Minimally Important Differences in Patient or Proxy-Reported Outcome Studies Relevant to Children: A Systematic Review

Shanil Ebrahim, MSc, PhD,[a,b,c] Kelsey Vercammen, BSc,[d] Arunima Sivanand, BSc,[e] Gordon H. Guyatt, MD, MSc,[b,f] Alonso Carrasco-Labra, DDS, PhD,[b,g] Ricardo M. Fernandes, MD, PhD,[h] Mark W. Crawford, MD,[i] Gihad Nesrallah, MD, MSc,[j,k,l] Bradley C. Johnston, PhD[a,b,i,m]

abstract

CONTEXT: No study has characterized and appraised all anchor-based minimally important differences (MIDs) associated with patient-reported outcome (PRO) instruments in pediatric studies.

OBJECTIVE: To complete a comprehensive systematic survey and appraisal of published anchor-based MIDs associated with PRO instruments used in children.

DATA SOURCES: Medline, Embase, and PsycINFO (1989 to February 11, 2015).

STUDY SELECTION: Studies reporting empirical ascertainment of anchor-based MIDs among PROs used in pediatric care.

DATA EXTRACTION: All pertinent data items related to the characteristics of PRO instruments, anchors, and MIDs.

RESULTS: Of 4179 unique citations, 30 studies (including 32 cohorts) proved eligible and reported on 28 unique PROs (8 generic, 13 disease-specific, 5 symptoms-specific, 2 function-specific), with 9 (32%) classified as patient-reported, 11 (39%) proxy-reported, and 8 (29%) both patient- and proxy-reported. Of the 30 studies, we rated 14 (44%) as providing highly credible estimates of the MID. Most cohorts (n = 20, 62%) recorded patients' direct response to the target PRO and the use of an independent standard of comparison (n = 25, 78%). Most, however, failed to effectively report measurement properties of the anchor (n = 24, 75%).

LIMITATIONS: We have not yet addressed the measurement properties of instrument to measure credibility; our search was restricted to 3 electronic sources, and we used a single data abstractor.

CONCLUSIONS: Our study found 28 PROs that have been developed for children, with fewer than half providing credible estimates. Clinicians, clinical trialists, systematic reviewers, and guideline developers seeking to effectively summarize and interpret results of studies addressing PROs in child health are likely to find our comprehensive compendium of MIDs of use, both in providing best estimates of MIDs and identifying credible estimates.

[a]*Systematic Overviews through advancing Research Technology (SORT), Child Health Evaluative Sciences, The Research Institute, and* [i]*Department of Anaesthesia and Pain Medicine, The Hospital for Sick Children, Toronto, Ontario, Canada; Departments of* [b]*Clinical Epidemiology and Biostatistics and* [f]*Medicine, and Faculty of Health Sciences, McMaster University, Hamilton, Ontario, Canada;* [c]*Meta-Research Innovation Center at Stanford (METRICS), Stanford University, Stanford, California;* [d]*Department of Epidemiology, Harvard University, Cambridge,*

To cite: Ebrahim S, Vercammen K, Sivanand A, et al. Minimally Important Differences in Patient or Proxy-Reported Outcome Studies Relevant to Children: A Systematic Review. *Pediatrics.* 2017;139(3):e20160833

Patient-reported outcomes (PROs) provide patients' perspectives regarding treatment benefits and harms, and are often the outcomes of greatest importance to patients. The PRO literature has grown considerably over the past 3 decades[1]; clinical trialists increasingly use PRO instruments as primary outcomes. PROs are also used in monitoring quality and performance in health systems, and have become a priority for research funding agencies worldwide.[2,3] Although evidence supporting reliability, validity, and responsiveness exists for many PROs, interpretation of their results remains a challenge; for a given instrument, what change in a score constitutes a trivial, small but important, moderate, or large treatment effect?

The minimally important difference (MID) provides a measure of the smallest change in the PRO of interest that patients perceive as important, either beneficial or harmful.[4,5] The MID can be helpful for patients, clinicians, and clinical practice guideline developers when considering the trade-off between beneficial and harmful outcomes, and for clinical trialists planning sample sizes for their studies. There are multiple methods to assess MIDs, including anchor-based methods, distribution-based methods, Delphi-methods, and scale-judgment methods with anchor- and distribution-based methods being the 2 primary approaches used to estimate MIDs.[6,7] The anchor-based approach is generally considered the optimal way to determine the MID because it directly captures the patients' preferences and values.[6,7]

In pediatric populations, clinicians and researchers use PRO instruments to measure symptoms, disease severity, mental health, development, functional ability, and other constructs. For children too young to answer for themselves, proxy-respondents must substitute for the patient; such measures represent the best way of assessing the patients' subjective health state in these circumstances. No study or database has thus far systematically documented all available anchor-based MIDs associated with patient- or proxy-reported instruments for children. Given that clinical trialists, systematic review authors, and guideline panels are likely to find a compendium of trustworthy MID estimates of considerable use, we conducted a systematic survey to summarize all published anchor-based MIDs associated with PRO instruments used to evaluate the effects of interventions on chronic medical and psychiatric conditions in pediatric populations.

METHODS

A previously published protocol, summarized briefly in this article, provides additional details of our methods.[1]

Eligibility Criteria

We included original reports of studies that document the development of anchor-based MIDs for PRO instruments designed for chronic medical and psychiatric conditions in pediatric populations (<18 years of age). We defined an anchor-based approach as any independent assessment to which the PRO instrument is compared, irrespective of the interpretability or the quality of the anchor. PROs of interest included self-reported patient-important outcomes of health-related quality of life, functional ability, symptom severity, and measures of psychological distress and well-being.

Although self-reported measures are likely to provide more valid anchors for generating MID estimates, proxy-respondents (parents, caregivers, or clinicians) are often called on to respond on behalf of pediatric patients, particularly when patients are too young or are incapable due to disability. We therefore included studies in which a proxy completed the PRO instrument and/or the anchor.

We excluded studies in which only the clinician completed the PRO instrument, and excluded studies reporting only distribution-based MIDs without an accompanying anchor-based MID.

Information Sources and Search

We searched Medline, Embase, and PsycINFO for studies published from 1989 to February 11, 2015, by using relevant medical subject headings. Our published protocol describes the Medline search strategy.[1]

Study Selection

Two reviewers independently screened titles and abstracts to identify potentially eligible citations. Subsequently, to determine eligibility, teams of 2 reviewers reviewed the full texts of citations identified as potentially eligible.

Data Collection, Items, and Extraction

Pairs of investigators independently extracted data by using a pilot-tested data collection form consisting of the following items: study design, description of population, interventions, outcomes, and characteristics of PRO instruments, anchors, and MID assessment (data collection forms provided in protocol manuscript[1]). We classified PROs as "generic" if they measured health profiles not specific to a disease state/population or symptom or function (eg, Child Health Questionnaire, Short Form-36), or "specific" if they were specific to a disease state or population (eg, Hydrocephalus Outcome Questionnaire), a particular symptom (Visual Analog Scale [VAS] for Pain), or a particular function (eg, Oxford Ankle Foot Questionnaire). One

independent extractor verified all data.

