International consensus on a standard set of outcome measures for child and youth anxiety, depression, obsessive-compulsive disorder, and post-traumatic stress disorder

Karolin R Krause, Sophie Chung, Abiodun O Adewuya, Anne Marie Albano, Rochelle Babins-Wagner, Laura Birkinshaw, Peter Brann, Cathy Creswell, Kathleen Delaney, Bruno Falissard, Christopher B Forrest, Jennifer L Hudson, Shin-ichi Ishikawa, Meghna Khatwani, Christian Kieling, Judi Krause, Kanika Malik, Vania Martínez, Faraz Mughal, Thomas H Ollendick, Say How Ong, George C Patton, Ulrike Ravens-Sieberer, Peter Szatmari, Evie Thomas, Lucas Walters, Bryan Young, Yue Zhao, Miranda Wolpert

Lancet Psychiatry 2021; 8: 76–86

Evidence Based Practice Unit, Faculty of Brain Sciences, University College London, London, UK (K R Krause PhD); **Anna Freud National Centre for Children and Families, London, UK** (K R Krause); **International Consortium for Health Outcomes Measurement (ICHOM), London, UK** (S Chung BA); **Department of Behavioral Medicine, Lagos State University College of Medicine, Lagos, Nigeria** (Prof A O Adewuya FWACP); **Department of Psychiatry, New York State Psychiatric Institute, New York, NY, USA** (Prof A M Albano PhD); **Columbia University Vagelos College of Physicians and Surgeons, Columbia University, New York, NY, USA** (Prof A M Albano); **Calgary Counselling Centre, Calgary, AB, Canada** (R Babins-Wagner PhD); **Faculty of Social Work, University of Calgary, Calgary, AB, Canada** (R Babins-Wagner); **Child and Youth Mental Health Service, Eastern Health, Melbourne, VIC, Australia** (P Brann PhD, E Thomas BA); **School of Clinical Sciences, Monash University, Melbourne, VIC, Australia** (P Brann); **Young Minds, London, UK** (L Birkinshaw HBA); **Departments of Experimental Psychology and Psychiatry, University of Oxford, Oxford, UK** (Prof C Creswell PhD); **Rush College of Nursing, Chicago, IL, USA** (Prof K Delaney PhD); **Université Paris-Saclay, Gif-sur-Yvette, France** (Prof B Falissard PhD); **Université de Versailles Saint-Quentin-en-Yvelines, Versailles, France** (Prof B Falissard); **Institut national de la santé et de la recherche médicale (INSERM), Paris, France** (Prof B Falissard); **Centre de recherche en Epidémiologie et Santé des**

A major barrier to improving care effectiveness for mental health is a lack of consensus on outcomes measurement. The International Consortium for Health Outcomes Measurement (ICHOM) has already developed a consensus-based standard set of outcomes for anxiety and depression in adults (including the Patient Health Questionnaire-9, the Generalised Anxiety Disorder 7-item Scale, and the WHO Disability Schedule). This Position Paper reports on recommendations specifically for anxiety, depression, obsessive-compulsive disorder, and post-traumatic stress disorder in children and young people aged between 6 and 24 years. An international ICHOM working group of 27 clinical, research, and lived experience experts formed a consensus through teleconferences, an exercise using an adapted Delphi technique (a method for reaching group consensus), and iterative anonymous voting, supported by sequential research inputs. A systematic scoping review identified 70 possible outcomes and 107 relevant measurement instruments. Measures were appraised for their feasibility in routine practice (ie, brevity, free availability, validation in children and young people, and language translation) and psychometric performance (ie, validity, reliability, and sensitivity to change). The final standard set recommends tracking symptoms, suicidal thoughts and behaviour, and functioning as a minimum through seven primarily patient-reported outcome measures: the Revised Children's Anxiety and Depression Scale, the Obsessive Compulsive Inventory for Children, the Children's Revised Impact of Events Scale, the Columbia Suicide Severity Rating Scale, the KIDSCREEN-10, the Children's Global Assessment Scale, and the Child Anxiety Life Interference Scale. The set's recommendations were validated through a feedback survey involving 487 participants across 45 countries. The set should be used alongside the anxiety and depression standard set for adults with clinicians selecting age-appropriate measures.

Introduction

Depression and anxiety affect an estimated 4·4% and 3·6% of the world's population, and rank as the first and sixth largest contributor to health-related disability, respectively.[1] These disorders frequently emerge in childhood and adolescence and, unless treated early and effectively, commonly adversely affect mental health and psychosocial outcomes across the life course.[2–4] Despite an increase in mental health-care provision, service systems have failed to reduce the prevalence of these disorders in children and young people.[5] The global response to this problem requires holistic strengthening, not only in specialist mental health services, but also in primary care, community health, child health, and school settings.

In addition to resourcing and training in evidence-based care, one essential element of service strengthening is the systematic monitoring of patient progress.[6,7] Valid data for treatment outcomes are an essential facet for evaluating the effectiveness of care and can inform the setting of strategic targets for health systems, comparisons between different systems or services, and clinical decision making on a case-by-case basis.[8,9]

Currently, there is neither agreement nor global guidance on how best to track the response to clinical care for anxiety and depression in children and young people. Uptake of routine outcome measurement is low and, when done, there is considerable variation in outcomes, instruments, and assessment timepoints; a review published in 2019 recorded 15 different measures used to assess primary outcomes across 19 studies of routine treatment for anxiety and depression in young people.[10] Resulting data gaps and inconsistencies severely limit the potential for comparing different models of clinical care, identifying good practice, and informing quality improvement efforts.

This initiative aimed to address this challenge by devising a standard set, that is, a consensus-based standardised collection of treatment outcomes to be measured and reported as a minimum by all those providing relevant care.[11] To ensure that this standard is meaningful and acceptable to its intended users, the International Consortium for Health Outcomes Measurement (ICHOM) convened an international working group of lived experience experts, practitioners, and researchers to agree on a set of outcome domains, measurement instruments, case-mix factors (ie, case characteristics that should be considered when adjusting for differences in the composition of service user populations, or care provision across settings), and measurement timepoints (panel 1).

An existing ICHOM standard set for adult anxiety and depression covers young people from the age of 14 years (appendix p 3).[12] It includes a recommendation for tracking symptom change via the Generalised Anxiety

Disorder 7-item Scale (GAD-7)[15] and the depression subscale of the Patient Health Questionnaire (PHQ-9),[16] as well as functioning through the WHO Disability Assessment Schedule (WHODAS) 2.0 12-item short form.[17] Our standard set aims to complement this effort by providing recommendations specifically tailored for use with children and young people. The combination of the two sets will provide for the transition from youth into adult services at any point between the ages of 14 and 24 years, allowing for local variation in transition and judgments about which set is most suitable for different ages.

This standard set is designed for children and young people aged between 6 and 24 years who access care for anxiety, depression, obsessive-compulsive disorder (OCD), or post-traumatic stress disorder (PTSD), as defined by standard diagnostic criteria. All are internalising disorders, typically characterised by high negative affect, with OCD and PTSD long classified as anxiety disorders.[18] This standard set is recommended for use by all those providing care to the population in scope worldwide, regardless of intervention setting or approach. The working group sought to combine self-reported, parent-reported, and clinician-reported outcome measures to account for the different perspectives these reporters can provide.[19–21] Parent report measures also serve as a proxy when children and young people are unable to complete measures because of young age or developmental constraints (appendix p 3).[21] The process and methods used to develop this standard set are summarised in panel 2 and shown in figure 1, with additional detail provided in the appendix (pp 3–18).

Recommended outcomes and measures

As per the working group consensus, this standard set recommends tracking response to treatment across the three outcome domains of symptoms, suicidal thoughts and behaviour, and functioning, as a minimum, using seven instruments that are primarily self reported. These seven measurement instruments were selected because they fulfilled most or all of the working group's appraisal criteria (table 1, panel 2). A detailed discussion of instrument properties, performance against appraisal criteria, and accessibility is provided in the appendix (pp 19–22). 463 practitioners and researchers, and 24 lived-experience experts provided feedback on the draft working group recommendations through an open review process (panel 2, appendix p 18). Overall, 75% of practitioners and researchers stated that they had confidence in the recommended outcomes and instruments. In addition, 75–100% of participants with lived experience rated each included outcome domain as important, and 100% of lived experience participants confirmed the acceptability of the recommended measures for use in clinical practice. No additional outcome domains were consistently highlighted as missing from the recommended set during the open review process.

