RUNNING HEAD: MEASURING RESPONSE TO CLINICAL CARE

Measuring Response to Clinical Care in Children and Young People with Anxiety, Depression, OCD or PTSD: An International Standard Set of Outcome Measures

Karolin Krause, Sophie Chung, Abiodun O. Adewuya, Anne Marie Albano, Rochelle Babins-Wagner, Peter Brann, Laura Birkinshaw, 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, Evie Thomas, George C. Patton, Ulrike Ravens-Sieberer, Peter Szatmari, Lucas Walters, Bryan Young, Yue Zhao, Miranda Wolpert

Author Affiliations

Evidence Based Practice Unit, University College London and Anna Freud National Centre for Children and Families, London, United Kingdom (KK, PhD)

International Consortium for Health Outcomes Measurement (ICHOM), London, United Kingdom (SC, BA)

Lagos State University College of Medicine, Ikeja, Lagos, Nigeria (AOA, Prof, FWACP)

New York State Psychiatric Institute and Columbia University Vagelos College of Physicians and Surgeons, Columbia University, New York, New York, USA (AMA, Prof, PhD)

Calgary Counselling Centre, Calgary, AB, Canada; Faculty of Social Work, University of Calgary, Calgary, Alberta, Canada (RBW, PhD)

Eastern Health Child and Youth Mental Health Service & Monash University, Melbourne, Victoria, Australia (PB, PhD)

Young Minds, London, United Kingdom (LB, HBA)

Departments of Experimental Psychology and Psychiatry, University of Oxford, Oxford, United Kingdom (CC, Prof, PhD)

Rush College of Nursing, Chicago, Illinois, USA (KD, Prof, PhD)

Université Paris-Saclay, Université de Versailles Saint-Quentin-en-Yvelines, Institut national de la santé et de la recherche médicale (INSERM), Centre de recherche en Epidémiologie et Santé des Populations (CESP), Villejuif, Île-de-France, France (BF, MD & PhD)

Children's Hospital of Philadelphia, Philadelphia, Pennsylvania, USA (CBF, Prof, PhD)

Centre for Emotional Health, Department of Psychology, Macquarie University, Sydney, New South Wales, Australia (JLH, Professor, PhD)

Faculty of Psychology, Doshisha University, Kyoto, Japan (SI, Prof, PhD)

Department of Psychology, University of Delhi, New Delhi, Delhi, India (MK, BA)

Hospital de Clínicas de Porto Alegre, Universidade Federal do Rio Grande do Sul, Porto Alegre, Rio Grande do Sul, Brazil (CK, MD)

Child and Youth Mental Health Service, Children's Health Queensland Hospital and Health Service, Brisbane, Queensland, Australia (JK, MA)

Sangath, New Delhi, Delhi, India (KM, PhD)

Centro de Medicina Reproductiva y Desarrollo Integral del Adolescente (CEMERA), Facultad de Medicina, Universidad de Chile; Millennium Nucleus to Improve the Mental Health of Adolescents and Youths (Imhay); Millennium Institute for Research in Depression and Personality (MIDAP), Santiago de Chile, Región Metropolitana, Chile (VM, Prof, PhD)

School of Primary, Community and Social Care, Keele University, Keele ,Staffordshire, United Kingdom (FM, MBChB)

Child Study Center, Department of Psychology, Virginia Tech, Blacksburg, Virginia, USA (THO, Prof, PhD)

Department of Child and Adolescent Psychiatry, Institute of Mental Health, Singapore, Singapore (SHO, MMED)

Eastern Health Child Youth Mental Health Service, Victoria, Australia (ET, BA)

Murdoch Children's Research Institute, Melbourne; Department of Paediatrics, The University of Melbourne, Melbourne, Victoria, Australia (GCP, Prof, MD)

Department of Child and Adolescent Psychiatry, Psychotherapy, and Psychosomatics, University Medical Center Hamburg-Eppendorf, Hamburg, Germany (URS, Prof, PhD)

The Hospital for Sick Children, the University of Toronto; the Cundill Centre for Child and Youth Depression, Centre for Addiction and Mental Health, Toronto, Ontario, Canada (PS, Prof, MD)

Jack.org, Ottawa, Ontario, Canada (LW, B.Sc.; BY)

Teaching and Learning Evaluation and Measurement Unit, The University of Hong Kong, Hong Kong (YZ, EdD)

Wellcome Trust, London, Greater London, United Kingdom (MW, Prof, DClinPsy)

Corresponding Author: Karolin Krause, Evidence Based Practice Unit, Anna Freud National Centre for Children and Families, The Kantor Centre of Excellence, 4-8 Rodney Street, London N1 9JH, United Kingdom; karolin.krause.16@ucl.ac.uk, +44 7850 296780.