Methodologists familiar with MID methods (S.E., B.C.J.) reviewed a pool of 10 eligible studies, and used standard thematic analysis techniques[8] to abstract concepts related to the methodological quality of MID determinations. They reviewed coding and revised the taxonomy of methodological factors iteratively until informational redundancy and consensus were achieved. Based on this initial survey of the literature and our group's experience with methods of ascertaining MIDs,[9–16] we developed criteria for evaluating the credibility of anchor-based MID determinations. Our group has previously used such methods successfully for developing methodological quality appraisal standards across a wide range of topics.[17–21]

The MID credibility instrument includes 6 items, with greater credibility if (1) patients directly responded to PRO; (2) investigators used an independent standard of comparison (instead of the same instrument to assess hypothetical scenarios); (3) the anchor was interpretable for patients; (4) the anchor was interpretable to clinicians; (5) the anchor was sufficiently closely empirically related to the target PRO; and (6) the anchor measurement properties (validity and reliability) were reported and satisfactory (validity and reliability coefficients >0.5). Each of the 6 criteria were judged by using 4 response options: definitely no, not so much, to a great extent, and definitely yes. Assessors resolved disagreements by discussion, and if needed, with the study team. We present results for each criterion as high risk, consisting of definitely no and not so much, and low risk, consisting of to a great extent and definitely yes.

To summarize the credibility of the MID, we compiled the 6 criteria rated for each study as high, moderate, and low overall credibility. Studies received a high credibility rating if all criteria were met, including report of satisfactory relation to target instrument or measurement properties of anchor specified (ie, all criteria rated low risk). Studies received a moderate-credibility rating if all criteria were met with the exception of the requirements for a demonstrated satisfactory relation to target and satisfactory demonstrated measurement properties of the anchor. A low-credibility rating was where the study failed to report satisfactory measurement properties of the anchor specified and also failed to meet >1 additional criteria.

This article summarizes MID estimates, along with study design, intervention, population characteristics, characteristics of the PRO, characteristics of the anchor, and credibility ratings.

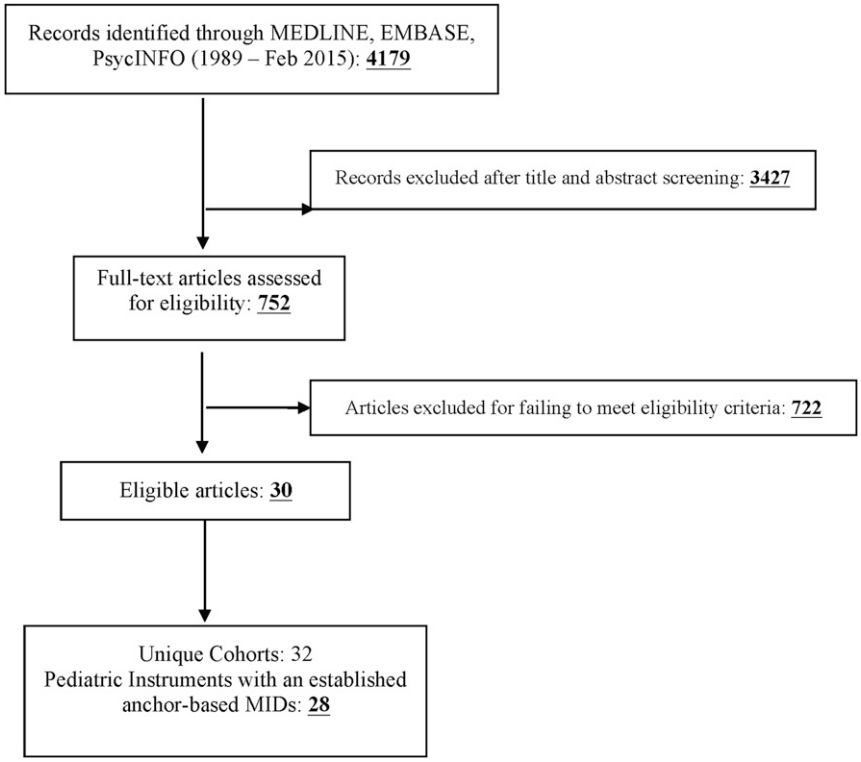

FIGURE 1
Flowchart of eligible studies.

RESULTS

Search Results

Of 4179 candidate citations, 752 were flagged as potentially eligible, of which 30 proved eligible. From the title and abstract screening, we excluded 3427 articles. Studies were excluded for not being a primary study, addressing an adult population, failing to evaluate an MID, not evaluating the MID by using an anchor-based approach, or having health status assessed by a clinician. Our final sample of 30 articles included 32 cohorts and reported on 28 unique PRO instruments; 2 articles reported the MID by using 2 separate datasets (Fig 1).[22,23] The Supplemental References provides a list of included studies.

Study Characteristics

Table 1 provides a summary of study characteristics and Supplemental Table 6 presents the characteristics per study. Among the 32 cohorts,

most (n = 27) were prospective observational studies, half were evaluated in autoimmune or respiratory conditions, and half in North America. All studies were published after 2000 with 9 (28%) published within the past 2 years. Sample size that was used to calculate the MID varied across studies, with a median (interquartile range) of 126 (41–281), and the mean age ranged from 2.8 to 14.6 years.

PRO Instrument Characteristics

Table 2 presents detailed characteristics of the PRO instruments. Of the 28 unique PRO instruments reported across the 32 cohorts, 8 (29%) were generic and 13 (46%) were disease-specific, 5 (18%) were symptom-specific, and 2 (7%) were function-specific. Most (n = 26, 93%) had been previously cited in existing literature; only 2 (7%), the Family Functioning Questionnaire and the Parent Cough-Specific Quality of Life Questionnaire (PC-QoL-8), being used for the first time.[23,24] The number of items in each PRO varied widely, with half having >10 items. Eighteen (64%) of the PROs were measured on the nominal/ordinal scale, with 7 (25%) being measured on the ratio/interval scale and 3 (11%) including elements of both. Ten (36%) of the instruments assessed a single domain, whereas 18 (64%) assessed ≥2 domains, with the most frequently reported domain being physical function (n = 12, 31%). Nine (32%) PROs were administered solely to the pediatric patient, 11 (39%) to the proxy only, and 8 (29%) to both the patient and a proxy.

Anchor Characteristics

Table 3 presents detailed characteristics of the anchors. Of the 51 anchors, 32 (63%) used a single-item instrument with the remaining 13 (25%) including >1 item; 6 (12%) did not present this information. Most of the anchor response options were nominal/ordinal (n = 36, 71%), whereas 7 (14%) were ratio/interval, 2 (4%) contained components of both, and 6 (11%) did not report this information. Forty-two (82%) anchors were limited to a single domain and 9 (18%) addressed ≥2 domains. Twenty (38%) anchors were administered to the pediatric patient only, 22 (42%) to the proxy only, and 9 (17%) to both the patient and the proxy.

