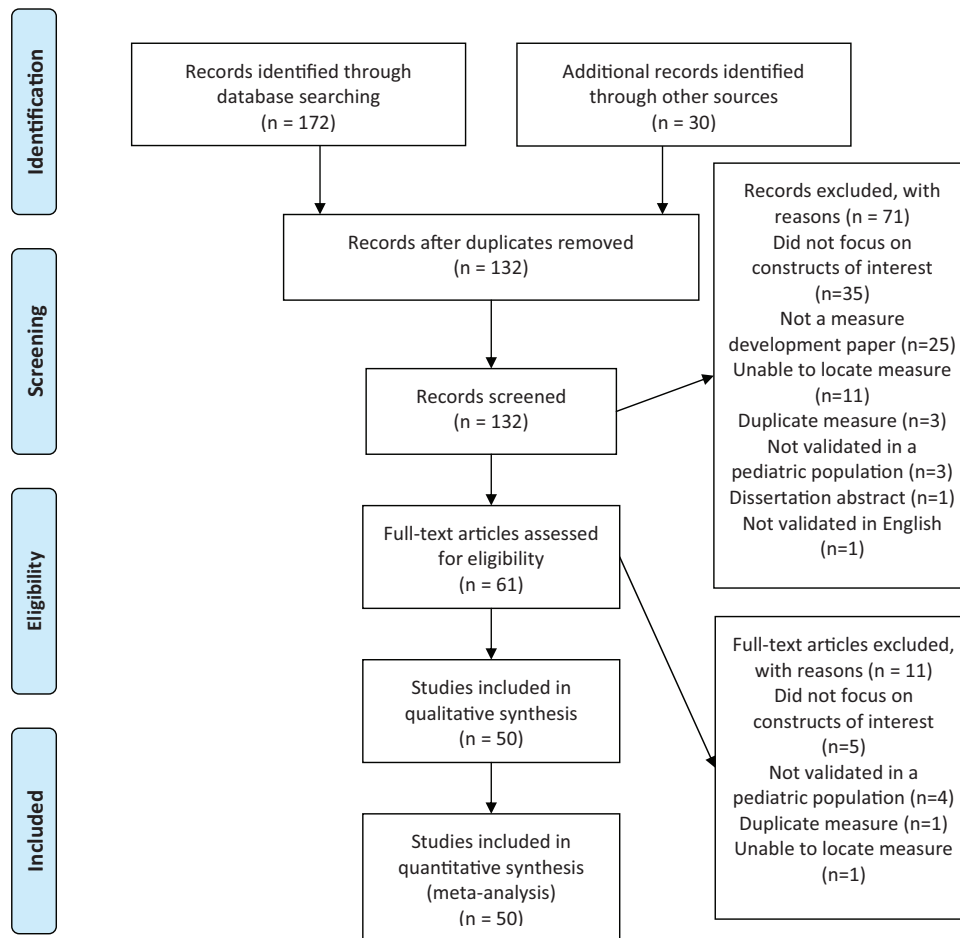


Figure 1. PRISMA flow diagram.

barriers, one assessed both adherence behaviors and treatment responsibility, and one assessed adherence behaviors, adherence barriers, and treatment responsibility.

### Quality of measures

Most measures reviewed received *good* ratings for content validity ( $n=41$ ) and either *adequate* ( $n=19$ ) or *good* ( $n=15$ ) ratings on internal consistency. Regarding validity generalization, most measures received *adequate* ( $n=31$ ) or *good* ( $n=14$ ) ratings on this criterion. Several measures received either *below adequate* ( $n=23$ ) or *good* ( $n=12$ ) ratings on construct validity. Most measures received either *below adequate* ratings for interrater reliability ( $n=14$ ) or could not be rated on this criterion due to insufficient data ( $n=27$ ). Thirty-three and 32 measures could not be rated on the domains of test–retest reliability or treatment sensitivity respectively, also due to insufficient data. All measures received *adequate* ratings for clinical utility. For ease of readability, measure names are included in the following subsections and the reader is referred to [Table III](#),

[Supplementary Table 1](#), and additional Supplementary data (sections titled Measure References and Additional References) which include references and full citations for each measure reviewed.

*Asthma.* Of the 10 asthma-specific PROs, four assessed adherence behaviors (Asthma Therapy Assessment Questionnaire; Family Asthma Management System Scale; Medication Intake Survey–Asthma; Pediatric Inhaler Adherence Questionnaire), three assessed adherence barriers (Illness Management Survey; Parent Barriers to Managing Asthma; Parent Belief in Treatment Efficacy), two assessed disease management skills (Child and Parent Asthma Self-Efficacy Scale; Reasoning About Managing Asthma), and one assessed treatment responsibility (Asthma Responsibility Questionnaire). Regarding internal consistency, seven measures received *adequate* or *good* ratings (*good*:  $n=4$ ; *adequate*:  $n=3$ ). Only two measures (Family Asthma Management System Scale; Asthma Responsibility Questionnaire) had published data on interrater reliability. Two measures reported test–retest