Panel 1: The International Consortium for Health Outcomes Measurement (ICHOM) approach to standard set development

ICHOM is a non-profit organisation that specialises in the development of condition-specific standard sets for clinical practice. ICHOM has supported the development of more than 28 existing standard sets, including one for adult depression and anxiety.[12] Outcome measurement is approached within a framework of person-centred and value-based health care, in which value is defined as the health outcomes achieved, relative to the resources invested, rather than the volume of services delivered.[13] Within a person-centred framework, value should be defined around outcomes that matter to service users. All ICHOM standard sets are condition specific, based on the understanding that the needs of service users and the treatment options available to them are at least partly shaped by the principal presenting problems.[13,14] Service users are directly involved in defining the standard set, which must include patient-reported outcomes. The final set of outcomes should represent the end result rather than the process of care, should balance a comprehensive approach to tracking outcome with a feasible recommendation that services can reliably implement, and should be responsive to quality improvement efforts.

Symptoms

The set recommends measuring anxiety and depression symptoms through youth and parent report for all children and young people within the prespecified scope (ie, those aged between 6 and 24 years presenting with anxiety, depression, OCD, or PTSD). To minimise the length and complexity of the standard set, the working group chose to recommend a joint measure of anxiety and depression symptoms—the 25-item short form of the Revised Children's Anxiety and Depression Scale (RCADS-25).[26] The group initially selected the scale's 47-item version on the basis of the appraisal criteria,[52] but decided to recommend the short form to reduce respondent burden. The RCADS-25 consists of 15 items that track anxiety symptoms and ten items that track symptoms of major depression, which can be summed to compute aggregate anxiety, depression, and total internalising symptom scores. Although less widely validated than the RCADS-47, the RCADS-25 met most inclusion criteria (table 1), although evidence of its sensitivity to change was not available at the time of the appraisal. Long or short versions of the RCADS have been applied in Africa, Europe, the Americas, and Asia.[53]

Symptoms of OCD and PTSD should be tracked separately for children and young people presenting with a diagnosis of OCD or PTSD, or with sub-threshold symptoms as appropriate. Symptoms of OCD and PTSD should be tracked via the self-reported 21-item Obsessive

Populations (CESP), Villejuif, Île-de-France, France (Prof B Falissard); **Children's Hospital of Philadelphia, Philadelphia, PA, USA** (Prof C B Forrest MD); **Department of Psychology, Centre for Emotional Health, Macquarie University, Sydney, NSW, Australia** (Prof J L Hudson PhD); **Faculty of Psychology, Doshisha University, Kyoto, Japan** (Prof S-I Ishikawa PhD); **Department of Psychology, University of Delhi, Delhi, India** (M Khatwani BA); **Hospital de Clínicas de Porto Alegre, Universidade Federal do Rio Grande do Sul, Porto Alegre, Rio Grande do Sul, Brazil** (C Kieling MD); **Child and Youth Mental Health Service, Children's Health Queensland Hospital and Health Service, Brisbane, QLD, Australia** (J Krause MA); **Sangath, New Delhi, Delhi, India** (K Malik PhD); **Centro de Medicina Reproductiva y Desarrollo Integral del Adolescente (CEMERA), Facultad de Medicina, Universidad de Chile, Santiago, Chile** (Prof V Martínez PhD); **Agencia Nacional de Investigación y Desarrollo (ANID), Millennium Science Initiative Program, Millennium Nucleus to Improve the Mental Health of Adolescents and Youths (IMHAY), and Millennium Institute for Research in Depression and Personality (MIDAP), Santiago, Chile** (Prof V Martínez); **School of Primary, Community and Social Care, Keele University, Staffordshire, UK** (F Mughal MBChB); **Department of Psychology, Child Study Center, Virginia Tech, Blacksburg, VA, USA** (Prof T H Ollendick PhD); **Department of Child and Adolescent Psychiatry, Institute of Mental Health, Singapore, Singapore** (S H Ong MMED); **Murdoch Children's Research Institute, Melbourne, VIC, Australia** (Prof G C Patton MD); **Department of Paediatrics, The University of Melbourne, Melbourne, VIC, Australia** (Prof G C Patton); **Department of Child and Adolescent Psychiatry, Psychotherapy, and Psychosomatics, University Medical Center Hamburg-Eppendorf,**

Hamburg, Germany (Prof U Ravens-Sieberer PhD); **The Hospital for Sick Children, University of Toronto, Toronto, ON, Canada** (Prof P Szatmari MD); **Cundill Centre for Child and Youth Depression, Centre for Addiction and Mental Health, Toronto, ON, Canada** (Prof P Szatmari MD); **Jack.org, Toronto, ON, Canada** (L Walters BSc, B Young BA); **Teaching and Learning Evaluation and Measurement Unit, The University of Hong Kong, Hong Kong** (Y Zhao EdD); **and Wellcome Trust, London, UK** (Prof M Wolpert DClinPsy)

Correspondence to: Dr Karolin R Krause, Evidence Based Practice Unit, Faculty of Brain Sciences, University College London, London, WC1E 6BT, UK **karolin.krause.16@ucl.ac.uk**

See **Online** for appendix

For more on **WHODAS** see https://www.who.int/standards/classifications/international-classification-of-functioning-disability-and-health/who-disability-assessment-schedule

Panel 2: Methods

The working group comprised 27 experts from 13 countries, including youth with lived experience, parents or carers (hereafter referred to as parents), mental health practitioners, and researchers working within relevant disciplines (eg, psychiatry, epidemiology, psychometrics; appendix pp 3–4). The central project team (KRK, SC, MW) coordinated and facilitated the consensus-building process and completed supporting research tasks but did not vote on the consensus recommendations. A detailed description of the process and methods is provided in the appendix (pp 3–18).

General process

From October, 2018, until December, 2019, the working group completed a 14-month structured and evidence-informed consensus-building process (figure 1). The group convened for eight teleconferences, completed a three-round adapted Delphi exercise (a method for reaching group consensus) to select outcome domains,[22,23] and participated in iterative rounds of anonymous voting to arrive at recommendations for outcomes, measurement instruments, case-mix factors, and timepoints (appendix pp 4–5). The process was informed by sequential research inputs done by the central project team (appendix pp 5–18).

Selection of outcome domains

In line with methodological recommendations by the Core Outcome Measures in Effectiveness Trials (COMET) initiative,[24] 70 possible outcomes and 507 measurement instruments were identified through multiple avenues, including a systematic scoping review of 257 treatment outcome studies, and supplemental sources (eg, cohort studies, qualitative outcome research, and instrument banks), and break-out groups with service user representatives (appendix pp 5–8).

Selection of measurement instruments

Once the group reached a consensus on the outcomes to include, 107 thematically relevant measures were systematically appraised to identify those most suitable for tracking the selected outcomes over time. Appraisal criteria included relevance (ie, comprehensive coverage of the selected outcome domain), feasibility and acceptability (ie, completion within less than 20 min, free availability for use in clinical settings, including in paper-and-pencil format, previous validation in children and young people, and translation into at least a second language), and psychometric performance (ie, inter-rater or retest reliability above 0·70, internal consistency above 0·70, and evidence of sensitivity to change), in line with International Society for Quality of Life Research recommendations (appendix pp 8–16).[25] The instruments judged to best satisfy these criteria were shortlisted, and their strengths and shortcomings discussed. At this stage, other aspects, such as content and construct validity, were also considered. The final set of measures was then selected via consensus through iterative rounds of anonymous voting. The working group aimed for a standard set that would be simple to use, impose a low burden on its intended users, and be applicable across different contexts.

Selection of case-mix factors and measurement timepoints

Based on the systematic scoping review and existing ICHOM standard sets, the central project team compiled a list of possible case-mix factors, and did a rapid review of reviews examining predictors, mediators, and moderators of treatment response (appendix pp 17–18). The working group discussed this information and formed consensus on the case-mix factors to include. Similarly, the central project team drew on existing ICHOM standard sets to present an initial proposal for measurement timepoints, which were discussed, refined, and voted on by the working group.

Open review of draft recommendations

To establish the generalisability and acceptability of the working group's recommendations, an open review web survey gathered external feedback from 463 practitioners, researchers, and policy makers from 45 countries, as well as 24 young people and parents from Denmark, the UK, and the USA (appendix p 18).

Compulsive Inventory for Children (OCI-CV)[27] and the Children's Revised Impact of Events Scale (CRIES)[54] for youth (CRIES-8) and parent report (CRIES-13). Both measures have been applied in Europe, Asia, and the Americas. As of March, 2020, a version of the OCI-CV for parents was not available.

It is important to note that these symptom measures are recommended for the purpose of tracking treatment outcomes over time. They are not considered primarily for the purpose of diagnosing presenting problems and are not intended to replace a thorough clinical assessment using the latest diagnostic tools. The latter require additional properties related to diagnostic validity (eg, sensitivity and specificity) that the working group did not explicitly consider during measure appraisal.