TABLE 1 Description of Eligible Cohorts (n = 32)

Variables	Cohorts, n (%)
Publication date	
2001–2003	8 (25)
2004–2006	3 (9)
2007–2009	4 (13)
2010–2012	8 (25)
2013–2014	9 (28)
Country of study origin	
United States	8 (25)
Europe	6 (19)
Canada	6 (19)
Australia	5 (16)
United Kingdom	4 (13)
Asia	3 (9)
Study design	
Clinical/experimental trial	5 (16)
Observational study	27 (84)
Clinical condition under investigation	
Allergy	3 (9)
Autoimmune disorders	7 (22)
Arthritis	5 (16)
Crohn disease	2 (6)
Musculoskeletal disorders	2 (6)
Foot/ankle injury	1 (3)
Scoliosis	1 (3)
Neurologic	3 (9)
Cerebral palsy	2 (6)
Hydrocephalus	1 (3)
Respiratory	7 (22)
Asthma	4 (13)
Chronic cough	3 (9)
Pain	5 (15)
Other	5 (16)
Acne	1 (3)
Otitis media	2 (6)
Sickle cell disease	1 (3)
Healthy population of adolescents	1 (3)
Total no. of participants at baseline	
0–100	13 (41)
100–200	5 (16)
200–300	5 (16)
300–400	3 (9)
400+	5 (16)
Not reported	1 (3)

MID Details

Table 4 provides a compendium of MIDs and the credibility of their estimates. Across the 32 cohorts, 3 studies reported >1 anchor-based method of determining the MID. In total, our review identified 35 distinct MID estimates for the 28 instruments.[24–26] Methods of determining the MID included specifying a change in score or an absolute threshold on the anchor that constituted a minimum improvement or deterioration (n = 33): hypothetical scenarios (n = 2) in which response options were based on a hypothetical response to the PRO instrument (which was also used as the anchor). Of the eligible MID estimates, all but 1 reported the MID as an absolute difference or threshold score rather than as a relative change (percentage of total instrument score).[27] A single article reported boy- and girl-specific estimates.[28] Twenty-three (66%) of the MID estimates included a measure of precision, and 12 (34%) did not.

Credibility Assessment

Table 5 summarizes our credibility assessment and Supplemental Table 7 presents the results of our credibility assessment per cohort. We rated 18 cohorts as having low-credibility MIDs, 13 cohorts as having moderate-credibility MIDs, and 1 cohort as having high credibility.

DISCUSSION

Main Findings

Our study represents the first systematic survey and appraisal of anchor-based MID estimates, and the first comprehensive compendium of such estimates for chronic medical and psychiatric conditions among children. We found 30 studies consisting of 32 cohorts of children reporting 28 unique PRO instruments; >50%

TABLE 2 Characteristics of PRO Instruments

PRO Instrument	Single or Multiple Domains	Domains	Established Validity and Responsiveness	Type of Response Options	PRO Administered to	Study Author, Year of Publication
Generic						
Assessment of Preschool Children's Participation	Multiple	Play, skill development, active physical recreation, social activities	Yes	Ratio/Interval	Parent	Chen, 2013
Childhood Health Assessment Questionnaire	Multiple	Disability, discomfort	Yes	Both	Self- and parent-reported	Brunner, 2005 Dempster, 2001
Family Functioning Questionnaire	Multiple	Family functioning	No	Nominal/Ordinal	Self- and parent-reported	Brouwer, 2007
Functional Status Questionnaire	Multiple	Functioning	Yes	Nominal/Ordinal	Parent	Brouwer, 2007
Kiddo-KINDL, Nepalese version	Multiple	Physical, emotional, self-esteem, family	Yes	Nominal/Ordinal	Self-reported	Yamaguchi, 2010
Quality of my life questionnaire	Multiple	Quality of life	Yes	Both	Self- and parent-reported	Gong, 2007
Rand general health rating index	Multiple	General health	Yes	Nominal/Ordinal	Parent	Brouwer, 2007
TNO-AZL Infant Quality of Life	Multiple	Quality of life	Yes	Nominal/Ordinal	Parent	Brouwer, 2007
Disease-specific						
Acne-Specific Quality of Life Questionnaire	Multiple	Self-perception, emotional, social, symptoms	Yes	Nominal/Ordinal	Self-reported	McLeod, 2003
Asthma Control Questionnaire	Single	Asthma control	Yes	Nominal/Ordinal	Self-reported	Juniper, 2010; Nguyen, 2014
Childhood Asthma Control Test	Single	Symptoms, daily functioning	Yes	Nominal/Ordinal	Self- and parent-reported	Voorend-van Bergen, 2013
Hydrocephalus Outcome Questionnaire	Multiple	Physical, socio-emotional, cognitive	Yes	Nominal/Ordinal	Parent	Kulkarni, 2006
Juvenile Arthritis Disease Activity Score	Multiple	Disease activity, well-being	Yes	Both	Self- and parent-reported	Bulatović Calasan, 2014
Otitis Media-6 Questionnaire	Single	Physical, emotional, quality of life	Yes	Nominal/Ordinal	Parent	Heidemann, 2013; Brouwer, 2007
Pediatric Allergic Disease Quality of Life Questionnaire	Multiple	Practical, symptoms, emotional	Yes	Nominal/Ordinal	Self-reported	Roberts, 2003; Roberts, 2005
PC-QoL-8	Multiple	Psychological, physical, and social functioning	No	Nominal/Ordinal	Parent	Newcombe, 2011; Newcombe, 2013
PC-QoL-27	Multiple	Psychological, physical, and social functioning	Yes	Nominal/Ordinal	Parent	Newcombe, 2013
Pediatric Crohn Disease Activity Index	Single	Disease activity	Yes	Ratio/Interval	Self- and parent-reported	Kundhal, 2003
Rhinoconjuctivitis Total Symptom Score	Single	Symptoms	Yes	Nominal/Ordinal	Self-reported	Devillier 2014

TABLE 2 Continued

PRO Instrument	Single or Multiple Domains	Domains	Established Validity and Responsiveness	Type of Response Options	PRO Administered to	Study Author, Year of Publication
Scoliosis Research Society-22	Multiple	Appearance, activity, pain, mental, satisfaction	Yes	Nominal/Ordinal	Self-reported	Carreon, 2010
Test for Respiratory and Asthma Control in Kids	Multiple	Impairment, risk	Yes	Nominal/Ordinal	Parent	Zeiger, 2011
Function-specific						
Oxford Ankle Foot Questionnaire	Multiple	Physical, emotional, school, play	Yes	Nominal/Ordinal	Self- and parent-reported	Morris, 2009
Pediatric Motor Activity Log	Multiple	Amount and quality of hand use	Yes	Nominal/Ordinal	Parent	Lin, 2012
Symptom-specific						
21-point VAS-Pain	Single	Disease activity, pain	Yes	Ratio/Interval	Parent	Filocamo, 2010
Color Analog Scale-Pain	Single	Pain	Yes	Ratio/Interval	Self-reported	Bulloch, 2002; McConahay, 2007
Faces Pain Scale	Single	Pain	Yes	Nominal/Ordinal	Self-reported	Bulloch, 2002
Numerical Rating Scale-Pain	Single	Pain	Yes	Ratio/Interval	Self-reported	Myrvik, 2013; Voepel-Lewis, 2011; Brouwer, 2007
VAS-Pain	Single	Pain	Yes	Ratio/Interval	Self- and parent-reported	Powell, 2001; Kelly, 2001; Myrvik, 2013

Studies are listed in the Supplemental References.

were published in the past 5 years. Among the instruments identified, more than half used a PRO that was administered directly to patients; however, it was most common for proxies to respond to both the PRO instrument and the anchor. We found that more than two-thirds of questionnaires were specific instruments as opposed to generic, with nearly half being disease-specific PROs.

Our systematic survey draws attention to the methodological issues and challenges involved in MID determinations. We classified more than half of the studies as low credibility with only 1 regarded as high quality. In particular, for all but 4 instruments, investigators neglected to either report a satisfactory relation between the PRO instrument and the anchor or they did not report satisfactory measurement properties of the anchor. Although it is promising to see the increasing frequency of MID ascertained in child and adolescent populations, our findings highlight the limitations in establishing highly credible estimates. Our credibility assessment tool helps with the transparency of this issue to readers.