**Table II.** Modified Criteria Used to Evaluate Quality of Evidence-Based Assessments of Pediatric Self-Management

	Hunsley and Mash's (2018) criteria for evidence-based assessment	Modified criteria
Internal Consistency <sup>1</sup>	<i>Adequate:</i> Majority of alphas between .70 and .79 <i>Good:</i> Majority of alphas between .80 and .89 <i>Excellent:</i> Majority of alphas $\geq .90$	Not modified
Interrater Reliability <sup>2</sup>	<i>Adequate:</i> Majority of K values between .60 and .74, or Pearson/interclass correlations between .70 and .79 <i>Good:</i> Majority of K values between .75 and .84, or Pearson/interclass correlations between .80 and .89 <i>Excellent:</i> Majority of K values $\geq .85$ , or Pearson/interclass correlations $\geq .90$	Not modified
Test-Retest Reliability <sup>3</sup>	<i>Adequate:</i> Majority of correlations $\geq .70$ over several days to several weeks <i>Good:</i> Majority of correlations $\geq .70$ over several months <i>Excellent:</i> Majority of correlations $\geq .70$ over at least one year	Not modified
Content Validity	<i>Adequate:</i> Domain of the construct assessed was clearly defined and items were representative of the domain <i>Good:</i> In addition to criteria for an <i>adequate</i> rating, measure was evaluated by judges (e.g., experts, participants) <i>Excellent:</i> In addition to criteria for a <i>good</i> rating, measure development involved multiple groups of judges and quantitative ratings	Not modified
Construct Validity	<i>Adequate:</i> Replicated evidence of various types of construct validity (e.g., predictive, concurrent, convergent, discriminant) by independent research groups <i>Good:</i> Majority of evidence by independent research groups across multiple types of construct validity <i>Excellent:</i> In addition to criteria for a <i>good</i> rating, there is evidence of incremental validity with respect to clinical data	<i>Adequate:</i> Two or more independent research groups assessed construct validity with respect to one health-related outcome, <sup>a</sup> sociodemographic factors (i.e., distinguishing differences between groups), or other measures of adherence/self-management <i>Good:</i> Two or more independent research groups assessed construct validity across two or more health-related outcomes, sociodemographic factors, or other measures of adherence/self-management <i>Excellent:</i> Two or more independent research groups assessed construct validity across two or more health-related outcomes, sociodemographic factors, or other measures of adherence/self-management and demonstrated incremental validity (e.g., iterative versions of the measure showing improvement over time or adaptations of the same measure comparing versions in the same sample)
Validity Generalization	<i>Adequate:</i> Some evidence supports using the measure with either more than one specific group or in multiple contexts <i>Good:</i> Majority of evidence supports using the measure with more than one specific group or in multiple contexts <i>Excellent:</i> Majority of evidence supports using the measure with more than one specific group and in multiple contexts	<i>Adequate:</i> One or two independent research groups demonstrated use of the measure in more than one demographic group <sup>b</sup> or in multiple contexts <sup>c</sup> <i>Good:</i> Three or more independent research groups demonstrated use of the measure in more than one demographic group or in multiple contexts <i>Excellent:</i> Three or more independent research groups demonstrated use of the measure in more than one demographic groups <i>and</i> in multiple contexts
Treatment Sensitivity	<i>Adequate:</i> Some evidence of sensitivity to change pre-post treatment <i>Good:</i> Majority of evidence by independent research groups demonstrates sensitivity to change pre-post treatment <i>Excellent:</i> In addition to criteria for a <i>good</i> rating, evidence of sensitivity to change across different treatments	<i>Adequate:</i> One study that demonstrates sensitivity to change pre-post treatment <i>Good:</i> Two studies from independent research groups that demonstrate sensitivity to change pre-post treatment <i>Excellent:</i> More than two studies by independent research groups that demonstrate sensitivity to change pre-post treatment

(continued)

Table II. (continued)

	Hunsley and Mash's (2018) criteria for evidence-based assessment	Modified criteria
Clinical Utility	<p><i>Adequate:</i> Measure data are clinically useful while taking into account factors including cost, ease/availability of administration and scoring instructions, etc.</p> <p><i>Good:</i> In addition to criteria for an <i>adequate</i> rating, there is published evidence that using the measure results in clinical benefits</p> <p><i>Excellent:</i> In addition to criteria for an <i>adequate</i> rating, there is evidence by independent research groups that using the measure results in clinical benefits</p>	Not modified

Notes. Ratings of *below adequate* were given if published evidence of a criterion was found, but it did not reach an *adequate* rating. Ratings of *N/A* were given if published evidence of a criterion was not found. These two additional rating categories were developed for this systematic review and not included in Hunsley and Mash's original ratings.

<sup>a</sup>Health outcomes included clinical outcomes, health-related quality of life, health care utilization, psychological functioning, and alternative measures of adherence (e.g., medication possession ratios, electronic monitoring, daily diaries).

<sup>b</sup>Demographic characteristics included age, gender, and race/ethnicity.

<sup>c</sup>Contexts included setting (e.g., outpatient, inpatient) and type of administration (interview, paper-pencil, web-based).

<sup>1</sup>Internal consistency refers to correlations between items on the same measure to examine whether items meant to measure the same construct produce similar ratings.

<sup>2</sup>Interrater reliability refers to the consistency of ratings on a measure by multiple respondents (e.g., youth-report and parent-report).

<sup>3</sup>Test-retest reliability refers to the closeness of the agreement between results of the same measure used at multiple time points under the same conditions.

reliability and both were rated as *adequate* (Pediatric Inhaler Adherence Questionnaire; Illness Management Survey). All measures received a rating of *good* for content validity. Findings regarding construct validity and validity generalization were mixed. Of the four measures with data on treatment sensitivity, three received an *adequate* rating (Asthma Therapy Assessment Questionnaire; Family Asthma Management System Scale; Reasoning about Managing Asthma) and one received a *below adequate* rating (Asthma Responsibility Questionnaire).

*Cystic fibrosis.* Of the two PROs reviewed, one assessed disease management skills (Knowledge of Disease Management-CF) and one assessed treatment responsibility (The Self-Care Independence Scale). Both measures demonstrated *good* content validity, *adequate* validity generalization, and *below adequate* construct validity. Neither measure had published data on interrater reliability or treatment sensitivity.

*Diabetes.* Of the 15 total PROs, five PROs assessed adherence behaviors (Diabetes Adherence, Parent and Child; Diabetes Family Adherence Measure; Diabetes Management Questionnaire; Diabetes Regimen Adherence Questionnaire; Self-Care Inventory), two assessed adherence barriers (Barriers to Diabetes Adherence Measure for Adolescents; Self-Care Adherence Interview), six assessed disease management skills (Adolescent Diabetes Needs Assessment Tool; Diabetes Behavior Rating Scale; Diabetes Self-Management Profile Revised; Diabetes Strengths and Resilience Measure; PedCarbQuiz; Self-Management of Type 1 Diabetes in Adolescents), and two assessed

treatment responsibility (Collaborative Parent Involvement Scale; Diabetes Family Responsibility Questionnaire). Eleven measures received *good* ratings for content validity (Diabetes Adherence, Parent and Child; Diabetes Family Adherence Measure; Diabetes Management Questionnaire; Barriers to Diabetes Adherence Measure for Adolescents; Adolescent Diabetes Needs Assessment Tool; Diabetes Behavior Rating Scale; Diabetes Self-Management Profile Revised; Diabetes Strengths and Resilience Measure; Self-Management of Type 1 Diabetes in Adolescents; Collaborative Parent Involvement Scale; Diabetes Family Responsibility Questionnaire). Most measures received *adequate* ratings on internal consistency ( $n = 7$ ), test-retest reliability ( $n = 7$ ), and validity generalization ( $n = 10$ ). Findings regarding other criteria appeared mixed, with about half of the measures not having published literature on interrater reliability ( $n = 8$ ) and treatment sensitivity ( $n = 8$ ).