Suicidal thoughts and behaviour

A consensus was reached on measuring suicidal thoughts and behaviour in young people aged 10 years and older, unless considered inappropriate (eg, due to the cultural context), using the Columbia Suicide Severity Rating Scale (C-SSRS) recent self-report screener[28]—a short self-report version of the clinician-administered C-SSRS interview protocol. The measure consists of six items that track the severity of suicidal ideation and behaviour in the previous month. We did not identify any study specifically validating the C-SSRS recent self-report screener in children and young people, but the clinician-rated C-SSRS has shown good internal consistency, inter-rater reliability, and sensitivity to change in adolescent samples.[28]

Functioning

Functioning describes a child's ability to engage in typical activities and meet role demands in line with age-specific sociocultural norms.[55,56] None of the identified functioning measures fully satisfied the working group's requirements and a consensus was reached on mitigating this issue by tracking a broad concept of global functioning or health-related quality of life (HRQoL), as well as condition-specific functional impairment through short dedicated measures of each concept. Generic measures allow for comparisons across conditions, although condition-specific measures might be more sensitive to change. As per group consensus, measures had to cover psychosocial functioning, peer relationships, and sleep functioning at least at an item level, although sleep was eventually covered through the RCADS-25 (ie, item 8 ["I have trouble sleeping"] and item 9 ["I feel scared if I have to sleep on my own"]).[26]

The KIDSCREEN-10 (completed by children or young people, and parents) was selected as a generic measure of global functioning. This unidimensional ten-item index of HRQoL tracks functioning in relation to physical health and energy levels, leisure activities, social and family relationships, and cognition.[29] The more comprehensive KIDSCREEN-52 was originally developed through a process of cross-cultural harmonisation involving 13 European countries. A simultaneous approach was used, taking into account different cultural perspectives, to ensure international comparability and cross-cultural applicability.[57] Its ten-item short form has been applied in Asia, eastern Africa, Europe, North America, and Latin America. Although originally designed for epidemiological studies, the KIDSCREEN-10 has been shown to discriminate well between children and young people with high and low levels of functioning, with few so-called ceiling or floor effects.[29] In addition to the KIDSCREEN-10, the Children's Global Assessment Scale (CGAS) was selected as a brief clinician-rated measure of global functioning.[30] On this widely used measure, clinicians perform a single rating by locating children and young people on a scale from 1 to 100, placing them into one of ten categories, from 1 to 10 (indicating high impairment) to 91 to 100 (indicating superior functioning).[30]

The Children's Anxiety Life Interference Scale (CALIS)[31] was selected to measure condition-specific impairment via nine or ten items that describe instances of anxiety that affected functioning at home, school, in social life, and in relation to other activities. The CALIS only covers anxiety-related impairment and the working group did not identify any eligible measure that tracked depression-specific impairment or that captured impairment from both anxiety and depression. However, as of March, 2020, the CALIS author team were revising the measure to cover impairment from both anxiety and depression, and validation studies involving children, young people, and parents in community, school, and clinical settings were ongoing. As soon as validation results and the revised measure become available the working group will consider replacing the original CALIS with the revised Children's Anxiety and Depression Life Interference Scale (CADLIS) in the standard set.

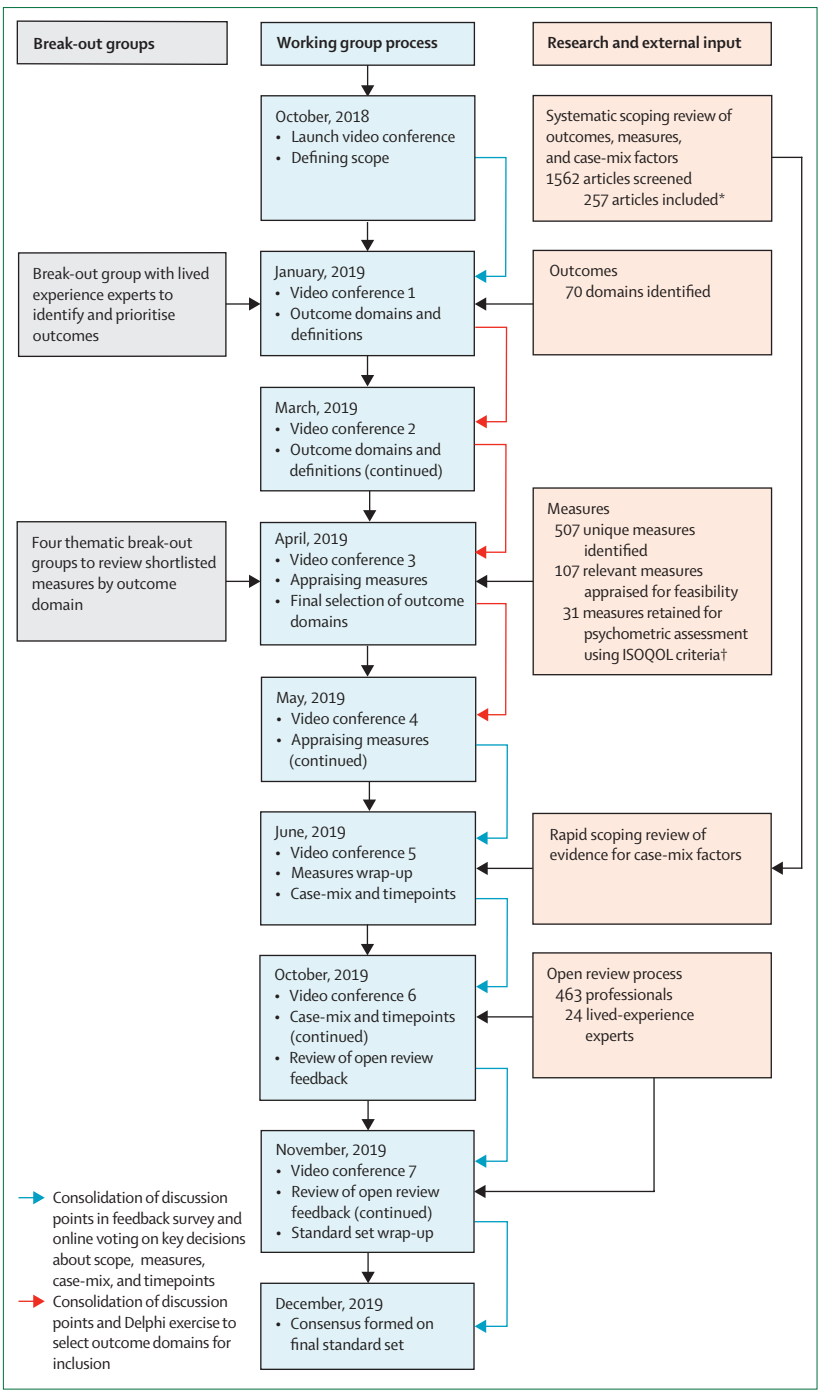

Figure 1: **Working group process**
*Supplementary sources are not included in this count. †Shortlisted measures were reviewed against the minimum standards for patient-reported outcome measures recommended by the International Society for Quality of Life Research (ISOQOL).[25]

	Measure	Relevance (domain coverage [ie, satisfactory domain coverage at subscale or item level])	Feasibility and acceptability				Psychometric performance*		
			Short (ie, <60 items, <20 min)	Free (ie, currently no cost for non-commercial use)	Validated (ie, at least one validation in a CYP sample)	Translated (ie, >1 language version available)	Reliability (ie, TRT or IRR ≥0·70)	Validity (ie, internal consistency [eg, Cronbach's α≥0·70])	Sensitivity to change (ie, any evidence of sensitivity to change)
Symptoms									
Anxiety and depression	RCADS-25	GAD, MDD, OCD, PD, SAD, SP	25 items (<10 min)	Yes	Ages 6–18 years (clinical and non-clinical)	5†	TRT[32]	Yes[26,33]	No evidence
OCD	OCI-CV	Doubting or checking, obsessing, hoarding, washing, ordering, neutralising	21 items (<10 min)	Yes	Ages 6–18 years (clinical and non-clinical)	>3	TRT[27,34–36]	Yes[27,34–38]	Some evidence[27]
PTSD	CRIES-8 (for children and youth) and CRIES-13 (for parent report)	Intrusion, avoidance (hyperarousal in CRIES-13)	8 items (CRIES-8) and 13 items (CRIES-13); (<5 min)	Yes	Ages 7–18 years (clinical and non-clinical)	>27	TRT[39]	Yes[39–41]	Some evidence[42]‡
Suicidal thoughts and behaviour									
Measure	C-SSRS	Severity of ideation, behaviour (ie, attempts)	3 or 6 items§ (<5 min)	Yes	Ages 12–18 years (clinical and non-clinical)	>100	IRR[43]¶	Yes[28]¶	Some evidence[28]¶
Functioning									
Measure	KIDSCREEN-10	Physical activity and energy, emotions, leisure time and participation, relationships with parents and peers, cognition, school performance	10 items (<5 min)	Yes	Ages 8–18 years (non-clinical)	45	TRT[29,44]	Yes[29,44,45]	No evidence
Measure	CGAS	Global functioning	1 item (<5 min)	Yes	Ages 4–18 years	>2	IRR[29,46–48] TRT[30,47]	NA	Some evidence[49]
Measure	CALIS	Enjoyable activities, relationships with siblings, parents, friends, peers, sports, schoolwork, distress	9 or 10 items (<5 min)	Yes	Ages 6–17 years (clinical and non-clinical)	7	TRT[31]	Yes[31,50,51]	Some evidence[31]