The interpretation or applicability of MIDs from the same instrument across 2 different populations may differ. For example, a 2001 study in our sample assessed the MID of the VAS in patients with acute pain of either traumatic or nontraumatic cause,[29] whereas a 2013 study assessed the MID of the VAS in patients with sickle cell disease.[30] The former ascertained an MID of +11 mm for improvement whereas the latter ascertained 9.7 mm for improvement. Readers may be confused as to which one is more accurate when they see 2 (or more) estimates or whether the difference can be explained by chance. Thus, it is crucial to provide readers with MID estimates for the population to which one intends to apply the estimates.

TABLE 3 Characteristics of Anchor Instruments

Anchor Instrument	No. of Items or Questions	Response Options	No. of Domains	Domain List	Who Reported	Corresponding PRO	Study Using Anchor/PRO
ACRPedi30 for disease improvement	Not reported	Not reported	Multiple	Disease improvement	Parent	Juvenile Arthritis Disease Activity Score	Bulatović Calasan, 2014
AOM frequency (AOM episodes/child)	1	Ratio/Interval	Single	Indices of disease	Self-reported	Rand general health rating index Functional Status Questionnaire TNO-AZL Infant Quality of Life Otitis Media-6 Numerical rating scales Family Functioning Questionnaire	Brouwer, 2007
AOM Severity	3	Nominal/Ordinal	Single	Severity of disease	Self-reported	Rand general health rating index Functional Status Questionnaire TNO-AZL Infant Quality of Life Otitis Media-6 Numerical rating scales Family Functioning Questionnaire	Brouwer, 2007
Asthma Control Test	5	Nominal/Ordinal	Single	Asthma control	Self-reported	Asthma Control Questionnaire	Nguyen, 2014
Caregiver report of change in respiratory symptom status	1	Nominal/Ordinal	Single	Symptoms	Parent	Test for Respiratory and Asthma Control in Kids	Zeiger, 2011
Categorical Rating of Pain	1	Nominal/Ordinal	Single	Pain	Self-reported	VAS	Kelly, 2001
Center for Epidemiologic Studies Depression Scale	20	Nominal/Ordinal	Single	Depression	Self-reported	Kiddo-KINDL	Yamaguchi, 2010
Change in respiratory control status-based physical guidelines	1	Nominal/Ordinal	Single	Symptoms	Parent	Test for Respiratory and Asthma Control in Kids	Zeiger, 2011
Childhood Asthma Control Test	7	Nominal/Ordinal	Single	Asthma control	Self- and parent-reported	Asthma Control Questionnaire	Nguyen, 2014
Childhood Health Assessment Questionnaire	Not reported	Not reported	Multiple	Disability, discomfort	Self-, parent-, and clinician-reported	Juvenile Arthritis Disease Activity Score	Bulatović Calasan, 2014
Child: Global Change of Well-being, categorical	1	Nominal/Ordinal	Single	Well-being	Self-reported	Childhood Health Assessment Questionnaire	Brunner, 2005
Child: Global Change of Well-being, linear	1	Ratio/Interval	Single	Well-being, disease activity	Self-reported	Childhood Health Assessment Questionnaire	Brunner, 2005
Child: Global Change of Well-being, Likert	1	Nominal/Ordinal	Single	Well-being	Self-reported	Childhood Health Assessment Questionnaire	Brunner, 2005
Clinician's Global Rating of Change Questionnaire	1	Nominal/Ordinal	Single	Overall health	Clinician	Asthma Control Questionnaire	Juniper, 2010
Cough Verbal Category Descriptive Score	1	Nominal/Ordinal	Single	Interference to usual activities	Parent	PC-QoL Questionnaire	Newcombe, 2011
Crohn Disease Activity Index	Not reported	Ratio/Interval	Single	Disease activity	Self- and other proxy-reported	Pediatric Crohn Disease Activity Index	Kundhal, 2003
CRV	1	Not reported	Single	Well-being	Clinician	Childhood Health Assessment Questionnaire	Brunner, 2005
Dichotomized Clinician's Global Rating of Range Questionnaire	1	Nominal/Ordinal	Single	Disease activity	Clinician	Pediatric Crohn Disease Activity Index	Kundhal, 2003

TABLE 3 Continued

Anchor Instrument	No. of Items or Questions	Response Options	No. of Domains	Domain List	Who Reported	Corresponding PRO	Study Using Anchor/PRO
ESR	Not reported	Not reported	Single	Hematology test	Self-, parent-, and clinician-reported	Juvenile Arthritis Disease Activity Score	Bulatović Calasan, 2014
Facial Acne Global Assessment	3	Nominal/Ordinal	Single	Appearance	Clinician	Acne-Specific Quality of Life Questionnaire	McLeod, 2003
Functional Independence Measure in Children	18	Nominal/Ordinal	Multiple	Self-care, mobility, cognition	Self- and other proxy-reported	Assessment of Preschool Children's Participation Pediatric Motor Activity Log	Chen, 2013; Lin, 2012
Frequency of episodes or attacks	1	Nominal/Ordinal	Single	Frequency of symptoms	Parent	Test for Respiratory and Asthma Control in Kids	Zeiger, 2011
Frequency of oral corticosteroid use	1	Nominal/Ordinal	Single	Medication use	Parent	Test for Respiratory and Asthma Control in Kids	Zeiger, 2011
Global Assessment of Pain Relief	1	Nominal/Ordinal	Single	Pain	Self-reported	VAS	Myrvik, 2013
Global Health Rating	1	Nominal/Ordinal	Single	Overall health	Parent- and clinician-reported	Numerical Rating Scale Hydrocephalus Outcome Questionnaire	Kulkarni, 2006
Global Rating of Change Scale	15	Nominal/Ordinal	Single	Symptoms	Self-reported	Rhinoconjuctivitis Total Symptom Score	Devillier, 2014
Global Rating of disease course	1	Nominal/Ordinal	Single	Disease course	Parent- and clinician-reported	21-point VAS	Filocamo, 2010
Global Rating of Pain	1	Nominal/Ordinal	Single	Pain	Self-reported	Color Analog Scale Faces of Pain Scale	Bulloch, 2002
Global Rating of problem	1	Nominal/Ordinal	Single	Global rating	Self- and parent-reported	Oxford Ankle Foot Questionnaire	Morris, 2009
Global Perceived Effect	1	Nominal/Ordinal	Single	Quality of life	Parent	Otitis Media-6 Questionnaire	Heidemann, 2013
Pain Relief Indicator	1	Nominal/Ordinal	Single	Pain	Self-reported	Numerical Rating Scale	Voepel-Lewis, 2011
Parent: Global Change of Well-being, categorical	1	Nominal/Ordinal	Single	Well-being	Parent	Childhood Health Assessment Questionnaire	Brunner, 2005
Parent: Global Change of Well-being, linear	1	Ratio/Interval	Single	Well-being	Parent	Childhood Health Assessment Questionnaire	Brunner, 2005
Parent: Global Change of Well-being, Likert	1	Nominal/Ordinal	Single	Well-being	Parent	Childhood Health Assessment Questionnaire	Brunner, 2005
Parent/Patient Assessment of Childhood Well-being	Not reported	Not reported	Single	Disability, discomfort	Self-, parent-, and clinician-reported	Juvenile Arthritis Disease Activity Score	Bulatović Calasan, 2014
Pediatric Asthma Caregivers Quality of Life	23	Nominal/Ordinal	Multiple	Emotions, activity, symptoms	Parent- and guardian-reported	Childhood Asthma Control Test	Voorend-van Bergen, 2013
Pediatric Asthma Quality of Life Questionnaire	23	Nominal/Ordinal	Multiple	Quality of life	Self-reported	Childhood Asthma Control Test	Voorend-van Bergen, 2013
Physician: Global Change of Well-being, linear	1	Ratio/Interval	Single	Well-being	Clinician	Childhood Health Assessment Questionnaire	Brunner, 2005