*Epilepsy.* Of the two PROs reviewed, one assessed adherence behaviors and adherence barriers (Pediatric Epilepsy Self-Management Questionnaire) and the other assessed disease management skills (Epilepsy Self-Management Scale). Both measures demonstrated *good* content validity, *adequate* internal consistency, and *below adequate* construct validity. One measure received an *adequate* rating for validity generalization and a *good* rating for treatment sensitivity (Pediatric Epilepsy Self-Management Questionnaire), while published data on these criteria was not available for the other measure (Epilepsy Self-Management Scale). Neither measure had published data on interrater or test-retest reliability.

*Food allergy.* One PRO was found for youth with food allergies (Food Allergy Management Perceptions Questionnaire). This measure demonstrated *good* content validity, *adequate* internal consistency and validity generalization, and *below adequate* construct validity. No data on interrater reliability, test–retest reliability, and treatment sensitivity were found.

*HIV.* One PRO was reviewed for youth with HIV. The P1042S Child/Adolescent Questionnaire and Parent/Caregiver Questionnaire assessed adherence behaviors. This measure received ratings of *good* for internal consistency, *adequate* for validity generalization, and *below adequate* for interrater reliability. There was no published data on test–retest reliability, content validity, and treatment sensitivity.

*Inflammatory Bowel Diseases.* One PRO assessed disease management skills (IBD-Knowledge Inventory Device) and the other assessed treatment responsibility (IBD Family Responsibility Questionnaire). Both measures demonstrated *good* content validity, *adequate* construct validity, and *adequate* validity generalization. Findings were mixed regarding internal consistency, interrater reliability, and test–retest reliability. Both measures also lacked published data on treatment sensitivity.

*Juvenile Idiopathic Arthritis.* One PRO assessing both adherence behaviors and treatment responsibility was found for youth with juvenile idiopathic arthritis (Parent Adherence Report Questionnaire). This measure received *good* ratings on content validity and construct validity. It demonstrated *adequate* validity generalization and *below adequate* internal consistency, interrater reliability, and test–retest reliability. No published data were available to assess treatment sensitivity.

*Spina Bifida.* Of the five PROs reviewed, one assessed adherence behaviors (Parent Report of Medical Adherence in Spina Bifida Scale), three assessed disease management skills (Adolescent/Young Adult Self-Management Independence Scale II; Kennedy Krieger Independence Scales – Spina Bifida Version; Spina Bifida Independence Survey), and one assessed treatment responsibility (Sharing of SB Management Responsibilities). All five measures received *adequate* ratings on validity generalization. No treatment sensitivity data were available for any of the measures, and data on test–retest reliability and interrater reliability was not available for most of the measures.

*Transplant (Solid Organ and Bone Marrow).* Of the four transplant-specific PROs, one assessed adherence barriers (Barriers Assessment Tool) and three assessed

disease management skills (Behavioral Affective and Somatic Experiences; Behavioral Affective and Somatic Experiences-Parent Version; Behavioral Affective and Somatic Experiences-Child Version). All four measures received *good* ratings for content validity. Regarding interrater reliability, one measure received a rating of *good* (Behavioral Affective and Somatic Experiences), while three were rated as *below adequate* (Barriers Assessment Tool; Behavioral Affective and Somatic Experiences-Parent Version; Behavioral Affective and Somatic Experiences-Child Version). Three measures had published data on treatment sensitivity, and all received *adequate* ratings (Behavioral Affective and Somatic Experiences; Behavioral Affective and Somatic Experiences-Parent Version; Behavioral Affective and Somatic Experiences-Child Version). Findings regarding internal consistency and validity generalization were mixed. There were no published data regarding test–retest reliability for any of the four transplant-specific measures. One measure had published data regarding construct validity and received a rating of *adequate* (Barriers Assessment Tool).

*General Measures and Measures Adapted for Multiple Chronic Illness Populations.* Although several measures were originally developed for a specific population, multiple they have been since translated for use in other chronic conditions. These measures include the 24-hr Recall Interview/Daily Phone Diary, the Medical Adherence Measure, the Adolescent and Parent Medication Barriers Scales, and the Allocation of Treatment Responsibility scale (Table III). In addition, two measures developed for general use across any disease population were reviewed (Chronic Disease Compliance Instrument; Self-Management Skills Scale).

Three measures in this group assessed adherence behaviors (24-hr Recall Interview/Daily Phone Diary; Chronic Disease Compliance Instrument; Medical Adherence Measure), two assessed adherence barriers (Adolescent Medication Barriers Scale; Parent Medication Barriers Scale), one assessed disease management skills (Self-Management Skills Scale), and one assessed treatment responsibility (Allocation of Treatment Responsibility). Seven measures received *good* ratings for content validity and one received a rating of *adequate* (24-hr Recall Interview/Daily Phone Diary). Four measures demonstrated *good* internal consistency (Adolescent Medication Barriers Scale; Parent Medication Barriers Scale; Self-Management Skills Scale; Allocation of Treatment Responsibility), one received a rating of *adequate* (Chronic Disease Compliance Instrument), and one demonstrated *below adequate* internal consistency (24-hr Recall Interview/Daily Phone Diary). Three



**Table III. Ratings Based on Hunsley and Mash's (2018) Criteria for Evidence-Based Assessments of Adherence Behaviors, Adherence Barriers, Disease Management Skills, and Treatment Responsibility**