CALIS=Child Anxiety Life Interference Scale. CGAS=Children's Global Assessment Scale. CRIES-8=Children's Revised Impact of Events Scale for youth. CRIES-13=Children's Revised Impact of Events Scale for parents. C-SSRS=Columbia Suicide Severity Rating Scale. CYP=children and young people. GAD=generalised anxiety disorder. IRR=inter-rater reliability. MDD=major depressive disorder. NA=not applicable. OCD=obsessive compulsive disorder. OCI-CV=Obsessive Compulsive Inventory for Children. PD=panic disorder. RCADS-25=Revised Children's Anxiety and Depression Scale. SAD=separation anxiety disorder. SP=social phobia. TRT=test-retest reliability. *These thresholds are based on the International Society for Quality of Life Research minimum standards for patient-reported outcome measures.[25] †The RCADS-25 is currently available in five languages on the source website (appendix p 22) but the 47-item long form is available in 18 languages, meaning item-level translations are available in more than five languages for those items included in the RCADS-25. ‡These validation studies tested a parent-report version only. §Two initial questions serve as screeners, with the three remaining severity items administered only to those endorsing the first two (all young people are asked about suicidal behaviour). ¶To our knowledge, the C-SSRS self-report screener had not been subject to a validation study in children and young people, as of March, 2020. For the psychometric appraisal, studies assessing the psychometric properties of the severity subscale in the clinician-led C-SSRS semi-structured interview schedule were considered instead.

Table 1: **Overview of recommended outcome measures and their evaluation criteria, as of March, 2020**

Recommended case-mix factors

The standard set aims to facilitate comparisons and benchmarking of outcomes, which requires the collection of additional data to adjust for variation in populations and intervention settings. As per the working group consensus, services should record demographic, clinical, complexity, and intervention factors through a mix of self-report, parent report, and clinician-report (table 2). Beyond the working group, the case-mix factors that were suggested were endorsed by more than 80% of practitioners and researchers who provided feedback during the open review process. These factors represent a minimum that should be assessed by all those providing relevant care, and services might wish to add other indicators to meet local information needs.

Demographic factors

The set recommends recording age, sex assigned at birth, gender, ethnic minority status, socioeconomic status, and the child's living situation. Socioeconomic status should be measured by recording the highest level of education completed by any of the parents of children and young people as a widely accepted proxy that can be mapped onto the International Standard Classification of Education for international comparisons.[58] The set includes one question about children and young people's living situation and two questions that capture ethnic

minority and marginalised group status via self-report (table 2).

Clinical factors

Several studies show that symptom burden, symptom duration, and the presence of comorbidities affect treatment response in children and young people with depression or anxiety (appendix pp 17–18).[59–68] The group recommends recording principal and comorbid presenting problems by administering the 30 problem descriptions of the current view tool.[69] Although not equivalent with formal diagnoses, these problem descriptions broadly align with the diagnostic categories of the ICD-11 for paediatric populations.[70] ICHOM standards are available for recording symptom duration and previous service use (table 2).

Complexity factors

Research suggests that parental mental health influences treatment response in children and young people with anxiety and depression (appendix pp 17–18).[64,71–73] The set includes two questions about experiences with or diagnoses of mental health problems in the immediate family, and previous use of mental health services by the reporting parent. Although evidence is scarce and conflicting on the influence of adverse experiences (eg, childhood maltreatment) on treatment response,[64,74] the set recommends recording trauma history via the selected complexity factors section of the current view tool. Here, clinicians can indicate a range of adverse experiences based on available information. Problems at school or work should be tracked as additional complexity factors, using the current view tool.

Intervention factors

Services should collect information about the intervention approach, including the intervention focus (ie, who is actively involved), treatment modality and prescribed medication (ie, as per lists of options compiled by the working group and through open review feedback), and intervention setting (ie, whether or not treatment was delivered via a digital platform or inpatient care, as opposed to other settings). Services might wish to record additional detail on intervention characteristics to meet local, regional, or national information needs.

Recommended measurement timepoints

Timelines for measurement are highly practical considerations and are likely to vary substantially across services. The standard set makes a minimum recommendation (figure 2) for measuring outcomes over the full cycle of care but encourages services to do so as often as is clinically helpful to inform decision making, or to align with the collection of local or national data. The suggested timepoints were widely endorsed by over 80% of practitioners and researchers participating in the open review process.

	Description	Reporter
Demographic factors		
Age*	Year of birth	CYP or parent
Sex*	Sex assigned at birth	CYP or parent
Gender identity	"Do you think of yourself as…?"	CYP
Parent or carer education*	Highest level of education completed by any of the parents or carers of a child or young person (using ISCED standards)	CYP or parent
Ethnicity	"Do you consider yourself to be in an ethnic minority where you live?"	CYP or parent
Marginalised group status	"Do you consider yourself to be a member of a marginalised group where you live?"	CYP or parent
Living situation	"Which of the following people live with you [your child] at your [their] home?"	CYP or parent
Clinical factors		
Diagnoses and comorbidities	Measured via the provisional problems list of the current view tool	Clinician
Duration of symptoms*	"For how many months have you [your child] been experiencing [specific condition] symptoms?"	CYP or parent
Previous service use	"During the last year, did you [your child] receive any of the following treatments for [specific condition]? Complete separately for (a) medication, (b) psychotherapy, and (c) other."	CYP or parent
Complexity factors		
Trauma history	Measured via the selected complexity factors of the current view tool	Clinician
Parental mental health	"It is useful to know whether there is a family history of mental health problems. Have you, or anyone else in the immediate family, ever experienced, or been diagnosed with, any of the following conditions: anxiety, depression, a substance use disorder (eg, alcohol or drugs), schizophrenic disorder, personality disorder, somatoform disorder (ie, unexplained physical symptoms), or other?"	Parent
Parental service use	"Have you ever sought help for your mental health?"	Parent
Education or work difficulties	Measured via the contextual problems section of the current view tool	Clinician
Intervention variables		
Intervention focus (1)	"Who is actively involved in the intervention? (select all that apply)"	Clinician
Intervention focus (2)	"Is the intervention delivered to an individual child or family, or to a group of children or families?"	Clinician
Intervention approach	"What is the treatment approach? (select all that apply)"	Clinician
Prescribed medication	"What type of medication is prescribed? (select all that apply)"	Clinician
Intervention setting (1)	"Does this intervention involve an overnight stay at an institution providing mental health support?"	Clinician
Intervention setting (2)	"Does this intervention involve the use of a digital platform?"	Clinician

CYP=children and young people. ICHOM=International Consortium for Health Outcomes Measurement. ISCED=International Standard Classification of Education. OCD=obsessive-compulsive disorder. PTSD=post-traumatic stress disorder. *Variables defined and operationalised as per an existing ICHOM standard. Response categories for each question are provided in the full standard set reference guide (panel 3).