[Accepted for publication by the Lancet Psychiatry on 03/08/2020]

Contributors

KK, MW, and SC oversaw the study design. KK and SC led data acquisition through the literature reviews, Delphi surveys, and Open Review surveys, as well as data analysis and interpretation. KK takes responsibility for the integrity of the data and the accuracy of data analysis. All authors attended teleconferences, and all but the core project team (KK, MW, SC) completed anonymous votes and feedback surveys as part of the Delphi process and formed consensus on the final recommendation. KK 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 across Australia of routine outcome measurement in public mental health. He chairs the National Mental Health Child and Adolescent Information Development Expert Advisory Group. There is an interest in supporting routine outcome measurement and benchmarking between organisations. This does not influence the content of the submitted work. SC was an employee of the International Consortium for Health Outcomes Measurement (ICHOM) during the conduct of this study. ICHOM received grants from the Government of New South Wales Agency for Clinical Innovation, Australia; Providence Health & Services, USA; Västra Götaland Regional Council, Sweden; and National Health Service, England, United Kingdom, in support of this work. CC has conducted 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 E. Lilly, BMS, Servier, SANOFI, GSK, HRA, Roche, Boeringer Ingelheim, Bayer, Almirall, Allergan, Stallergene, Genzyme, Pierre Fabre, AstraZeneca, Novartis, Janssen, Astellas, Biotronik, Daiichi- Sankyo, Gilead, MSD, Lundbeck, Stallergene, 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. SI has participated in several studies which 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. KK received personal fees from the International Consortium for Health Outcomes Measurement (ICHOM) during the conduct of the study. URS 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 is head of the new Mental Health Priority Area at the Wellcome Trust, which may 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 advisor to NHS England on informatics. AOA, RBW, LB, KD, CBF, MK, CK, JK, FM, KM, VM, THO, SHO, ET, GCP, PS, LW, BY, YZ have nothing to disclose.

Acknowledgments

The authors 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. The authors would also like to thank Luz Sousa Fialho for providing administrative support, Katherine G. Young for providing research support, and Christiana Ilesanmi for providing both administrative and research support to this project, as well as Zofia Das-Gupta for providing guidance on the application and management of the ICHOM process until May 2019. The authors would like to thank one further service user representative who participated in the early stages of the process, including the Delphi process, but did not wish to be a named co-author.

Funding

This work has been funded by the National Health Service (NHS) England; Västra Götaland Regional Council, Sweden; the Government of New South Wales Agency for Clinical Innovation, Australia; and Providence Health & Services, USA. The funders had no role in the study design, data analysis, interpretation, or writing of the report. The funders did support 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 (RP_2014-04-018). FM was supported by an NIHR In-Practice Fellowship (IPF-2017-11-002), and is supported by a NIHR School for Primary Care Research GP Career Progression Fellowship. The views expressed in this manuscript of those of the authors and not necessarily those of the NHS, NIHR, or Department for Health and Social Care. VM received funding by the Millennium Science Initiative of the Ministry of Economy, Development and Tourism, grant Millennium Nucleus to Improve the Mental Health of Adolescents and Youths, Imhay, and the Innovation Fund for Competitiveness [Fondo de Innovación para la Competitividad, FIC], Project IS130005. CK is a Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) researcher and an Academy of Medical Sciences Newton Advanced Fellow.

Keywords

Anxiety; Depression; Stress Disorders, Post-Traumatic; Obsessive-Compulsive Disorder; Patient Reported Outcomes; Outcome Measures; Patient-Relevant Outcome; Quality Improvement; Reference Standards; Delphi Technique; Child Health; Adolescent Health; Core Outcome Set.