	Measure and Authors	Number of Items/ Respondent	Internal Consistency	Interrater Reliability	Test-Retest Reliability	Content Validity	Construct Validity	Validity Generalization	Treatment Sensitivity	Clinical Utility
Asthma Adherence Behaviors	Asthma Therapy Assessment Questionnaire (ATAQ; Skinner et al., 2004)	20 items; caregiver report	○	—	—	***	***	***	**	**
	Family Asthma Management System Scale (FAMSS; Klinnert, McQuaid, & Gavin, 1997; McQuaid, Walders, Kopel, Fritz, & Klinnert, 2005)	Semi-structured in- tervention with parents of chil- dren with asthma ages 11-17	***	***	—	***	***	***	**	**
Adherence Barriers	Medication Intake Survey – Asthma (MIS-A; Dima et al., 2017)	6 items; youth ages ≥ 12 and parents of children ages 6-12	—	—	—	***	○	**	—	**
	Pediatric Inhaler Adherence Questionnaire (Martinez, Sossa, & Rand, 2008)	6 items; parent- report	—	—	**	***	○	**	—	**
	Illness Management Survey (IMS; Logan, Zeilikovsky, Labay, & Spengel, 2003)	36 items; youth- report	***	—	**	***	***	***	—	**
	Parent Barriers to Managing Asthma (Bursch, Schwankovsky, Gilbert, & Zeiger, 1999)	9 items; parent- report	**	—	—	***	○	**	—	**
	Parent Belief in Treatment Efficacy (Bursch et al., 1999)	5 items; parent- report	**	—	—	***	○	**	—	**
	Child and Parent Asthma Self-Efficacy Scale (Bursch et al., 1999)	13 items on the parent version; 14 items on the child version	***	—	—	***	○	**	—	**
Disease Management Skills	Reasoning About Managing Asthma (Kintner, Cook, Hull, & Meeder, 2013)	8 items; youth- report	**	—	—	***	○	**	**	**
	Asthma Responsibility Questionnaire (McQuaid, Penza-Clyve, Nassau, & Fritz, 2001)	10 items; youth- and parent- report	***	○	—	***	***	***	○	**
Cystic fibrosis	Knowledge of Disease Management – CF (KDM- Management – CF)	23 items; adoles- cent-report	**	—	—	***	○	**	—	**

(continued)

Table III. (continued)

	Measure and Authors	Number of Items/ Respondent	Internal Consistency	Intrater Reliability	Test-Retest Reliability	Content Validity	Construct Validity	Validity Generalization	Treatment Sensitivity	Clinical Utility
Disease Management Skills	CF; Bernstein, Riekert, & Quitner, 2018)									
	The Self-Care Independence Scale (SCIS; Patton, Graham, Varlotta, & Holsclaw, 2003)	44 items; parents of children ages 4-17	****	-	**	***	○	**	-	**
Treatment Responsibility	Diabetes Adherence, Parent and Child (Lehmkuhl et al., 2009)	20 items; parent- and child-report	**	○	**	***	○	**	-	**
	Diabetes Family Adherence Measure (DFAM; Lewin et al., 2010)	19 items and 4 validity items; youth-report	**	-	**	***	○	**	-	**
Diabetes Adherence Behaviors	Diabetes Management Questionnaire (DMQ; Mehta et al., 2015)	20 items; youth- and caregiver report	**	○	○	***	○	**	-	**
	Diabetes Regimen Adherence Questionnaire (DRAQ; Brownlee-Duffeck et al., 1987)	15 items; adolescents ages 8-17	**	-	-	-	○	**	-	**
Adherence Barriers	Self-Care Inventory (SCI; Davis et al., 2001; Delamater et al., 1997; Greco et al., 1990; La Greca, Swales, Klemm, & Madigan, 1988, April; Lewin et al., 2009; Wysocki et al., 2000)	14 items; parent- and adolescent-report	***	○	**	-	***	○	**	**
	Barriers to Diabetes Adherence Measure for Adolescents (Mulvaney et al., 2011)	21 items; adolescent-report	***	-	-	***	○	**	-	**
Disease Management Skills	Self-Care Adherence Interview (Hanson et al., 1989, 1992, 1996)	15 item semi-structured interview; youth ages 10-20 and parents	○	****	**	-	○	**	-	**
	Adolescent Diabetes Needs Assessment Tool (Cooper et al., 2014)	117 items; adolescent-report	-	-	**	***	○	**	-	**
	Diabetes Behavior Rating Scale (Iannotti et al., 2006)	Injection version: 36 items; pump version: 37 items;	***	○	**	***	**	***	**	**

(continued)

Table III. (continued)

Measure and Authors	Number of Items/ Respondent	Internal Consistency	Interrater Reliability	Test-Retest Reliability	Content Validity	Construct Validity	Validity Generalization	Treatment Sensitivity	Clinical Utility
Diabetes Self-Management Profile Revised (Harris et al., 2000; Iannotti et al., 2006)	youth and parents 29 items; semi-structured interview with youth and parents	**	****	○	***	****	***	**	**
Diabetes Strengths and Resilience Measure (Hilliard, Iturralde, Weissberg-Benchell, & Hood, 2017)	12 items; youth ages 8-18	***	-	-	***	○	**	**	**
PedCarbQuiz (Koonz et al., 2010)	78 items; youth and caregivers	***	-	-	****	**	***	-	**
Self-Management of Type 1 Diabetes in Adolescents (SMOD-A; Rechenberg, Whittemore, Grey, Jaser, & Group, 2016; Rechenberg, Whittemore, Holland, & Grey, 2017; Schilling et al., 2009)	52 items; youth ages 8-21	**	-	**	***	***	****	**	**
Collaborative Parent Involvement Scale (Nansel et al., 2009)	12 items; adolescent-report	****	-	-	***	***	**	○	**
Diabetes Family Responsibility Questionnaire (DFRQ; Anderson, Auslander, Jung, Miller, & Santiago, 1990)	17 item interview; youth ages 6-21 and their parents	**	**	-	***	***	**	**	**
Pediatric Epilepsy Self-Management Questionnaire (Carbone, Zebrack, Plegue, Joshi, & Shellhaas, 2013; Modi, Monahan, Daniels, & Glauser, 2010)	27 items; caregivers of youth ages 2-14 and adolescents ages 12-17	**	-	-	***	○	**	***	**
Epilepsy Adherence Behaviors and Barriers									
Disease Management Skills	26 items; older adolescents and adults ages 17-66	**	-	-	***	○	-	-	**