Table 2: **Case-mix factors in the standard set for child and youth anxiety, depression, OCD, and PTSD**

As per working group consensus, all case-mix factors and outcomes should be measured at the point of assessment or intake (ie, baseline), or as near to these timepoints as possible if services wish to collect data at second contact (figure 2). As a guideline, all outcome measures should be administered at least every 3 months after baseline measurements were taken. Services are encouraged to consider more frequent intervals, including session by session measurement, to help

For more on the **ICHOM standard** see https://www.ichom.org/portfolio/anxiety-depression-ocd-and-ptsd-in-children-and-young-people/

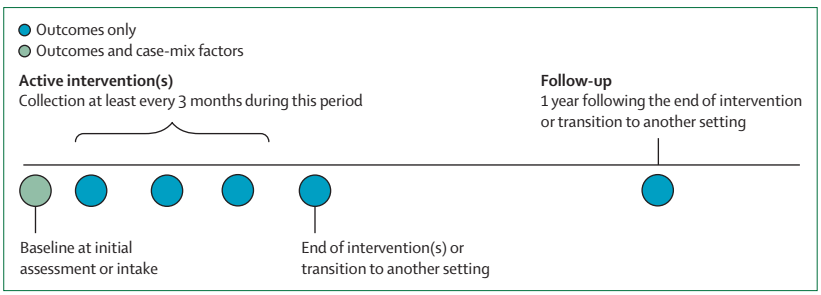

Figure 2: **Timeline for data collection**

For more on **ICHOM Connect** see connect.ichom.org

Panel 3: **How to access the standard set resources**

A standard set reference guide, flyer, and data dictionary is available free of charge and can be accessed via ICHOM Connect. The reference guide defines each outcome domain and provides a detailed description of the recommended measures, case-mix factors, and timepoints.

embed monitoring into the clinical process[75] and reduce the risk of missing data due to drop-out before follow-up. The group recognises that effective session by session measurement requires well established systems and can otherwise be perceived as unnecessarily burdensome.

To mark the end of an active treatment cycle, outcomes should be measured upon transition into adult services, into a different level of care (eg, from outpatient to inpatient care), or upon completion when no further activities are planned (figure 2). Outcomes should then be measured again at a follow-up assessment 1 year after baseline measurements were taken. This can enable important insights into longer-term outcomes, but could require substantial adjustments to the way services are currently organised and funded.

Strengths of the standard set

Of the working group's 27 voting members, four were young adults aged between 18 and 24 years, and two were parents. This was the first ICHOM working group to involve young people with lived experience of service use—rather than only their parents—as full working group members at all stages of the consensus-building process, including the teleconferences, the adapted Delphi exercise (a method for reaching group consensus),[22] and subsequent surveys. Although limited familiarity with specific methodological questions around outcome measurement can form a barrier to the meaningful participation of lived-experience experts,[76] the teleconferences helped foster a common understanding within the group ahead of each round of voting. The working group chair (MW) solicited input from all call participants on all key discussion points to promote equal participation and manage power imbalances. Additional feedback from a wider group of young people and parents was sought (via the open review process) towards the end of the consensus-building process. Variable requirements for ethical review for this stakeholder group meant that the survey was only accessible in three countries, while the professional survey was accessible globally. This difference in availability led to a comparatively small sample of lived-experience participants (n=24) in the open review process. However, separate analysis and review of feedback from the professional and lived-experience surveys meant that the working group was able to consider each feedback stream in its own right.

The working group included experts from low-income and middle-income countries (LMICs) who have rarely been represented by similar initiatives to date (appendix p 3).[77] Although 90% of children and young people live in LMICs, 90% of research on youth mental health comes from high-income countries.[78] Experts from LMICs were also consulted through the open review process (appendix p 18). A range of local needs, challenges, and cultural factors could thus be considered. The high endorsement of the standard set in the open review process underscores its relevance and acceptability beyond the immediate working group.

Implementation

Resources that can guide the implementation of this standard set are available from ICHOM (panel 3). Members of the working group will form a steering group to oversee the implementation of the standard set and to review recommendations while taking into consideration the lessons from pilot initiatives and new developments in the field. In several health systems (eg, from Australia, Canada, the Netherlands, the UK, or the USA), the routine collection of outcome data is becoming increasingly embedded,[79,80] which should facilitate the adoption of the standard set but can also cause issues of alignment with existing local or national systems.[81,82] As services adopt routine outcome monitoring, it is imperative that data are handled safely and securely in accordance with relevant data protection frameworks, and that informed consent protocols are in place.

This standard set was designed for clinical practice and might be suitable for use in clinical trials, depending on the requirements and the type of intervention tested. Since work on this standard set has ended, the Wellcome Trust and the National Institute of Mental Health (NIMH) have recommend the RCADS-25 as an outcome measure for research with children and adolescents experiencing depression or anxiety (suitable for ages 8–18). The funders suggest that the PHQ-9 and GAD-7 could be used instead of the RCADS-25 with adolescents and young adults, where this is judged developmentally appropriate. The funders recommend the WHODAS 2.0 as a functioning measure for young adults from the age of 18 years. A standard set designed specifically for clinical trials in adolescent depression is under development as part of the International Network for Research Outcomes in Adolescent Depression Studies (IN-ROADS) initiative at

the Hospital for Sick Children (Toronto, ON, Canada).[83] UNICEF is leading another complementary initiative focusing on mental health outcomes for adolescents with anxiety or depression in population surveys.[84] Although there is a unique opportunity to encourage further harmonisation, a degree of divergence might persist because of differences in priorities, processes, and methodologies (eg, outcome measures for clinical trials might not need to be freely available). The present work adds to existing review efforts to identify meaningful, feasible, and acceptable outcome measures suitable for use with children and young people in practice settings.[85–89] The steering group will consider opportunities for alignment with complementary initiatives, as well as with existing ICHOM standard sets (eg, with a view to linking scores obtained from sets for children and young people on the one hand, and adult sets on the other hand, for the purpose of longitudinal analysis).[90]

Limitations and areas for future research

The scope of this standard set is limited to anxiety, depression, OCD, and PTSD. The working group has sought to make a minimum recommendation focusing on core outcomes for these target disorders. Clinical judgment is warranted when tracking additional symptoms over time that are not covered by the recommended outcome measures. Three complementary ICHOM standard sets are available that focus on psychotic disorders (covering bipolar disorder), personality disorders, and substance use and addictive behaviour disorders in young people and adults. The present standard set captures the presence of these disorders and other comorbid presenting problems at baseline via the current view screening tool for the purpose of case-mix adjustment.

This standard set was developed by a small working group of consistent members that convened at frequent intervals. There was not complete parity in representation across different strata of experts, including young people and parents. In future efforts, large-scale Delphi surveys might be more suitable for consulting equal numbers of different stakeholder groups, albeit at the expense of more in-depth and continuous group deliberation.

Because of the nature of consensus building, compromises had to be made and not all individual views and priorities could be reflected in the final standard set. For example, two additional outcome domains (ie, coping, and interference of treatment with daily life) reached consensus among the group's lived experience experts, but not within the wider group. To promote person-centred care, services could consider tracking additional outcomes on the basis of shared decision making with children, young people, and families, either through suitable standardised measures or by including personalised outcome measures that track progress in relation to individual problems or treatment goals.[91]

To promote uptake of this standard set, simplicity and feasibility were prioritised over detail and specificity. Brief and freely available instruments were prioritised over more complex and, at times, more established ones, and short forms over long versions. As the former tend to be less widely validated than the latter, it is hoped that the standard set will accelerate validation efforts and generate new data for their psychometric properties.[32] To maintain simplicity, the set does not recommend separate measures for different age groups. Although the selected measures have generally been validated in children and young people aged between 6 and 18 years, the set is less specifically tailored to the experiences of those aged between 18 and 24 years. As the set moves into implementation, the steering group will consider whether, on balance, the gains from increased feasibility can be seen to justify this design.

None of the recommended measures are perfect.[92] The goal was to identify the best-possible suite of instruments within the working group's feasibility and psychometric criteria, based on the evidence available at the time, as a starting point for generating wider insights into how outcome measurement might be strengthened in the future. Although the working group considered sensitivity to change an essential measurement property, its appraisal was hindered by the scarcity of data and objective appraisal guidelines. The International Society for Quality of Life Research (ISOQOL) recommends that for longitudinal research, patient-reported outcome measures "should have evidence of responsiveness, including empirical evidence of changes in scores consistent with predefined hypotheses" about the expected treatment outcome.[25] Although the working group considered such evidence when available, no standard thresholds could be applied to determine whether sensitivity was sufficient.

For more on **ICHOM standard sets for other disorders** see https://www.ichom.org/standard-sets/

An important area that could not be considered as part of the appraisal is the measurement invariance of the selected tools across languages and cultural backgrounds, and the extent to which cultural differences might affect the validity of the selected measurement instruments.[93] In the absence of invariance, comparisons between different groups are not fully meaningful, and the working group hopes that this initiative will enhance data availability and spur efforts to examine how consistently the recommended measures track their designated outcome concepts across cultural and language contexts.

The working group encountered challenges with identifying suitable measures of functioning, with common issues including overlap in item coverage between symptoms and functioning, a perceived overemphasis on bodily functions as opposed to psychosocial functioning, an absence of validation across the full age span or in relevant clinical populations, and cross-cultural validity. Overall, additional research is needed to understand how suitable the recommended measures are for capturing change over the course of treatment, how suitable these measures are in different real-world clinical settings (eg, primary *vs* specialist care), and how acceptable to services and service users.