Wordcount

4760 words

Abstract

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 PHQ-9, GAD-7 and WHODAS 2.0). This paper reports on recommendations specifically for anxiety, depression, obsessive-compulsive disorder (OCD), and post-traumatic stress disorder (PTSD) in children and young people (CYP) aged 6-24 years. An international ICHOM working group of 27 clinical, research, and lived-experience experts formed consensus through teleconferences, a Delphi exercise, and iterative anonymous voting. A systematic review identified 70 possible outcomes and 107 relevant measurement instruments. Measures were appraised for their feasibility in routine practice (i.e., brevity, free availability, validation in CYP, translation) and psychometric performance (i.e., validity, reliability, 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 RCADS-25, OCI-CV, CRIES-8/13, C-SSRS, KIDSCREEN-10, CGAS, and CALIS. 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.

Measuring Response to Clinical Care in Children and Young People with Anxiety, Depression, OCD or PTSD: An International Standard Set of Outcome 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 over recent decades, service systems have failed to reduce the prevalence of these disorders in children and young people (CYP; 5). The global response 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 on treatment outcome is an essential facet in evaluating care effectiveness and can inform the setting of strategic targets for health systems, comparisons between systems and 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 CYP. Uptake of routine outcome measurement remains low, and where it occurs, there is considerable variation in outcomes, instruments, and assessment time points, with a recent review recording 20 different measures used to assess primary outcome across 38 studies of routine treatment for youth anxiety and depression (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 service user representatives, practitioners, and researchers to build consensus on a set of outcome domains, measurement instruments, case-mix factors (i.e., 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 to recommend.

Panel 1. The ICHOM Approach to Standard Set Development

ICHOM is a non-profit organisation specialised 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 healthcare, where 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 service user needs and treatment options 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; balance a comprehensive approach to tracking outcome with a feasible recommendation that services can reliably implement; and be responsive to quality improvement efforts.

An existing ICHOM Standard Set for adult anxiety and depression covers young people from the age of 14 years (12; Appendix pg. 3). It includes a recommendation for tracking symptom change via the Generalized 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 World Health Organization Disability Schedule 2.0 (WHODAS 2.0) 12-item short form (17; the full recommendation is available from ichom.org). The present Standard Set aims to complement this effort by providing a set of recommendations specifically tailored for use with CYP. The combination of the two Sets will provide for transition from youth into adult services at any point between the ages of 14 and 24 years, allowing for local variation in transition and judgements about which Set is most suitable for different ages.

This Standard Set is designed for CYP aged 6 to 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 levels of negative affect, with OCD and PTSD long classified as anxiety disorders (18). It 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-, parent-, and clinician-reported outcome measures, to account for the different perspectives these reporters tend to provide (19-21). Parent-report also serves as a proxy where CYP are unable to complete measures due to young age or developmental constraints (21). A more detailed discussion of the scope of this Standard Set is provided in the Appendix (pg. 3).

Panel 2. Methodology

The Working Group comprised 27 experts from 13 countries, including service users, parents or carers (henceforth "parents"), mental health practitioners and researchers working within relevant disciplines (e.g., psychiatry, epidemiology, psychometrics; see Appendix pg. 3-4). A central project team (KK, SC, MW) coordinated and facilitated the consensus-building process and completed supporting research tasks, but did not vote on the consensus recommendation. The Appendix (pg. 3-18) provides a detailed description of the process and methodology. A flow chart is provided in figure 1.

General process. Over 14 months, the Working Group completed a structured and evidence-informed consensus-building process (figure 1). The group convened for eight teleconferences, completed a three-round Delphi exercise 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 time points (Appendix pg. 4-5). The process was informed by sequential research inputs performed by the central project team (Appendix pg. 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 review of 257 treatment outcome studies, a narrative review of supplemental sources (e.g., cohort studies; qualitative outcome research, instrument banks), and break-out groups with service user representatives (Appendix pg. 5-8)