(continued)

Table III. (continued)

Measure and Authors	Number of Items/ Respondent	Internal Consistency	Interrater Reliability	Test-Retest Reliability	Content Validity	Construct Validity	Validity Generalization	Treatment Sensitivity	Clinical Utility
Food Allergy Adherence Behaviors and Barriers	Food Allergy Management Perceptions Questionnaire (FAMPQ; Herbert, Lin, Matsui, Wood, & Sharma, 2016)	25 items; youth ages 13-21	**	-	***	○	**	-	**
HIV Adherence Behaviors, Treatment Responsibility	P1042S Child/Adolescent Questionnaire and Parent/Caregiver Questionnaire (Buchanan et al., 2012; Farley, Hines, Musk, Ferrus, & Tepper, 2003)	30 items; youth ages 8-18 and their caregivers	***	○	-	****	**	-	**
Inflammatory Bowel Disease Management Skills	IBD-Knowledge Inventory Device (IBD-KID; Haaland, Day, & Otley, 2014)	25 items; youth ages 10-17 and their parents	**	**	***	**	**	-	**
Treatment Responsibility	IBD Family Responsibility Questionnaire (IBD-FRQ; Greenley, Doughly, Stephens, & Kugathasan, 2010)	23 items; adolescents ages 11-18 and their caregivers	○	○	***	**	**	-	**
Juvenile Idiopathic Arthritis Adherence Behaviors, Treatment Responsibility	Parent Adherence Report Questionnaire (PARQ; De Civita, Dobkin, Ehrmann-Feldman, Karp, & Duffy, 2005)	Caregivers of youth ages 15-18	○	○	***	***	**	-	**
Spina Bifida Adherence Behaviors	Parent Report of Medical Adherence in Spina Bifida Scale (PROMASB; Holmbeck et al., 1998)	39 items; parents of youth ages 8-9	○	○	**	-	**	-	**
Disease Management Skills	Adolescent/Young Adult Self-Management and Independence Scale II (Buran, Brei, Sawin, Stevens, & Neufeld, 2006; Sawin, Heffelfinger, Cashin, & Brei, 2018)	17 items; adolescents ages 12-21 and their parents	**	**	***	***	**	-	**
	Kennedy Krieger Independence Scales –	18 items; caregivers of youth and	***	-	***	○	**	-	**

(continued)

**Table III. (continued)**

Measure and Authors	Number of Items/ Respondent	Internal Consistency	Interrater Reliability	Test-Retest Reliability	Content Validity	Construct Validity	Validity Generalization	Treatment Sensitivity	Clinical Utility
Spina Bifida Version (Jacobson et al., 2013)	young adults ages 10-29	○	-	-	**	○	**	-	**
Spina Bifida Independence Survey (Psihogios, Kolbuck, & Holmbeck, 2015)	50 items; youth ages 8-15 and their parents								
Treatment Responsibility	34 items; youth ages 8-15 and their parents	○	-	-	**	○	**	-	**
Solid Organ Transplant Adherence Barriers	Barriers Assessment Tool (Varnell et al., 2017)	-	○	-	***	-	**	-	**
Disease Management Skills	Behavioral Affective and Somatic Experiences (BASES; Phipps, Hinds, Channell, & Bell, 1994)	**	***	-	***	-	***	**	**
Disease Management Skills	Behavioral Affective and Somatic Experiences, Parent Version (BASES-P; Phipps et al., 1994)	**	○	-	***	-	***	**	**
General Measures and Behaviors	Behavioral Affective and Somatic Experiences, Child Version (BASES-C; Phipps et al., 1994)	○	○	-	***	**	***	**	**
Adherence	24-hr Recall Interview/Daily Phone Diary (Johnson, Silverstein, Rosenbloom, Carter, & Cunningham, 1986; Marhefka, Tepper, Farley, Sleasman, & Mellins, 2006; Naar-King, Frey, Harris, & Arfken, 2005)* Initially validated in youth with diabetes and subsequently validated in cystic fibrosis, asthma, HIV, epilepsy, obesity, spina bifida, and autism spectrum disorder	○	**	**	**	***	***	**	**

(continued)



Table III. (continued)

Measure and Authors	Number of Items/ Respondent	Internal Consistency	Interrater Reliability	Test-Retest Reliability	Content Validity	Construct Validity	Validity Generalization	Treatment Sensitivity	Clinical Utility
Chronic Disease Compliance Instrument (Kyngas, Skaar-Chandler, & Duffy, 2000)	13 background questions and 41 items; adolescents ages 13-17 with diabetes (expanded to asthma, epilepsy, and rheumatoid arthritis)	**	-	-	***	○	***	-	**
Medical Adherence Measure (Zelikovsky & Schast, 2008) *Initially validated in renal transplant and subsequently validated in inflammatory bowel disease	Semi-structured clinical interview; youth ages 13-23	-	-	****	***	***	**	-	**
Adherence Barriers Adolescent Medication Barriers Scale (AMBS; Simons & Blount, 2007) *Initially validated in solid organ transplant and subsequently validated in inflammatory bowel disease	17 items; adolescents ages 11-21	***	○	-	***	****	***	**	**
Parent Medication Barriers Scale (PMBS; Simons & Blount, 2007) *Initially validated in solid organ transplant and subsequently validated in stem cell transplant	16 items; parents of youth ages 11-21	***	○	-	***	****	***	**	**
Disease Management Skills Self-Management Skills Scale (Klassen et al., 2015)	15 items; adolescent-report	***	-	-	***	-	-	-	**
Treatment Responsibility Allocation of Treatment Responsibility (Pai et al., 2010) *Initially validated in transplant and subsequently validated in epilepsy and sickle cell disease	18 items; youth ages 7-18 and their caregivers	***	**	-	***	***	***	-	**

Note. ○ = Below Adequate; \*\* = Adequate; \*\*\* = Good; and \*\*\*\* = Excellent; - = Not Applicable; For measures with parent and child reports, ratings reflect psychometric properties of both reporters.

measures had published data on treatment sensitivity (24-hr Recall Interview/Daily Phone Diary; Adolescent Medication Barriers Scale; Parent Medication Barriers Scale) and all three were rated as *adequate*. Of the four studies with published data available on interrater reliability, two received *adequate* ratings (24-hr Recall Interview/Daily Phone Diary; Allocation of Treatment Responsibility), and two were rated *below adequate* (Adolescent Medication Barriers Scale; Parent Medication Barriers Scale). Ratings for validity generalization and construct validity were mixed.