Call for action

This standard set provides the first global guideline for promoting the quality and consistency of routine outcome measurement for children and young people with anxiety, depression, OCD, or PTSD. It is person centred and devised specifically for use in clinical practice, with special attention given to acceptability and feasibility, including in resource-poor contexts. It also has great relevance to the provision of mental health care outside the health service system, such as in school settings. It forms an essential step towards enhancing evidence on service effectiveness, enabling comparisons and benchmarking of results across care systems, and promoting care quality in mobilising a comprehensive, forceful, and evidence-based response to the global burden of anxiety and depression. Future research should continue to expand the evidence base in relation to the sensitivity to change and measurement invariance of the included measures, and implementation initiatives should provide feedback on the relevance and acceptability of this recommendation to practitioners and service users. Both will be crucial to ensure that this standard set makes a viable recommendation that meaningfully captures change for children and young people across contexts.

Contributors

KRK, MW, and SC oversaw the study design. KRK and SC led data acquisition through the literature reviews, Delphi surveys, and open review surveys, as well as data analysis and interpretation. KRK took responsibility for the integrity of the data and the accuracy of data analysis. All authors attended teleconferences, and all but the core project team (KRK, MW, SC) completed anonymous votes and feedback surveys as part of the Delphi process and formed consensus on the final recommendation. KRK drafted the manuscript with support from SC, and all co-authors critically reviewed the working draft and agreed to the revisions and final submission. Administrative, technical, and organisational support was provided by SC.

Declaration of interests

AMA receives royalties from Oxford University Press for the Anxiety Disorders Interview Schedule (ADIS), child and parent versions. PB is involved with the development of routine outcome measurement in public mental health across Australia. He chairs the National Child and Adolescent Mental Health Information Development Expert Advisory Group which supports routine outcome measurement and benchmarking between organisations, but has no financial interest in any system. SC was an employee of the International Consortium for Health Outcomes Measurement (ICHOM) during this study. ICHOM received grants from the New South Wales Agency for Clinical Innovation (NSW, Australia), Västra Götaland Regional Council (Västra Götaland, Sweden), and NHS England & NHS Improvement, in support of this work. CC has led several studies and ongoing evaluations of commonly used scales, including the Spence Children's Anxiety Scale (SCAS), Revised Children's Anxiety and Depression Scale (RCADS), and the Child Anxiety Interference Scale (CAIS). BF reports personal fees from Eli Lilly, Bristol Myers Squibb, Servier, Sanofi, GlaxoSmithKline, HRA Pharma, Roche, Boehringer Ingelheim, Bayer, Almirall, Allergan, Stallergene, Genzyme, Pierre Fabre, AstraZeneca, Novartis, Janssen, Astellas, Biotronik, Daiichi- Sankyo, Gilead, MSD, Lundbeck, Actelion, UCB, Otsuka, Grunenthal, ViiV, outside the submitted work. JLH is one of the developers of the Child Anxiety Life Interference Scale (CALIS), which is a freely available measure and there is no financial conflict of interest. S-II has participated in several studies that aimed to develop Japanese versions of scales, including the Spence Children's Anxiety Scale, Children's Depression Scale, and the Social Phobia and Anxiety Inventory for Children. KRK received personal fees from ICHOM during the study. KRK is also involved with the International Network for Research Outcomes in Adolescent Depression Studies (IN-ROADS) initiative that aims to develop a core outcome set specifically for youth depression clinical trials. UR-S is one of the developers of the KIDSCREEN-10, which is a freely available measure and there is no financial conflict of interest. Since May, 2019, MW has been head of the new Mental Health Priority Area at the Wellcome Trust, which might be developing standard sets in mental health in the future. She has been involved in the development of the current view tool, which is a freely available measure and there is no financial conflict of interest. MW was previously head of the Child Outcomes Research Consortium (CORC), which advises on measurement in child mental health, and an adviser to National Health Service (UK) on informatics. All other authors declare no competing interests.

Acknowledgments

We would like to gratefully acknowledge all individuals and organisations who took the time to share or complete the open review feedback survey towards improving the standard set recommendation. We would also like to thank Luz Sousa Fialho (International Consortium for Health Outcomes Measurement [ICHOM], London, UK) for providing administrative support, Katherine G Young (ICHOM, London, UK) for providing research support, and Christiana Ilesanmi (ICHOM, London, UK) for providing both administrative and research support to this project, as well as Zofia Das-Gupta (ICHOM, London, UK) for providing guidance on the application and management of the ICHOM process until May, 2019. We would like to thank one further lived experience expert who participated in the early stages of the process, including the Delphi process, but did not wish to be a named co-author. This work has been funded by NHS England & NHS Improvement, Västra Götaland Regional Council (Västra Götaland, Sweden), New South Wales Agency for Clinical Innovation (NSW, Australia). The funders had no role in the study design, data analysis, interpretation, or writing of the report. The funders supported data collection by sharing the open review survey for practitioners within their networks. CC was supported by a National Institute for Health Research (NIHR) Research Professorship. FM was supported by an NIHR in-practice fellowship and is supported by an NIHR school for primary care research GP career progression fellowship. The views expressed in this manuscript are those of the authors and not necessarily those of the NHS, NIHR, or Department for Health and Social Care. VM received funding from Agencia Nacional de Investigación y Desarrollo—Millennium Science Initiative Program—NCS17_03, and ICS13_005. CK is a Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) researcher and an Academy of Medical Sciences Newton Advanced Fellow.

References

1 WHO. Depression and other common mental disorders: global health estimates. Geneva; 2017. https://apps.who.int/iris/bitstream/handle/10665/254610/WHO-MSD-MER-2017.2-eng.pdf (accessed Jan 10, 2020).

2 Kessler RC, Amminger GP, Aguilar-Gaxiola S, Alonso J, Lee S, Ustün TB. Age of onset of mental disorders: a review of recent literature. *Curr Opin Psychiatry* 2007; **20:** 359–64.

3 Albano AM, Chorpita BF, Barlow DH. Childhood anxiety disorders. In: Mash EJ, Barkley RA, eds. Child psychopathology, 2nd edn. New York, NY: The Guilford Press, 2003: 279–329.

4 Clayborne ZM, Varin M, Colman I. Systematic review and meta-analysis: adolescent depression and long-term psychosocial outcomes. *J Am Acad Child Adolesc Psychiatry* 2019; **58:** 72–79.

5 Jorm AF, Patten SB, Brugha TS, Mojtabai R. Has increased provision of treatment reduced the prevalence of common mental disorders? Review of the evidence from four countries. *World Psychiatry* 2017; **16:** 90–99.

6 Boswell JF, Kraus DR, Miller SD, Lambert MJ. Implementing routine outcome monitoring in clinical practice: benefits, challenges, and solutions. *Psychother Res* 2015; **25:** 6–19.

7 Scott K, Lewis CC. Using measurement-based care to enhance any treatment. *Cognit Behav Pract* 2015; **22:** 49–59.

8 Barkham M, Mellor-Clark J. Bridging evidence-based practice and practice-based evidence: psychological therapies. *Clin Psychol Psychother* 2003; **10:** 319–27.

9 Bickman L. A measurement feedback system (MFS) is necessary to improve mental health outcomes. *J Am Acad Child Adolesc Psychiatry* 2008; **47:** 1114–19.

10 Bear HA, Edbrooke-Childs J, Norton S, Krause KR, Wolpert M. Systematic review and meta-analysis: outcomes of routine specialist mental health care for young people with depression and/or anxiety. *J Am Acad Child Adolesc Psychiatry* 2019; **59:** 810–41.

11 Clarke M, Williamson PR. Core outcome sets and systematic reviews. *Syst Rev* 2016; **5:** 11.

12 Obbarius A, van Maasakkers L, Baer L, et al. Standardization of health outcomes assessment for depression and anxiety: recommendations from the ICHOM Depression and Anxiety Working Group. *Qual Life Res* 2017; **26:** 3211–25.

13 Porter ME. What is value in health care? *N Engl J Med* 2010; **363:** 2477–81.

14 Kazdin AE. The meanings and measurement of clinical significance. *J Consult Clin Psychol* 1999; **67:** 332–39.

15 Spitzer RL, Kroenke K, Williams JBW, Löwe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. *Arch Intern Med* 2006; **166:** 1092–97.

16 Spitzer RL, Kroenke K, Williams JBW. Validation and utility of a self-report version of PRIME-MD: the PHQ primary care study. Primary care evaluation of mental disorders. Patient health questionnaire. *JAMA* 1999; **282:** 1737–44.

17 Ustün TB, Chatterji S, Kostanjsek N, et al. Developing the World Health Organization disability assessment schedule 2.0. *Bull World Health Organ* 2010; **88:** 815–23.