Selection of measurement instruments. Once the Group reached consensus on outcomes for inclusion, 107 thematically relevant measures were systematically appraised to identify those most suitable for tracking the selected outcomes over time. Appraisal criteria included *relevance*; (i.e., comprehensive coverage of the selected outcome domain); *feasibility* and *acceptability* (i.e., completion within less than 20 minutes; free availability for use in clinical settings, including in paper-and-pencil format; prior validation in CYP; translation into at least a second language); and *psychometric performance* (i.e., interrater or retest reliability, and internal consistency above 0.70; evidence of sensitivity to change), in line with International Society for Quality of Life Research (ISOQOL) recommendations (25; Appendix pg 8-16). The instruments judged to best satisfy these criteria were shortlisted, and their relative 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 time points. Based on the systematic review and existing ICHOM Standard Sets, the central project team compiled a list of possible case mix factors, and conducted a rapid review

of reviews examining predictors, mediators, and moderators of treatment response (Appendix pg. 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 time points, 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 policymakers from 45 countries, as well as 24 young people and parents from Denmark, the United States (US), and the United Kingdom (UK; see Appendix pg. 18).

Recommendations of the Working Group

Outcomes and Measures

As per Working Group consensus, the Standard Set recommends tracking response to treatment across the three outcome domains of *symptoms*, *suicidal thoughts and behaviour*, and *functioning*, using seven primarily self-reported instruments. These were selected for fulfilling most or all of the appraisal criteria (table 1). A detailed discussion of instrument properties, performance against appraisal criteria, and accessibility is provided in the Appendix (pg. 19-22). During the Open Review, 75% of the 463 practitioners and researchers who provided feedback stated their overall confidence in the recommended outcomes and instruments and participants with lived experience confirmed the importance of outcomes included (75-100% of participants rated each included outcome domain as important) and the acceptability of the recommended measures (100% of 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.

Insert Table 1

Symptoms. The Set recommends measuring anxiety and depression symptoms through youth and parent-report for all CYP in scope. 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 Set recommends the 25-item short form of the Revised Children's Anxiety and Depression Scale (RCADS-25; 26) for youth and parent-report. Based on the appraisal criteria, the group initially selected the scale's 47-item version (27), but decided to recommend the short form to reduce respondent burden. The RCADS-25 consists of 15 items tracking anxiety symptoms, and 10 items tracking 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 (28).

Symptoms of OCD and PTSD should be tracked separately for CYP 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 Compulsive Inventory for Children (OCI-CV; 29); and the Children's Revised Impact of Events Scale (CRIES; 30) for youth (CRIES-8) and parent report (CRIES-13). Both measures have been applied in Europe, Asia, and the Americas. A parent version of the OCI-CV is not currently available.

It is important to note that the above-mentioned 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 state-of-the-art diagnostic tools. The latter require additional properties related to diagnostic validity (e.g., sensitivity and specificity), which the Working Group did not explicitly consider during measure appraisal.

Suicidal thoughts and behaviour. Consensus was reached on measuring suicidal thoughts and behaviour in all young people aged 10 years and older (unless considered inappropriate) using the Columbia Suicide Severity Rating Scale (C-SSRS) Recent Self-Report Screener (31) – a short self-report version of the clinician-administered C-SSRS interview protocol. The measure consists of six items tracking the severity of suicidal ideation and behaviour in the previous month. The self-report screener has not been validated in CYP, but the clinician-rated C-SSRS has demonstrated good internal consistency, interrater reliability, and sensitivity to change in adolescent samples (31).

Functioning. Functioning describes a child's ability to engage in typical activities and meet role demands in line with age-specific socio-cultural norms (32,33). None of the identified functioning measures fully satisfied the Working Group's requirements, and consensus was reached on mitigating this 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, while condition-specific measures may 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 (i.e., item 8: *I have trouble sleeping*; and item 9, *I feel scared if I have to sleep on my own*).