## Discussion

Our review highlights the widespread use of PROs assessing pediatric adherence and self-management and identified 50 measures across several illness populations. Although objective measures, including electronic monitoring and biomarkers, are typically considered the gold standard and viewed as more reliable than self-report (Hommel, Davis, & Baldassano, 2009; Landier et al., 2017), PROs continue to be widely utilized in both research methodology and clinical practice due to accessibility and ease of use (Duncan, Mentrikoski, Wu, & Fredericks, 2014; Muller et al., 2011). The present review expands upon Quittner et al.'s (2008) review of evidence-based assessment of pediatric adherence by focusing on PROs specifically and evaluating new measures across disease populations using newly disseminated and rigorous criteria put forth by Hunsley and Mash (2018).

In general, most of the measures included in this review received *good* or *adequate* ratings for content validity and internal consistency, whereas construct validity was mostly *below adequate* with only about one third receiving a rating of *good*. Notably, *below adequate* ratings, particularly in construct validity, were often the result of the measure only being used by one research team, or lack of studies replicating findings. In addition, only one of the measures (PedCarbQuiz; Koontz et al., 2010) achieved an *excellent* rating for content validity because none of the other original validation studies presented data on quantitative ratings of items by stakeholders during measure development as suggested by Hunsley and Mash's (2018) criterion for content validity of evidence-based assessments.

Most of the measures reviewed received *adequate* or *good* ratings on Hunsley and Mash's (2018) criterion for validity generalization, in part due to use across settings or by multiple groups (e.g., youth, caregiver, and medical provider). Measures reviewed were used in a variety of settings (e.g., community, clinical, and research) and most involved caregiver report,

which reflects existing recommendations for multi-informant reporting for adherence and self-management behaviors (Greenley, Kunz, Walter, & Hommel, 2013). However, only 11 measures included both adolescents and young adults ( $\geq 18$  years) in their validation samples, despite the significant risk of poor adherence and self-management during this developmental period. Unfortunately, this significantly limits generalization of many of the reviewed measures across developmental stages (Pai & Ostendorf, 2011). Moreover, merely seven measures were noted to be validated with racially diverse samples. This is likely a function of the diagnoses included such as type 1 diabetes, epilepsy, and inflammatory bowel diseases which tend to be racially and ethnically homogenous patient populations. Overall, while our review highlights that measures of adherence and self-management have been validated for youth with a range of chronic conditions, much work needs to be done to ensure these measures are invariant and reliable across age ranges, socioeconomic status, and minority groups.

There are important limitations in the extant literature from which this systematic review was conducted which carry substantial implications for future research. For example, it is clear from this review that some chronic condition populations are much further developed in terms of utilizing illness-specific adherence and self-management PROs. Type 1 diabetes and asthma together comprised 50% of the measures included in this review, indicating a need for research in other pediatric chronic conditions. Cystic fibrosis, despite the highly complex, time-consuming, and taxing self-management demands, had only two disease-specific measures included in this review. Other conditions reviewed in this study, such as inflammatory bowel diseases, epilepsy, and food allergies have only recently (i.e., last 15 years) been studied in the context of self-management research. Still, other conditions such as cancer, headache, sleep disorders, obesity, and sickle cell disease did not have measures that met inclusion criteria for this review. Although several PROs were identified in these disease groups, they did not meet inclusion for this systematic review due to the absence of psychometric data. This highlights the critical need for more research evaluating the psychometric properties of existing measures in order to establish their evidence-base and guide recommendations for use.

Another limitation of the current review is that Hunsley and Mash's (2018) rating system does not include a specific benchmark for what constitutes an "evidence-based" PRO of pediatric adherence and self-management. Hunsley and Mash's framework combined with the lack of PROs achieving an *adequate* rating (representative of a "minimal level of

scientific rigor”) across all criteria preclude us from making recommendations regarding the “best” PROs. Instead, we encourage researchers and clinicians to consider the strengths and weaknesses of each PRO in the context of its intended use as some criteria may be more salient than others in certain situations (e.g., treatment sensitivity to evaluate change as a result of an intervention versus validity generalization if you want to use the PRO in a particular subgroup). In addition, we were unable to rate important psychometric domains (e.g., including interrater reliability, test-retest reliability, or treatment sensitivity) for a large number of measures. Few studies reported this type of data, potentially due to page limitations in journals and a lack of interest in measurement-focused research from both a funding and publishing standpoint. Last, none of the measures reviewed received *good* or *excellent* ratings on clinical utility, which would require that published data demonstrate clinical improvement as a result of using a particular measure. Thus, it will be important for future studies to document these psychometric data in order to allow for a higher level of scientific rigor and, consequently, more accurate and useful adherence and self-management PROs for youth with chronic conditions.

Based on our systematic review, future research and clinical initiatives should address the following areas. First, our field has been prolific at the development of adherence and self-management PROs in several chronic conditions. However, substantial variation exists in clinical practice and research with respect to the use of specific PROs. The field will advance more rapidly when collaborations across sites are formed, with the goals of increasing uniform measurement and decreasing variability in the measures employed to assess adherence and self-management PROs. For example, the PedsQL has become a well-used PRO to assess health-related quality of life across diseases (Varni, Seid, & Kurtin, 2001), with disease-specific PedsQL modules used in adjunct. A core adherence and self-management PRO with adjunct disease-specific modules is one potential area for future research and collaboration. Related to this, our field has seen a significant increase in the number of clinical trials targeting adherence and self-management in pediatric populations. It will be important to use validated PROs that have the strongest psychometric properties in trials examining the effects of adherence and self-management interventions on behavioral or health outcomes.