TABLE 3 Continued

Anchor Instrument	No. of Items or Questions	Response Options	No. of Domains	Domain List	Who Reported	Corresponding PRO	Study Using Anchor/PRO
Physician: Global Change of Well-being, categorical	1	Nominal/Ordinal	Single	Well-being	Clinician	Childhood Health Assessment Questionnaire	Brunner, 2005
Physician Guidelines-based control table rating	1	Nominal/Ordinal	Single	Control of disease	Self- and clinician-reported	Test for Respiratory and Asthma Control in Kids	Zeiger, 2011
Physician's therapy recommendations	1	Nominal/Ordinal	Single	Therapy recommended	Self- and clinician-reported	Test for Respiratory and Asthma Control in Kids	Zeiger, 2011
Responses to hypothetical scenarios on Quality of My Life PRO	3	Both	Multiple	Quality of life	Self- and parent-reported	Quality of my life questionnaire	Gong, 2007
Responses to hypothetical scenarios on Childhood Health Assessment Questionnaire PRO	2	Ratio/Interval	Multiple	Disability, discomfort	Parent	Childhood Health Assessment Questionnaire	Dempster, 2001
Response to Likert scale about pain	1	Nominal/Ordinal	Single	Pain	Self-reported	Color Analog Scale	McConahay, 2007
Rhinoconjunctivitis Quality of Life Questionnaire	Not reported	Not reported	Multiple	Quality of life	Self-reported	Rhinoconjuctivitis Total Symptom Score	Devillier, 2014
Scoliosis Research Society-30 (last 8 questions), SAQ items 26, 32	10	Nominal/Ordinal	Multiple	Appearance, activity, pain	Self-reported	Scoliosis Research Society-22	Carreon, 2010
Subject Self-Assessment of Facial Acne	1	Nominal/Ordinal	Single	Appearance	Self-reported	Acne-Specific Quality of Life Questionnaire	McLeod, 2003
VAS for Asthma or Hayfever	1	Ratio/Interval	Single	Trouble with disease	Self-reported	Pediatric Allergic Disease Quality of Life Questionnaire	Roberts, 2005
VAS for how troubled they were about asthma, hayfever, allergies	3	Ratio/Interval	Single	Trouble with disease	Self-reported	Pediatric Allergic Disease Quality of Life Questionnaire	Roberts, 2003
Verbal Categorical Descriptive	1	Nominal/Ordinal	Single	Trouble with disease	Parent	PC-QoL Questionnaire-27 PC-QoL Questionnaire-8	Newcombe, 2013
Verbal Categorical Rating of "heaps better" to "heaps worse"	1	Nominal/Ordinal	Single	Pain	Self-reported	VAS Pain Score	Powell, 2001

Studies are listed in the Supplemental References. AOM, acute otitis media; CRV, core response variable; ESR, erythrocyte sedimentation rate; PRO, patient reported outcome; SAQ, scoliosis appearance questionnaire.

One needs to be also cautious when interpreting these findings to similar clinical populations but where other factors, such as the region in which the study was conducted, could result in different manifestations or interpretations of the disease.

Developing an estimate of the MIDs is needed for interpreting the magnitude of improvement or deterioration of PROs relevant to children; however, the measurement and interpretation in this population poses additional challenges to that of an adult population. First, what is considered to be minimally important can vary depending on if it is the patient (child) or a proxy (eg, parent) answering the question; sometimes a proxy report may be misleading,[31] whereas in other instances, a proxy report may be a reasonable estimate of patient status.[32] In our study, we thus restricted proxies to include only parents/guardians. Second, previous authors have speculated that patients presenting with different baseline scores may have different ratings or interpretations of a minimal change. For example, a child with severe functional impairment or pain may require a smaller change in rating to be considered meaningful as opposed to one who has a low degree of impairment or pain. However, a recent study demonstrated that the MID, on average, does not change relative to the baseline scores.[33] As well, the reader should note that Food and Drug Administration guidance does not use the term MID but rather emphasizes establishing meaningful change in PRO measures at the individual level (ie, defined as a responder) versus at the treatment group level. The Food and Drug Administration defines a responder threshold as "a score change in a measure, experienced by an individual patient over a predetermined time period that has been demonstrated in the target population to have a significant

TABLE 4 Characteristics of MID

PRO	Type of Measure	Range of Scores	MID Estimates on the PRO	Credibility of Estimates by Study
21-point VAS-Pain	Symptom	0 to 21	Physician (Slightly Improved, Worsened): Global: (−)0.71, (+)1.68 Pain: (−)2.20, (+)1.87 CI: Parent global: (−)1.77 to (+) 0.34 and (+)0.74 to (+)2.63 Parent pain: (−)3.58 to (−) 0.82 and (+)0.78 to (+)2.95 Parent (Slightly Improved, Worsened): Parent global: (−)0.64, (+)1.61 Parent pain: (−)1.09, (+)2.26 CI: Parent global: (−)1.60 to (+)0.33 and (+) 0.80 to (+)2.41 Parent pain: (−)2.45 to (−) 0.17 and (+) 0.75 to (+) 3.78	Low (Filocamo, 2010)
Acne-Specific Quality of Life Questionnaire	Disease	0 to 100	Subject Self-Assessment of Facial Acne: Self-perception: (+)5.15 Role-emotional: (+)4.73 Role-social: (+)3.08 Acne symptoms: (+)4.62 SD Self-perception: (+)7.6 Role-emotional: (+)7.4 Role-social: (+)5.5 Acne symptoms: (+)5.0 Facial Acne Global Assessment: Self-perception: (+)3.36 Role-emotional: (+)3.86 Role-social: (+)2.07 Acne symptoms: (+)3.59	Moderate (McLeod, 2003)
Assessment of Preschool Children's Participation	Generic	Not reported	Diversity: PA: (+)16.7%, SD: (+)19.4%, AP: (+)11.0%, SA: (+)16.5% Total: (+)16.3% Intensity: PA: (+)1.1, SD: (+)1.2, AP: (+)0.8, SA: (+)0.9 Total: (+)1.0	Low (Chen, 2013)
Asthma Control Questionnaire	Disease	0 to 6	Juniper 2010 Singer-Guyatt anchor-based method: (+)0.52 -symptoms alone: (+)0.65 -symptoms plus FEV1%pred: (+)0.52 -symptoms plus SABA use: (+)0.63 SEM: (+)0.45 (total) Geometric regression method: (+)0.5 SEM: (+)0.05 Nguyen 2014 6 to 11 y: (+)0.33 12 to 17 y: (+)0.42 Overall: (+)0.40	Low (Juniper, 2010); Moderate (Nguyen, 2014)
Childhood Asthma Control Test	Disease	0 to 27	(+)1.6 CI: (+)1.1 to (+)2.1	Moderate (Voorend-van Bergen, 2013)