The KIDSCREEN-10 was selected as a generic measure of global functioning, to be completed by CYP and parents. This unidimensional 10-item index of HRQoL tracks functioning in relation to physical health and energy levels, leisure activities, social and family relationships, and cognition (34). The more comprehensive KIDSCREEN-52 was originally developed through a process of cross-cultural harmonisation involving 13 European countries. Its 10-item short form has been applied in Asia, Eastern Africa, Europe, North and Latin America. Although originally designed for epidemiological studies, the KIDSCREEN-10 has been shown to discriminate well between CYP with high and low levels of functioning, with few ceiling or floor effects (34). In addition to the KIDSCREEN-10, the Children's

Global Assessment Scale (CGAS) was selected as a brief clinician-rated measure of global functioning (35). On this widely used measure, clinicians preform a single rating by locating CYP on a scale from 1 to 100, placing them into one of ten categories, from 1 – 10 (*extremely impaired*) to 91 – 100 (*doing very well*).

The Children's Anxiety Life Interference Scale (CALIS; 36) was selected to measure condition-specific impairment via 9 items (10 for the parent version) that describe instances of anxiety impacting on functioning at home, at school, on social life, or on activities. The CALIS currently only covers anxiety-related impairment, and the Working Group did not identify any eligible measure that tracks depression-specific impairment, or captures impairment from both anxiety and depression. However, the CALIS author team recently revised the measure to cover impairment from anxiety *and* depression, and a validation study involving children, young people, and parents in community, school, and clinical settings is ongoing. As soon as validation results 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.

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 Working Group consensus, services should record demographic, clinical, complexity, and intervention factors, through a mix of self-, parent-, and clinician-report (table 2). Beyond the Working Group, the suggested case-mix factors were endorsed by more than 80% of practitioners and researchers who provided feedback during the Open Review. These factors represent a minimum that should be assessed by all those providing relevant care, and services may wish to add other indicators to meet local information needs.

Insert Table 2

Demographic factors. The Set recommends recording age, gender, ethnic minority status, socio-economics status, and the child's living situation. Services should record the sex assigned at birth, and the gender reported by CYP. Socio-economic status should be measured by recording the highest level of education completed by any of the CYP's parents, as a widely accepted proxy that can be mapped onto the International Standard Classification of Education for international comparisons (ISCED; 37). The Set includes one question about CYP's living situation, and two questions capturing ethnic minority and marginalised group status via self-report (table 2).

Clinical factors. Several studies suggest that symptom burden, symptom duration, and the presence of comorbidities affect treatment response in CYP with depression or anxiety (38–47; Appendix pg. 17-18). The group recommends recording principal and comorbid presenting problems by administering the 30 problem descriptions of the Current View Tool (48). While not equivalent with formal diagnoses, these broadly align with the diagnostic categories of the

ICD-11 for pediatric populations (49). ICHOM standards are available for recording symptom duration and prior service use (table 2).

Complexity factors. Research suggests that parental mental health influences treatment response in CYP with anxiety, and depression (43,50-52; Appendix pg. 17-18). The Set includes two questions about experiences with or diagnoses of mental health problems in the immediate family, and prior use of mental health services by the reporting parent. While evidence is limited and conflicting on the influence of adverse experiences (e.g., childhood maltreatment) on treatment response (43,53), 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 intervention focus (i.e., in terms of who is actively involved), treatment modality and prescribed medication (i.e., as per lists of options compiled by the Working Group and through Open Review feedback), and intervention setting (i.e., in terms of whether or not treatment was delivered via a digital platform or inpatient care, as opposed to other settings). Services may wish to record additional detail on intervention characteristics to meet local, regional, or national information needs.

Measurement Time Points

Timelines for measurement are highly practical considerations, 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 local or national data collection. The suggested timepoints were widely endorsed by over 80% of practitioners and researchers participating in the Open Review.

As per Working Group consensus, all case-mix factors and outcomes should be measured at assessment or intake (i.e., baseline), or as near to these time points as possible where services wish to collect data at second contact (figure 2). As a guideline, all outcome measures should be administered every three months following baseline. Services are encouraged to consider more frequent intervals, including session-by-session measurement, to help embed monitoring into the clinical process (54), and reduce the risk of missing data due to drop-out prior to follow-up. The group recognises that effective session-by-session measurement requires well established systems and can otherwise be perceived as unduly 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 (e.g., 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, one year after baseline. This

can enable important insights into longer-term outcomes, but may require significant adjustments to the way services are currently organised and funded.

Insert Figure 2.