A second area for future research and clinical work involves the use of mHealth or other health technology solutions to advance pediatric self-management PROs. Using emerging technologies to electronically collect pediatric self-management PRO data may increase the feasibility of implementation into clinical

care, including integration into the medical record (Jensen et al., 2015). If PROs can be routinely collected using health technologies, it may be increasingly feasible for health care systems to consider and use PROs such as those included in this review as health care quality metrics (Bevans et al., 2014).

There are numerous mobile applications that include self-management PROs and most provide some level of behavioral intervention (Carmody, Denson, & Hommel, 2018); however, they lack empirical testing and support. These apps provide real-time PRO data that researchers and clinicians can access on the back-end through agreements with technology companies and app developers. In instances where researchers and clinicians do not have access to back-end data, patient-facing data may be accessed to inform clinical decisions and answer basic research questions. To enable access to patient-facing data, clinicians and researchers may select an app whose interface allows providers access to patient-reported data (e.g., MediSafe Providers Portal) or ask youth and families to generate their own reports from the patient-facing interface and bring them to clinic or research visits (e.g., Asthma Health Storylines). Ecological momentary assessments can also be deployed via mobile devices and tailored to fit the needs of specific patients or research questions.

While this study identified several promising PROs, research examining the methods or workflow needed to optimize widespread use of pediatric adherence and self-management PROs in various health care settings is needed, including identifying both the best individuals to champion their use and the best health care providers to administer or retrieve data from them. There are opportunities for dissemination and implementation of pediatric adherence and self-management PROs through professional psychology, medical, and disease-specific organizations, as well as national groups, focused on quality of care and quality improvement. This area also requires some consensus on which tools should be widely used, so uniformity of assessment is essential. While several of the PROs we evaluated have *adequate*, *good*, and *excellent* ratings, more rigorous validation and reporting of psychometric properties (e.g., content validity, retest reliability, and treatment sensitivity) are needed. We recommend that additional research examine which PROs work best for patients within a specific disease population to ensure that the most effective tools are utilized and disseminated broadly.

Ultimately, the continued use of pediatric adherence and self-management PROs depends on improved psychometric data quality and quantity, a shared commitment to using a limited set of validated tools that best assess a given construct, optimized methods of obtaining data, and spreading the use of

the PROs in clinical practice so that patients receive the most benefit in care provision.

### Supplementary Data

Supplementary data can be found at: <https://academic.oup.com/jpepsy>.

### Funding

This work was supported by the National Institutes of Health (grant number NICHD T32 HD 68223-7).

*Conflicts of interest:* None declared.

### Acknowledgements

The authors gratefully acknowledge the support of Elanie Grigg Dean and supporting staff at the Edward L. Pratt Library at Cincinnati Children's Hospital in providing guidance on the systematic review process. We also gratefully acknowledge the assistance of clinical research coordinators for their involvement with this project.

### References

- Annunziato, R. A., Bucuvalas, J. C., Yin, W., Arnand, R., Alonso, E. M., Mazariegos, G. V., ... Shemesh, E. (2018). Self-management measurement and prediction of clinical outcomes in pediatric transplant. *The Journal of Pediatrics*, *193*, 128–133.
- Bevans, K. B., Moon, J., Carle, A. C., Mara, C. A., Lai, J.-S., DiMarco, L., ... Woods, D. (2014). Patient reported outcomes as indicators of pediatric health care quality. *Academic Pediatrics*, *14*, S90–S96.
- Bhatia, S., Landier, W., Shangquan, M., Hageman, L., Schaible, A. N., Carter, A. R., ... Wong, F. L. (2012). Nonadherence to oral mercaptopurine and risk of relapse in Hispanic and non-Hispanic white children with acute lymphoblastic leukemia: A report from the children's oncology group. *Journal of Clinical Oncology*, *30*, 2094–2101.
- Carmody, J. K., Denson, L. A., & Hommel, K. A. (2018). Content and usability evaluation of medication adherence mobile applications for use in pediatrics. *Journal of Pediatric Psychology*, *33*, 333–342.
- Delamater, A. M., Applegate, B., Shaw, K. H., Edison, M., Szapocznik, J., & Nemery, R. (1997). What accounts for poor metabolic control in minority youths with diabetes. *Annals of Behavioral Medicine*, *19*, SO64.
- Ducharme, F. M., Zemek, R. L., Chalut, D., McGillivray, D., Noya, F. J. D., Resendes, S., ... Zhang, X. (2011). Written action plan in pediatric emergency room improves asthma prescribing, adherence, and control. *American Journal of Respiratory and Critical Care Medicine*, *183*, 195–203.
- Duncan, C. L., Mentrikoski, J. M., Wu, Y. P., & Fredericks, E. M. (2014). Practice-based approach to assessing and treatment non-adherence in pediatric regimens. *Clinical Practice in Pediatric Psychology*, *2*, 322–336.
- Fredericks, E. M., Lopez, M. J., Magee, J. C., Shieck, V., & Oipari-Arrigan, L. (2007). Psychological functioning, nonadherence, and health outcomes after pediatric liver transplantation. *American Journal of Transplantation*, *7*, 1974–1983.
- Goodhand, J. R., Kamperidis, N., Sirwan, B., Macken, L., Tschuma, N., Koodun, Y., ... Lindsay, J. O. (2013). Factors associated with thiopurine non-adherence in patients with inflammatory bowel disease. *Alimentary Pharmacology & Therapeutics*, *38*, 1097–1108.
- Graves, M. M., Roberts, M. C., Rapoff, M., & Boyer, A. (2010). The efficacy of adherence interventions for chronically ill children: A meta-analytic review. *Journal of Pediatric Psychology*, *35*, 368–382.
- Greenley, R. N., Kunz, J. H., Walter, J., & Hommel, K. A. (2013). Practical strategies for enhancing adherence to treatment regimen in inflammatory bowel disease. *Inflammatory Bowel Diseases*, *19*, 1534–1545.
- Hilliard, M. E., Ramey, C., Rohan, J. M., Drotar, D., & Cortina, S. (2011). Electronic monitoring feedback to promote adherence in an adolescent with fanconi anemia. *Health Psychology*, *30*, 503–509.
- Holmbeck, G. N., Thill, A. W., Bachanas, P., Garber, J., Miller, K. B., Abad, M., ... Zukerman, J. (2008). Evidence-based assessment in pediatric psychology: Measures of psychosocial adjustment and psychopathology. *Journal of Pediatric Psychology*, *33*, 958–980.
- Hommel, K. A., Davis, C. M., & Baldassano, R. N. (2009). Objective versus subjective assessment of oral medication adherence in pediatric inflammatory bowel disease. *Inflammatory Bowel Diseases*, *15*, 589–593.
- Hommel, K. A., McGrady, M. E., Peugh, J., Zacur, G., Loreaux, K., Saeed, S., ... Denson, L. A. (2017). Longitudinal patterns of medication nonadherence and associated health care costs. *Inflammatory Bowel Diseases*, *23*, 1577–1583.
- Hood, K. K., Peterson, C. M., Rohan, J. M., & Drotar, D. (2009). Association between adherence glycemic control in pediatric type 1 diabetes: A meta-analysis. *Pediatrics*, *124*, e1171–e1179.
- Hunsley, J., & Mash, E. J. (2018). *A guide to assessments that work* (2nd edn). New York, NY: Oxford University Press.
- Jensen, R. E., Rothrock, N. E., DeWitt, E. M., Spiegel, B., Tucker, C. A., Crane, H. M., ... Crane, P. K. (2015). The role of technical advances in the adoption and integration of patient-reported outcomes in clinical care. *Medical Care*, *53*, 153–159.
- Kahana, S., Drotar, D., & Frazier, T. (2008). Meta-analysis of psychological interventions to promote adherence to treatment in pediatric chronic health conditions. *Journal of Pediatric Psychology*, *33*, 590–611.
- Koontz, M. B., Cuttler, L., Palmert, M. R., O'Riordan, M., Borawski, E. A., McConnell, J., & Kern, E. O. (2010). Development and validation of a questionnaire to assess carbohydrate and insulin-dosing knowledge in youth with type 1 diabetes. *Diabetes Care*, *33*, 457–462.
- Landier, W., Chen, Y., Hageman, L., Kim, H., Bostrom, B. C., Casillas, J. N., ... Bhatia, S. (2017). Comparison of self-report and electronic monitoring of 6MP intake in