18 Regier DA, Kuhl EA, Kupfer DJ. The DSM-5: classification and criteria changes. *World Psychiatry* 2013; **12:** 92–98.

19 De Los Reyes A, Augenstein TM, Wang M, et al. The validity of the multi-informant approach to assessing child and adolescent mental health. *Psychol Bull* 2015; **141:** 858–900.

20 Weisz JR, Kuppens S, Ng MY, et al. What five decades of research tells us about the effects of youth psychological therapy: a multilevel meta-analysis and implications for science and practice. *Am Psychol* 2017; **72:** 79–117.

21 Creswell C, Waite P, Hudson J. Practitioner review: anxiety disorders in children and young people—assessment and treatment. *J Child Psychol Psychiatry* 2020; **61:** 628–43.

22 Linstone HA, Turoff M. The Delphi method: techniques and applications. Reading, MA, USA: Adison Wesley, 1975.

23 Prinsen CAC, Vohra S, Rose MR, et al. How to select outcome measurement instruments for outcomes included in a "core outcome set"—a practical guideline. *Trials* 2016; **17:** 449.

24 Williamson PR, Altman DG, Bagley H, et al. The COMET Handbook: version 1.0. *Trials* 2017; **18** (suppl 3): 280.

25 Reeve BB, Wyrwich KW, Wu AW, et al. ISOQOL recommends minimum standards for patient-reported outcome measures used in patient-centered outcomes and comparative effectiveness research. *Qual Life Res* 2013; **22:** 1889–905.

26 Ebesutani C, Reise SP, Chorpita BF, et al. The Revised Child Anxiety and Depression Scale-Short Version: scale reduction via exploratory bifactor modeling of the broad anxiety factor. *Psychol Assess* 2012; **24:** 833–45.

27 Foa EB, Coles M, Huppert JD, Pasupuleti RV, Franklin ME, March J. Development and validation of a child version of the obsessive compulsive inventory. *Behav Ther* 2010; **41:** 121–32.

28 Posner K, Brown GK, Stanley B, et al. The Columbia-Suicide Severity Rating Scale: initial validity and internal consistency findings from three multisite studies with adolescents and adults. *Am J Psychiatry* 2011; **168:** 1266–77.

29 Ravens-Sieberer U, Erhart M, Rajmil L, et al. Reliability, construct and criterion validity of the KIDSCREEN-10 score: a short measure for children and adolescents' well-being and health-related quality of life. *Qual Life Res* 2010; **19:** 1487–500.

30 Shaffer D, Gould MS, Brasic J, et al. A children's global assessment scale (CGAS). *Arch Gen Psychiatry* 1983; **40:** 1228–31.

31 Lyneham HJ, Sburlati ES, Abbott MJ, et al. Psychometric properties of the Child Anxiety Life Interference Scale (CALIS). *J Anxiety Disord* 2013; **27:** 711–19.

32 Klaufus L, Verlinden E, van der Wal M, Kösters M, Cuijpers P, Chinapaw M. Psychometric evaluation of two short versions of the Revised Child Anxiety and Depression Scale. *BMC Psychiatry* 2020; **20:** 47.

33 Bagheri Z, Akbari M, Jafari P, Huedo-Medina TB. Assessing the measurement invariance of the RCADS-25 questionnaire across gender and child-parent dyads in the presence of multilevel data. *Qual Life Res* 2019; **28:** 829–40.

34 Rodríguez-Jiménez T, Piqueras JA, Lázaro L, Moreno E, Ortiz AG, Godoy A. Metric invariance, reliability, and validity of the Child Version of the Obsessive Compulsive Inventory (OCI-CV) in community and clinical samples. *J Obsessive Compuls Relat Disord* 2016; **9:** 1–8.

35 Rosa-Alcázar AI, Ruiz-García B, Iniesta-Sepúlveda M, López-Pina JA, Rosa-Alcázar Á, Parada-Navas JL. Obsessive Compulsive Inventory-Child Version (OCI-CV) in a Spanish community sample of children and adolescents. *Psicothema* 2014; **26:** 174–79.

36 Martínez-González AE, Rodríguez-Jiménez T, Piqueras JA, Vera-Villarroel P, Godoy A. Psychometric properties of the obsessive-compulsive Inventory-Child Version (OCI-CV) in Chilean children and adolescents. *PLoS One* 2015; **10:** e0136842.

37 Jones AM, De Nadai AS, Arnold EB, et al. Psychometric properties of the obsessive compulsive inventory: child version in children and adolescents with obsessive-compulsive disorder. *Child Psychiatry Hum Dev* 2013; **44:** 137–51.

38 Opakunle T, Aloba O, Akinsulore A. Obsessive Compulsive Inventory - Child Version (OCI-CV): confirmatory factor analysis, reliability, validity and correlates among Nigerian adolescents. *Malawi Med J* 2018; **30:** 262–69.

39 Verlinden E, van Meijel EPM, Opmeer BC, et al. Characteristics of the Children's Revised Impact of Event Scale in a clinically referred Dutch sample. *J Trauma Stress* 2014; **27:** 338–44.

40 Magalhães SS, Miranda DK, de Paula JJ, de Miranda DM, Romano-Silva MA, Malloy-Diniz LF. Psychometric properties of a Brazilian Portuguese version of the Children's Revised Impact of Event Scale (CRIES-8). *Rev Psiquiatr Clin (Santiago)* 2018; **45:** 27–32.

41 Salari R, Malekian C, Linck L, Kristiansson R, Sarkadi A. Screening for PTSD symptoms in unaccompanied refugee minors: a test of the CRIES-8 questionnaire in routine care. *Scand J Public Health* 2017; **45:** 605–11.

42 Gigengack MR, Hein IM, Lindeboom R, Lindauer RJL. Increasing resource parents' sensitivity towards child posttraumatic stress symptoms: a descriptive study on a trauma-informed resource parent training. *J Child Adolesc Trauma* 2017; **12:** 23–29.

43 Kilincaslan A, Gunes A, Eskin M, Madan A. Linguistic adaptation and psychometric properties of the Columbia-Suicide Severity Rating Scale among a heterogeneous sample of adolescents in Turkey. *Int J Psychiatry Med* 2019; **54:** 115–32.

44 Nik-Azin A, Shairi MR, Naeinian MR, Sadeghpour A. The health-related quality of life index KIDSCREEN-10: confirmatory factor analysis, convergent validity and reliability in a sample of Iranian students. *Child Indic Res* 2014; **7:** 407–20.

45 Stevanovic D, Tadic I, Novakovic T, Kisic-Tepavcevic D, Ravens-Sieberer U. Evaluating the Serbian version of the KIDSCREEN quality-of-life questionnaires: reliability, validity, and agreement between children's and parents' ratings. *Qual Life Res* 2013; **22:** 1729–37.

46 Weissman MM, Warner V, Fendrich M. Applying impairment criteria to children's psychiatric diagnosis. *J Am Acad Child Adolesc Psychiatry* 1990; **29:** 789–95.

47 Bird HR, Canino G, Rubio-Stipec M, Ribera JC. Further measures of the psychometric properties of the Children's Global Assessment Scale. *Arch Gen Psychiatry* 1987; **44:** 821–24.

48 Dyrborg J, Larsen FW, Nielsen S, Byman J, Nielsen BB, Gautrè-Delay F. The Children's Global Assessment Scale (CGAS) and Global Assessment of Psychosocial Disability (GAPD) in clinical practice—substance and reliability as judged by intraclass correlations. *Eur Child Adolesc Psychiatry* 2000; **9:** 195–201.

49 Sourander A, Helenius H, Piha J. Child psychiatric short-term inpatient treatment: CGAS as follow-up measure. *Child Psychiatry Hum Dev* 1996; **27:** 93–104.

50 Orgilés M, Fernández-Martínez I, Morales A, Melero S, Espada JP. Spanish Validation of the Child Anxiety Life Interference Scale (CALIS-C): psychometric properties, factorial structure and factorial invariance across gender. *Child Psychiatry Hum Dev* 2019; **50:** 756–63.

51 Marques T, Pereira AI, Pedro M, Russo V, Goes AR, Barros L. Preliminary analysis of the Portuguese version of the child anxiety life interference scale. *Psicol Reflex Crit* 2015; **28:** 454–62.

52 Chorpita BF, Yim L, Moffitt C, Umemoto LA, Francis SE. Assessment of symptoms of DSM-IV anxiety and depression in children: a revised child anxiety and depression scale. *Behav Res Ther* 2000; **38:** 835–55.