TABLE 4 Continued

PRO	Type of Measure	Range of Scores	MID Estimates on the PRO	Credibility of Estimates by Study
Childhood Health Assessment Questionnaire	Generic	0 to 3	Brunner 2005 Worsening: 1. (+)0.063 2. (−)0.167 3. (+)0.25 4. (−)0.113 5. 0 6. (+)0.238 7. (−)0.2 8. (−)0.231 9. (−)0.102 Improvement 1. (−)0.016 2. (−)0.205 3. (−)0.188 4. (+)0.125 5. (−)0.272 6. (−)0.023 7. (−)0.118 8. (−)0.170 0. (−)0.115 SD Worsening: 1. (+)0.217 2. (+)0.395 3. (+)0.375 4. (+)0.579 5. (+)0.702 6. (+)0.419 7. (+)0.736 8. (+) 0.618 9. (+)0.530 Improvement 1. (+)0.766 2. (+)0.642 3. (+)0.418 4. (+)0.407 5. (+)0.688 6. (+)0.273 7. (+)0.381 8. (+)0.459 9. (+)0.662 Dempster 2001 Improvement: (−)0.13 Deterioration: (+)0.75	Moderate (Brunner, 2005); Moderate (Dempster, 2001)
Color Analog Scale-Pain	Symptom	0 to 10	Bulloch 2002 (+)1.7cm CI: (+)1.1 to (+)2.3 McConahay 2007 (−)2.40cm CI: (−)3.15 to (−) 1.72cm	Low (Bulloch, 2002); Moderate (McConahay, 2007)
Faces Pain Scale	Symptom	0 to 6	(+)1.0 face IQR: (+)1 to (+)2	Low (Bulloch, 2002)
Family Functioning Questionnaire	Generic	0 to 100	AOM frequency: 3–10 points change AOM Severity: 2–10 points change	Low (Brouwer, 2007)
Functional Status Questionnaire	Generic	0 to 100	AOM frequency: 3–10 points change AOM Severity: 2–10 points change	Low (Brouwer, 2007)

TABLE 4 Continued

PRO	Type of Measure	Range of Scores	MID Estimates on the PRO	Credibility of Estimates by Study
Hydrocephalus Outcome Questionnaire	Disease	0 to 1.0	Using Mothers' Responses (+) 0.10 CI: (+) 0.07 to (+)0.20 Using Surgeons' Responses (+)0.12 CI: (+)0.04 to (+)0.21	Low (Kulkarni, 2006)
Juvenile Arthritis Disease Activity Score	Disease	0 to 57	Using ACRPedi30 for MID improvement: (−)5.5 IQR: (−)9.5 to (−)2.7 Using flare for MID deterioration: (+)1.7 IQR: (+)0.3 to (+)0.5	Low (Bulatović Calasan, 2014)
Kiddo-KINDL (Nepalese version)	Generic	0 to 100	(+)4 SD: (+)2.9 95% CI: (−)5.16 to (−)2.76	
Numerical Rating Scale-Pain	Symptom	0 to 10	Myrvik 2013 (+)0.9 SES: (+)0.24 Voepel-Lewis 2011 Improvement: (−)1 Deterioration: (+)1 CI Improvement: (−)0.5 to (+)1 CI Deterioration: (+)0.5 to (+)2.7 Brouwer 2007 [on 0 to 100] AOM Frequency: 5–15 points change, except during 0–7 mo interval with 29 points change AOM Severity: 4–8 points, except for Child with 16 and 17 points change	Moderate (Myrvik, 2013); Low (Voepel-Lewis, 2011); Low (Brouwer, 2007)
Otitis Media-6 Questionnaire	Disease	0 to 100	Heidemann 2013 FHS: (+)16.7, (+)22.2 (with ROC) NRS-child: (+)30.0 Brouwer 2007 AOM frequency: 5–15 points change AOM Severity: 4–8 points	Low (Heidemann, 2013); Low (Brouwer, 2007)
Oxford Ankle Foot Questionnaire	Function	0 to 100	Physical: (+)10.0, (+)17.3 School and Play: (+)8.7, (+)6.6 Emotional: (+)8.3, (+)6.9	Moderate (Morris, 2009)
Pediatric Allergic Disease Quality of Life Questionnaire	Disease	0 to 6	Roberts 2003 (+)0.33 CI: (+)0.11 to (+)0.54 Roberts 2005 (+)0.20 CI: (−)0.09 to (+)0.49	Moderate (Roberts, 2003); Moderate (Roberts, 2005)

TABLE 4 Continued

PRO	Type of Measure	Range of Scores	MID Estimates on the PRO	Credibility of Estimates by Study
PC-QoL-27	Disease	1 to 7	Newcombe 2011 Overall: (+)0.9 Social: (+)0.71 Physical: (+)0.99 Psychological: (+) 0.95 SD Overall: (+)0.66 Social: (+)0.83 Physical: (+)0.62 Psychological: (+)0.82 Newcombe 2013 A (+) 0.90 SD: (+)0.66 Newcombe 2013 B (+)1.42 SD: (+)1.07	Low (Newcombe, 2011); Low (Newcombe, 2013)
PC-QoL-8	Disease	1 to 7	Newcombe 2013 A (+)0.83 SD: (+)0.53 Newcombe 2013 B (+)1.42 SD: (+) 1.10	Low (Newcombe, 2013)
Pediatric Crohn Disease Activity Index	Disease	0 to 100	Kundhal 2003 A (−)12.5 points Kundhal 2003 B (−10) points	Low (Kundhal, 2003)
PMAL	Function	0 to 6	PMAL-AOU (+)0.94 PMAL-QOM (+)0.74	Low (Lin, 2012)
Quality of my life questionnaire	Generic	0 to 100	QoL: Improvement: (+) 7 mm Deterioration: (−)33 mm HRQOL: Improvement: (+)11 mm Deterioration: (−)38 mm	Low (Gong, 2007)
Rand general health rating index	Generic	0 to 100	AOM frequency: 3–10 points change AOM Severity: 2–10 points change	Low (Brouwer, 2007)
Rhinoconjuctivitis Total Symptom Score	Disease	NR	Using Global Rating of Change as Anchor ROC Curve Children: (+)1.28 Adolescents: (+)1.14 ROC Curve 95% CI Children: (+)0.64–0.78 Adolescents: (+)0.72–0.86 Using Rhinoconjunctivitis Quality of Life Questionnaire as Anchor: ROC Curve Children: (+)1.62 Adolescents: (+)0.88	Low (Devillier, 2014)
Scoliosis Research Society-22	Disease	1 to 5	Pain: (+)0.20 Activity: (+)0.08 Appearance: (+)0.98	High (Carreon, 2010)
Test for Respiratory and Asthma Control in Kids	Disease	0 to 100	(+)11 Range: Greater than or equal to (+)10	Low (Zeiger, 2011)
TNO-AZL Infant Quality of Life	Generic	0 to 100	AOM frequency: 3–10 points change AOM Severity: 2–10 points change	Low (Brouwer, 2007)

TABLE 4 Continued

PRO	Type of Measure	Range of Scores	MID Estimates on the PRO	Credibility of Estimates by Study
VAS-Pain	Symptom	0 to 100	Powell 2001 Improvement: (+)11 mm Deterioration: (+)8 mm Overall: (+)10 mm IQR: Improvement: (+)15 to (+)19 mm Deterioration: 0 to (−)14 mm Overall: (+)3 to (+)19 mm CI: Overall: (+)7 to (+)12 mm Kelly 2001 (+)24 CI: (+)17 to (+)31 Myrvik 2013 (+)0.97 cm SE: (+)0.27 cm	Moderate (Powell, 2001); Moderate (Kelly, 2001); Moderate (Myrvik, 2013)

Studies are listed in Supplemental References. ACRPedi30, American College of Rheumatology Pediatric 30; AOM, acute otitis media; AOU, amount of use; AP, active physical recreation; CI, confidence interval; HRQOL, health-related quality of life; IQR, interquartile range; PA, areas of play; PMAL, Pediatric Motor Activity Log; QoL, quality of life; QOM, quality of movement; ROC, receiver operating characteristic; SA, social activities.