- childhood ALL: A Children's Oncology Group study. *Blood*, 129, 1919–1926.
- Mash, E. J., & Hunsley, J. (2005). Evidence-based assessment of child and adolescent disorders: Issues and challenges. *Journal of Clinical Child and Adolescent Psychology*, 34, 362–379.
- McGrady, M. E., Holbein, C. E., Smith, A. W., Morrison, C. F., Hommel, K. A., Modi, A. C., . . . Ramsey, R. R. (2018). An independent evaluation of the accuracy and usability of electronic adherence monitoring devices. *Annals of Internal Medicine*, 169, 419–422.
- Modi, A. C., Pai, A. L., Hommel, K. A., Hood, K. K., Cortina, S., Hilliard, M. E., . . . Drotar, D. (2012). Pediatric self-management: A framework for research, practice, and policy. *Pediatrics*, 129, e473–e485.
- Modi, A. C., Rausch, J. R., & Glauser, T. A. (2014). Early pediatric antiepileptic drug nonadherence is related to lower long-term seizure freedom. *Neurology*, 82, 671–673.
- Modi, A. C., Wu, Y. P., Guilfoyle, S. M., & Glauser, T. A. (2012). Uninformed clinical decisions resulting from lack of adherence assessment in children with new onset epilepsy. *Epilepsy & Behavior*, 25, 481–484.
- Muller, A. D., Jaspan, H. B., Myer, L., Hunter, A. L., Harling, G., Bekker, L., & Orrell, C. (2011). Standard measures are inadequate to monitor pediatric adherence in a resource-limited setting. *AIDS and Behavior*, 15, 422–431.
- Pai, A. L., & McGrady, M. (2014). Systematic review and meta-analysis of psychological interventions to promote treatment adherence in children, adolescents, and young adults with chronic illness. *Journal of Pediatric Psychology*, 39, 918–931.
- Pai, A. L., & Ostendorf, H. M. (2011). Treatment adherence in adolescents and young adults affected by chronic illness during the health care transition from pediatric to adult health care: A literature review. *Children's Health Care*, 40, 16–33.
- Pizzo, H. P., Ettenger, R. B., Gjertson, D. W., Reed, E. F., Zhang, J., Gritsch, H. A., & Tsai, E. W. (2016). Sildenafil and tacrolimus coefficient of variation is associated with rejection, donor-specific antibodies, and nonadherence. *Pediatric Nephrology*, 31, 2345–2352.
- Quittner, A. L., Modi, A. C., Lemanek, K. L., Ievers-Landis, C. E., & Rapoff, M. A. (2008). Evidence-based assessment of adherence to medical treatments in pediatric psychology. *Journal of Pediatric Psychology*, 33, 916–936.
- Rapoff, M. (Ed.) (2010). *Adherence to pediatric medical regimens* (2nd edn). New York, NY: Springer Science+Business Media.
- Stirratt, M. J., Dunbar-Jacob, J., Crane, H. M., Simoni, J. M., Czajkowski, S., Hilliard, M. E., . . . Nilsen, W. J. (2015). Self-report measures of medication adherence behavior: Recommendations on optimal use. *Translational Behavioral Medicine*, 5, 470–482.
- Van Cleave, J., Gortmaker, S. L., & Perrin, J. M. (2010). Dynamics of obesity and chronic health conditions among children and youth. *JAMA*, 303, 623–630.
- Varni, J., Seid, M., & Kurtin, P. (2001). PedsQL 4.0: Reliability and validity of the Pediatric Quality of Life Inventory version 4.0 generic core scales in health and patient populations. *Medical Care*, 39, 800–812.
- Wu, Y. P., & Pai, A. L. (2014). Health care provider-delivered adherence promotion interventions: A meta-analysis. *Pediatrics*, 133, e1698–e1707.