53 Piqueras JA, Martín-Vivar M, Sandin B, San Luis C, Pineda D. The Revised Child Anxiety and Depression Scale: a systematic review and reliability generalization meta-analysis. *J Affect Disord* 2017; **218:** 153–69.

54 Yule W. Anxiety, depression and post-traumatic stress in childhood. In: Sclare I, ed. Child Psychology Portfolio. Windsor, England: NFER-Nelson, 1997.

55 Hoagwood K, Jensen PS, Petti T, Burns BJ. Outcomes of mental health care for children and adolescents: I. A comprehensive conceptual model. *J Am Acad Child Adolesc Psychiatry* 1996; **35:** 1055–63.

56 Rapee RM, Bőgels SM, van der Sluis CM, Craske MG, Ollendick T. Annual research review: conceptualising functional impairment in children and adolescents. *J Child Psychol Psychiatry* 2012; **53:** 454–68.

57 Ravens-Sieberer U, Herdman M, Devine J, et al. The European KIDSCREEN approach to measure quality of life and well-being in children: development, current application, and future advances. *Qual Life Res* 2014; **23:** 791–803.

58 United Nations Educational, Scientific and Cultural Organization. International Standard Classification of Education: ISCED 2011. 2012. http://uis.unesco.org/sites/default/files/documents/international-standard-classification-of-education-isced-2011-en.pdf (accessed June 18, 2020).

59 Southam-Gerow MA, Kendall PC, Weersing VR. Examining outcome variability: correlates of treatment response in a child and adolescent anxiety clinic. *J Clin Child Psychol* 2001; **30:** 422–36.

60 Berman SL, Weems CF, Silverman WK, Kurtines WM. Predictors of outcome in exposure-based cognitive and behavioral treatments for phobic and anxiety disorders in children. *Behav Ther* 2000; **31:** 713–31.

61 Liber JM, van Widenfelt BM, van der Leeden AJM, Goedhart AW, Utens EMWJ, Treffers PDA. The relation of severity and comorbidity to treatment outcome with Cognitive Behavioral Therapy for childhood anxiety disorders. *J Abnorm Child Psychol* 2010; **38:** 683–94.

62 Emslie GJ, Kennard BD, Mayes TL. Predictors of treatment response in adolescent depression. *Pediatr Ann* 2011; **40:** 300–06.

63 Lundkvist-Houndoumadi I, Hougaard E, Thastum M. Pre-treatment child and family characteristics as predictors of outcome in cognitive behavioural therapy for youth anxiety disorders. *Nord J Psychiatry* 2014; **68:** 524–35.

64 Brent DA, Kolko DJ, Birmaher B, et al. Predictors of treatment efficacy in a clinical trial of three psychosocial treatments for adolescent depression. *J Am Acad Child Adolesc Psychiatry* 1998; **37:** 906–14.

65 Rapee RM, Lyneham HJ, Hudson JL, Kangas M, Wuthrich VM, Schniering CA. Effect of comorbidity on treatment of anxious children and adolescents: results from a large, combined sample. *J Am Acad Child Adolesc Psychiatry* 2013; **52:** 47–56.

66 Storch EA, Merlo LJ, Larson MJ, et al. Impact of comorbidity on cognitive-behavioral therapy response in pediatric obsessive-compulsive disorder. *J Am Acad Child Adolesc Psychiatry* 2008; **47:** 583–92.

67 Skriner LC, Chu BC, Kaplan M, et al. Trajectories and predictors of response in youth anxiety CBT: integrative data analysis. *J Consult Clin Psychol* 2019; **87:** 198–211.

68 Hudson JL, Keers R, Roberts S, et al. Clinical predictors of response to cognitive-behavioral therapy in pediatric anxiety disorders: the genes for treatment (GxT) study. *J Am Acad Child Adolesc Psychiatry* 2015; **54:** 454–63.

69 Jones M, Hopkins K, Kyrke-Smith R, Davies R, Vostanis P, Wolpert M. Current view tool: completion guide. London, UK: CAMHS Press, 2013.

70 WHO. International classification of diseases for mortality and morbidity statistics (11th Revision). 2018. https://icd.who.int/browse11/l-m/en (accessed June 18, 2020).

71 Bodden DHM, Bögels SM, Nauta MH, et al. Child versus family cognitive-behavioral therapy in clinically anxious youth: an efficacy and partial effectiveness study. *J Am Acad Child Adolesc Psychiatry* 2008; **47:** 1384–94.

72 Aggelosopoulou A, Roussi-Vergou C, Zafeiropoulou M. "Anxioussaure": a CBT psychoeducational programme for children's anxiety. Programme development and pilot study. *Psychol J Hell Psychol Soc* 2012; **19:** 245–62.

73 Legerstee JS, Huizink AC, van Gastel W, et al. Maternal anxiety predicts favourable treatment outcomes in anxiety-disordered adolescents. *Acta Psychiatr Scand* 2008; **117:** 289–98.

74 Nanni V, Uher R, Danese A. Childhood maltreatment predicts unfavorable course of illness and treatment outcome in depression: a meta-analysis. *Am J Psychiatry* 2012; **169:** 141–51.

75 Batty MJ, Moldavsky M, Foroushani PS, et al. Implementing routine outcome measures in child and adolescent mental health services: from present to future practice. *Child Adolesc Ment Health* 2013; **18:** 82–87.

76 Chevance A, Tran VT, Ravaud P. Improving the generalizability and credibility of core outcome sets (COSs) by a large and international participation of diverse stakeholders. *J Clin Epidemiol* 2020; **125:** 206–12.

77 Davis K, Gorst SL, Harman N, et al. Choosing important health outcomes for comparative effectiveness research: an updated systematic review and involvement of low and middle income countries. *PLoS One* 2018; **13:** e0190695.

78 Kieling C, Rohde LA. Child and adolescent mental health research across the globe. *J Am Acad Child Adolesc Psychiatry* 2012; **51:** 945–47.

79 Trauer T. Outcome measurement in mental health: theory and practice. Cambridge, UK: Cambridge University Press, 2010.

80 Delespaul PAEG. Routine outcome measurement in the Netherlands—a focus on benchmarking. *Int Rev Psychiatry* 2015; **27:** 320–28.

81 Department of Health and Ageing. Mental health national outcomes and casemix collection: overview of clinician-rated and consumer self-report measures V1.50. 2003. https://www.amhocn.org/sites/default/files/publication_files/mhnocc_overview_doha.pdf (accessed June 18, 2020).

82 National Quality Forum. Measure applications partnership. 2019. https://www.qualityforum.org/Setting_Priorities/Partnership/Measure_Applications_Partnership.aspx (accessed June 22, 2020).

83 Monga S, Monsour A, Stallwood E, et al. Core outcome set development for adolescent major depressive disorder clinical trials: a registered report. *J Am Acad Child Adolesc Psychiatry* 2020; **59:** 1297–98.

84 UNICEF. Measurement of Mental Health Among Adolescents at the Population Level (MMAP): an Overview. 2019. https://data.unicef.org/wp-content/uploads/2018/02/Mental-health-measurement_MMAP_overview_UNICEF.pdf (accessed Dec 6, 2019).

85 Myers K, Winters NC. Ten-year review of rating scales. II: scales for internalizing disorders. *J Am Acad Child Adolesc Psychiatry* 2002; **41:** 634–59.

86 Winters NC, Collett BR, Myers KM. Ten-year review of rating scales, VII: scales assessing functional impairment. *J Am Acad Child Adolesc Psychiatry* 2005; **44:** 309–42.

87 Becker-Haimes EM, Tabachnick AR, Last BS, Stewart RE, Hasan-Granier A, Beidas RS. Evidence base update for brief, free, and accessible youth mental health measures. *J Clin Child Adolesc Psychol* 2020; **49:** 1–17.

88 Beidas RS, Stewart RE, Walsh L, et al. Free, brief, and validated: standardized instruments for low-resource mental health settings. *Cognit Behav Pract* 2015; **22:** 5–19.

89 Mash EJ, Hunsley J. Evidence-based assessment of child and adolescent disorders: issues and challenges. *J Clin Child Adolesc Psychol* 2005; **34:** 362–79.

90 Reeve BB, Thissen D, DeWalt DA, et al. Linkage between the PROMIS® pediatric and adult emotional distress measures. *Qual Life Res* 2016; **25:** 823–33.

91 Sales CMD. Seeing the person in the patient: making the case for individualized PROMs in mental health care. *Curr Psychiatry Rev* 2017; **13:** 184–87.

92 The Lancet Psychiatry. A good enough measure. *Lancet Psychiatry* 2020; **7:** 825.

93 Milfont TL, Fischer R. Testing measurement invariance across groups: applications in cross-cultural research. *Int J Psychol Res (Medellin)* 2010; **3:** 111–30.