TABLE 5 Credibility Assessment (n = 32)

Criterion	n (%)
1. Did the patients respond directly to the PRO and the anchor?	
High risk	12 (38)
Low risk	20 (62)
2. Was an independent standard of comparison used for an anchor?	
High risk	7 (22)
Low risk	25 (78)
3. Is the anchor itself interpretable for the patients?	
High risk	3 (9)
Low risk	29 (91)
4. Is the anchor itself interpretable for the clinicians?	
High risk	1 (3)
Low risk	31 (97)
5. Is the anchor construct empirically sufficiently closely related to the target instrument (eg, correlation coefficient >0.5)?	
High risk	24 (75)
Low risk	8 (25)
6. Have the authors reported measurement properties of the anchor?	
High risk	24 (75)
Low risk	8 (25)

treatment benefit."[34] This guidance may have implications on different interpretations in the magnitude of treatment effects for PROs developed for child health.

Strengths and Limitations

The strengths of our study include a comprehensive and transparent search strategy and independent eligibility assessment and data extraction. We used a methodological quality appraisal method to assess MID determination across 6 criteria, allowing us to classify MIDs as high, moderate, and low credibility.

Our study has limitations, one of which is that although our credibility assessment instrument consisting of 6 criteria was developed to provide transparency to the credibility of the MID, we have not yet addressed the measurement properties of our instrument to measure credibility. The development of the instrument was, however, informed by a review of the relevant literature and our own extensive experience with the generation of MIDs.[17–21] We are currently addressing the validity of the instrument as part of an ongoing larger study.

Another limitation is the restriction of our search to 3 electronic sources: Medline, Embase, and PsycINFO. There is a possibility that other databases, such as CINAHL, may have included original studies of MID estimates in children that were not indexed in the databases we searched. However, given that our focus was chronic medical and psychiatric conditions, it is likely that our selected databases effectively capture most, if not all, studies. Last, we did not conduct duplicate data abstraction and thus there is the possibility of errors in data extracted. However, to mitigate this risk, 1 reviewer checked all data and made corrections where necessary.

CONCLUSIONS

Our systematic presentation of all anchor-based MID estimates relevant to children will help promote informed decision-making by allowing clinical trialists, systematic review authors, guideline developers, and clinicians to better interpret the magnitude (size) of treatment effects for PROs. If, for example, we have identified a credible MID of 1.0 on an 11-point pain instrument, a mean difference

of 2.0 is twice the MID and likely considered a moderate to large treatment effect. Further, our work allows knowledge users to identify anchor-based MIDs that are more or less credible. Our credibility instrument, currently in the process of being validated, also provides guidance for subsequent work in developing, further establishing, or confirming MID estimates for PROs among the pediatric population.

ACKNOWLEDGMENTS

We thank Ms Tamsin Adams-Webber at the Hospital for Sick Children and Dr Paul Alexander for their assistance with developing the initial literature search.

ABBREVIATIONS

MID: minimally important difference
PC-QoL: Parent Cough-Specific Quality of Life Questionnaire
PRO: patient-reported outcome
VAS: visual analog scale

Massachusetts; [e]Department of Medicine, University of Toronto, Toronto, Ontario, Canada; [g]Evidence-Based Dentistry Unit, Faculty of Dentistry, Universidad de Chile, Independencia, Santiago, Chile; [h]Department of Pediatrics, Santa Maria Hospital, Lisboa, Portugal; [j]Nephrology Program, Humber River Regional Hospital, Toronto, Ontario, Canada; [k]Division of Nephrology, University of Western Ontario, London, Ontario, Canada; [l]Li Ka Shing Knowledge Institute, St. Michael's Hospital, Toronto, Ontario, Canada; and [m]Institute of Health Policy, Management and Evaluation, Dalla Lana School of Public Health, University of Toronto, Toronto, Ontario, Canada

Dr Ebrahim conceptualized and designed the study, was involved in the study selection and extraction process of the study and analysis, and drafted the initial manuscript; Ms Vercammen, Ms Sivanand, and Dr Carrasco-Labra were involved in the study selection and extraction process of the study; Drs Fernandes and Crawford reviewed and revised the manuscript; Drs Guyatt, Nersrallah, and Johnston conceptualized and designed the study and reviewed and revised the manuscript; and all authors approved the final manuscript to be submitted.

DOI: 10.1542/peds.2016-0833

Accepted for publication Nov 28, 2016

Address correspondence to Bradley Johnston, PhD, Systematic Overviews through advancing Research Technology (SORT), Department of Anaesthesia and Pain Medicine, The Hospital for Sick Children, University of Toronto, 686 Bay St, Toronto, Ontario, Canada M5G 0A4. E-mail: bradley.johnston@sickkids.ca

PEDIATRICS (ISSN Numbers: Print, 0031-4005; Online, 1098-4275).

FINANCIAL DISCLOSURE: The authors have indicated they have no financial relationships relevant to this article to disclose.

FUNDING: This project is funded by the Canadian Institutes of Health Research, Knowledge Synthesis grant DC0190SR. Dr Ebrahim was supported by a MITACS Elevate and SickKids Restracomp Postdoctoral Fellowship Award.

POTENTIAL CONFLICT OF INTEREST: The authors have indicated they have no potential conflicts of interest to disclose.

REFERENCES

1. Johnston BC, Ebrahim S, Carrasco-Labra A, et al. Minimally important difference estimates and methods: a protocol. *BMJ Open*. 2015;5(10):e007953
2. Schünemann HJ, Guyatt GH. Commentary—goodbye M(C)ID! Hello MID, where do you come from? *Health Serv Res*. 2005;40(2):593–597
3. Methodology Committee of the Patient-Centered Outcomes Research Institute (PCORI). Methodological standards and patient-centeredness in comparative effectiveness research: the PCORI perspective. *JAMA*. 2012;307(15):1636–1640
4. Guyatt GH, Osoba D, Wu AW, Wyrwich KW, Norman GR; Clinical Significance Consensus Meeting Group. Methods to explain the clinical significance of health status measures. *Mayo Clin Proc*. 2002;77(4):371–383
5. Jaeschke R, Guyatt GH, Keller J, Singer J. Interpreting changes in quality-of-life score in N of 1 randomized trials. *Control Clin Trials*. 1991;12(suppl 4):226S–233S
6. King MT. A point of minimal important difference (MID): a critique of terminology and methods. *Expert Rev Pharmacoecon Outcomes Res*. 2011;11(2):171–184
7. Morse J. Designing funded qualitative research. In: Denzin NK, Lincoln YS, eds. *Handbook for Qualitative Research*. Thousand Oaks, CA: Sage; 1994:220–235
8. Fallah A, Akl EA, Ebrahim S, et al. Anterior cervical discectomy with arthroplasty versus arthrodesis for single-level cervical spondylosis: a systematic review and meta-analysis. *PLoS One*. 2012;7(8):e43407
9. Turner D, Schünemann HJ, Griffith LE, et al. The minimal detectable change cannot reliably replace the minimal important difference. *J Clin Epidemiol*. 2010;63(1):28–36
10. Turner D, Schünemann HJ, Griffith LE, et al. Using the entire cohort in the receiver operating characteristic analysis maximizes precision of the minimal important difference. *J Clin Epidemiol*. 2009;62(4):374–379
11. Juniper EF, Guyatt GH, Willan A, Griffith LE. Determining a minimal important change in a disease-specific quality of life questionnaire. *J Clin Epidemiol*. 1994;47(1):81–87
12. Johnston BC, Thorlund K, Schünemann HJ, et al. Improving the interpretation of quality of life evidence in meta-analyses: the application of minimal important difference units. *Health Qual Life Outcomes*. 2010;8:116
13. Johnston BC, Thorlund K, da Costa BR, Furukawa TA, Guyatt GH. New methods can extend the use of

minimal important difference units in meta-analyses of continuous outcome measures. *J Clin Epidemiol*. 2012;65(8):817–826

14. Jaeschke R, Singer J, Guyatt GH. Measurement of health status. Ascertaining the minimal clinically important difference. *Control Clin Trials*. 1989;10(4):407–415
15. Yalcin I, Patrick DL, Summers K, Kinchen K, Bump RC. Minimal clinically important differences in Incontinence Quality-of-Life scores in stress urinary incontinence. *Urology*. 2006;67(6):1304–1308
16. Levine M, Ioannidis J, Haines T, Guyatt G. Harm (observational studies). In: Guyatt GRD, Meade M, Cook D, eds. *Users' Guides to the Medical Literature: A Manual for Evidence-Based Clinical Practice*. New York, NY: McGraw-Hill; 2008
17. Randolph A, Cook DJ, Guyatt G. Prognosis. In: Guyatt GRD, Meade M, Cook D, eds. *Users' Guides to the Medical Literature: A Manual for Evidence-Based Clinical Practice*. New York, NY: McGraw-Hill; 2008
18. Furukawa T, Jaeschke J, Cook DJ, Jaeschke R, Guyatt G. Measuring patients' experience. In: Guyatt GRD, Meade M, Cook D, eds. *Users' Guides to the Medical Literature: A Manual for Evidence-Based Clinical Practice*. New York, NY: McGraw-Hill; 2008
19. Sun X, Briel M, Walter SD, Guyatt GH. Is a subgroup effect believable? Updating criteria to evaluate the credibility of subgroup analyses. *BMJ*. 2010;340:c117
20. Akl EA, Sun X, Busse JW, et al. Specific instructions for estimating unclearly reported blinding status in randomized trials were reliable and valid. *J Clin Epidemiol*. 2012;65(3):262–267
21. Kundhal PS, Critch JN, Zachos M, Otley AR, Stephens D, Griffiths AM. Pediatric Crohn Disease Activity Index: responsive to short-term change. *J Pediatr Gastroenterol Nutr*. 2003;36(1):83–89
22. Newcombe PA, Sheffield JK, Chang AB. Parent cough-specific quality of life: development and validation of a short form. *J Allergy Clin Immunol*. 2013;131(4):1069–1074
23. Brouwer CN, Schilder AG, van Stel HF, et al. Reliability and validity of functional health status and health-related quality of life questionnaires in children with recurrent acute otitis media. *Qual Life Res*. 2007;16(8):1357–1373
24. Bulatović Calasan M, de Vries LD, Vastert SJ, Heijstek MW, Wulffraat NM. Interpretation of the Juvenile Arthritis Disease Activity Score: responsiveness, clinically important differences and levels of disease activity in prospective cohorts of patients with juvenile idiopathic arthritis. *Rheumatology (Oxford)*. 2014;53(2):307–312
25. Devillier P, Chassany O, Vicaut E, et al. The minimally important difference in the Rhinoconjunctivitis Total Symptom Score in grass-pollen-induced allergic rhinoconjunctivitis. *Allergy*. 2014;69(12):1689–1695
26. McLeod LD, Fehnel SE, Brandman J, Symonds T. Evaluating minimal clinically important differences for the acne-specific quality of life questionnaire. *Pharmacoeconomics*. 2003;21(15):1069–1079
27. Chen CL, Shen IH, Chen CY, Wu CY, Liu WY, Chung CY. Validity, responsiveness, minimal detectable change, and minimal clinically important change of Pediatric Balance Scale in children with cerebral palsy. *Res Dev Disabil*. 2013;34(3):916–922
28. Voepel-Lewis T, Burke CN, Jeffreys N, Malviya S, Tait AR. Do 0–10 numeric rating scores translate into clinically meaningful pain measures for children? *Anesth Analg*. 2011;112(2):415–421
29. Powell CV, Kelly AM, Williams A. Determining the minimum clinically significant difference in visual analog pain score for children. *Ann Emerg Med*. 2001;37(1):28–31
30. Myrvik MP, Brandow AM, Drendel AL, Yan K, Hoffmann RG, Panepinto JA. Clinically meaningful measurement of pain in children with sickle cell disease. *Pediatr Blood Cancer*. 2013;60(10):1689–1695
31. Schel SH, Bouman YH, Bulten BH. Quality of life in long-term forensic psychiatric care: comparison of self-report and proxy assessments. *Arch Psychiatr Nurs*. 2015;29(3):162–167
32. Kim EJ, Song DH, Kim SJ, et al. Proxy and patients ratings on quality of life in patients with schizophrenia and bipolar disorder in Korea. *Qual Life Res*. 2010;19(4):521–529
33. Zhang Y, Zhang S, Thabane L, Furukawa TA, Johnston BC, Guyatt GH. Although not consistently superior, the absolute approach to framing the minimally important difference has advantages over the relative approach. *J Clin Epidemiol*. 2015;68(8):888–894
34. US Department of Health and Human Services. Guidance for industry. Patient-reported outcome measures: use in medical product development to support labeling claims. 2009 December. Available at: www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatory Information/Guidances/UCM193282.pdf. Accessed August 9, 2016

Minimally Important Differences in Patient or Proxy-Reported Outcome Studies Relevant to Children: A Systematic Review

Shanil Ebrahim, Kelsey Vercammen, Arunima Sivanand, Gordon H. Guyatt, Alonso Carrasco-Labra, Ricardo M. Fernandes, Mark W. Crawford, Gihad Nesrallah and Bradley C. Johnston

Pediatrics 2017;139;
DOI: 10.1542/peds.2016-0833 originally published online February 14, 2017;

PEDIATRICS®
OFFICIAL JOURNAL OF THE AMERICAN ACADEMY OF PEDIATRICS

Minimally Important Differences in Patient or Proxy-Reported Outcome Studies Relevant to Children: A Systematic Review

Shanil Ebrahim, Kelsey Vercammen, Arunima Sivanand, Gordon H. Guyatt, Alonso Carrasco-Labra, Ricardo M. Fernandes, Mark W. Crawford, Gihad Nesrallah and Bradley C. Johnston

Pediatrics 2017;139;
DOI: 10.1542/peds.2016-0833 originally published online February 14, 2017;