Strengths of the Standard Set

Of the Working Group's 27 voting members, four were young adults between the ages of 18 and 24, and two were parents. This was the first ICHOM Working Group to involve young people with lived experience of service use – rather than just their parents – as full Working Group members at all stages of the consensus-building process, including the teleconferences, Delphi exercise, and subsequent surveys. While limited familiarity with outcome measurement has been described as a barrier to the meaningful participation of lived-experience experts (55), 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 towards the end of the consensus-building process, through the Open Review. 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 led to a comparatively small sample of Open Review participants with lived experience (N = 24). 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- and middle-income countries (LMIC, Appendix pg. 3), who have rarely been represented by similar initiatives to date (56). While 90% of CYP live in LMICs, 90% of research on youth mental health currently comes from high-income countries (57). Experts from LMIC were also consulted through the Open Review (Appendix pg. 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 underscores its relevance and acceptability beyond the immediate Working Group.

Implementation

Working Group members will form a steering group to oversee the Standard Set's implementation, and to review recommendations in light of learning from pilot initiatives and new developments in the field. In health systems such as Australia, Canada, the Netherlands, the UK, or the US, the routine collection of outcome data is becoming increasingly embedded (58,59), which should facilitate the adoption of the Standard Set, but can also cause issues of alignment with existing local or national systems (e.g., 60,61). As services adopt routine outcome monitoring, it is imperative that data is 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 may or may not 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, along with the PHQ-9, GAD-7 and WHODAS for older youth. A Standard Set specifically for clinicial trials for adolescent depression is currently under development at the Hospital for Sick Children in Toronto (62). The United Nations Children's Fund (UNICEF; 63) is leading another complementary initiative focusing on measuring health outcomes for adolescents with anxiety or depression in population surveys. While there is a unique opportunity to encourage further harmonisation, a certain degree of divergence may persist in light of different priorities, processes, and methodologies (e.g., outcome measures for clinical trials do not need to be freely available). The present work further adds to existing review efforts to identify meaningful, feasible, and acceptable outcome measures suitable for use with CYP in practice settings (64-68). The steering group will consider opportunities for alignment with complementary initiatives, as well as with existing ICHOM Standard Sets, for example with a view to linking scores obtained from the CYP and adult Sets for the purpose of longitudinal analysis (e.g., 69).

Panel 3. How to Access the Standard Set Resources

A Standard Set reference guide, flyer, and data dictionary are available free of charge and can be accessed via ICHOM Connect (www.connect.ichom.org). The reference guide defines each outcome domain, and describes the recommended measures, case-mix factors, and time points.

Limitations and Areas for Future Research

The scope of this Standard Set is limited to anxiety, depression, OCD, and PTSD, and the Working Group has sought to make a *minimum* recommendation focusing on core outcomes for these target disorders. Clinical judgement is warranted in tracking additional symptoms over time that are not covered by the recommended outcome measures. Two complementary ICHOM Standard Sets are currently under development that will 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 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 relatively small Working Group of consistent membership, which convened at frequent intervals. There was not complete parity in representation across different strata of experts, including CYP and parents. In future efforts, large-scale Delphi-surveys may be more suitable for consulting equal numbers of different stakeholder groups, albeit at the expense of more in-depth and continuous group deliberation.

As per 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 (i.e., *coping* and *interference of treatment with daily life*) reached consensus amongst the group's service user representatives, but not within the wider group. To promote person-centred care, services may want to consider tracking additional outcomes based on shared decision-making with CYP and families, either through suitable standardised measures, or by including personalised outcome measures that track progress in relation to idiosyncratic presenting problems or treatment goals (70).

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 on their psychometric properties (e.g., 71). To maintain simplicity, the Set does not recommend separate measures for different age groups. While the selected measures have generally been validated in CYP aged 6-18 years, the Set is less specifically tailored to the experiences of 18-24-year-olds. 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. 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. While the Working Group considered sensitivity to change an essential measurement property to consider, its appraisal was limited by a general lack of data and objective appraisal guidelines. 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). While the Working Group considered such evidence where available, no standard thresholds could be applied to determine whether sensitivity was sufficient.

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 may impact on the validity of the selected measurement instruments (72). In the absence of invariance, comparisons between different groups are not fully meaningful, and the 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.

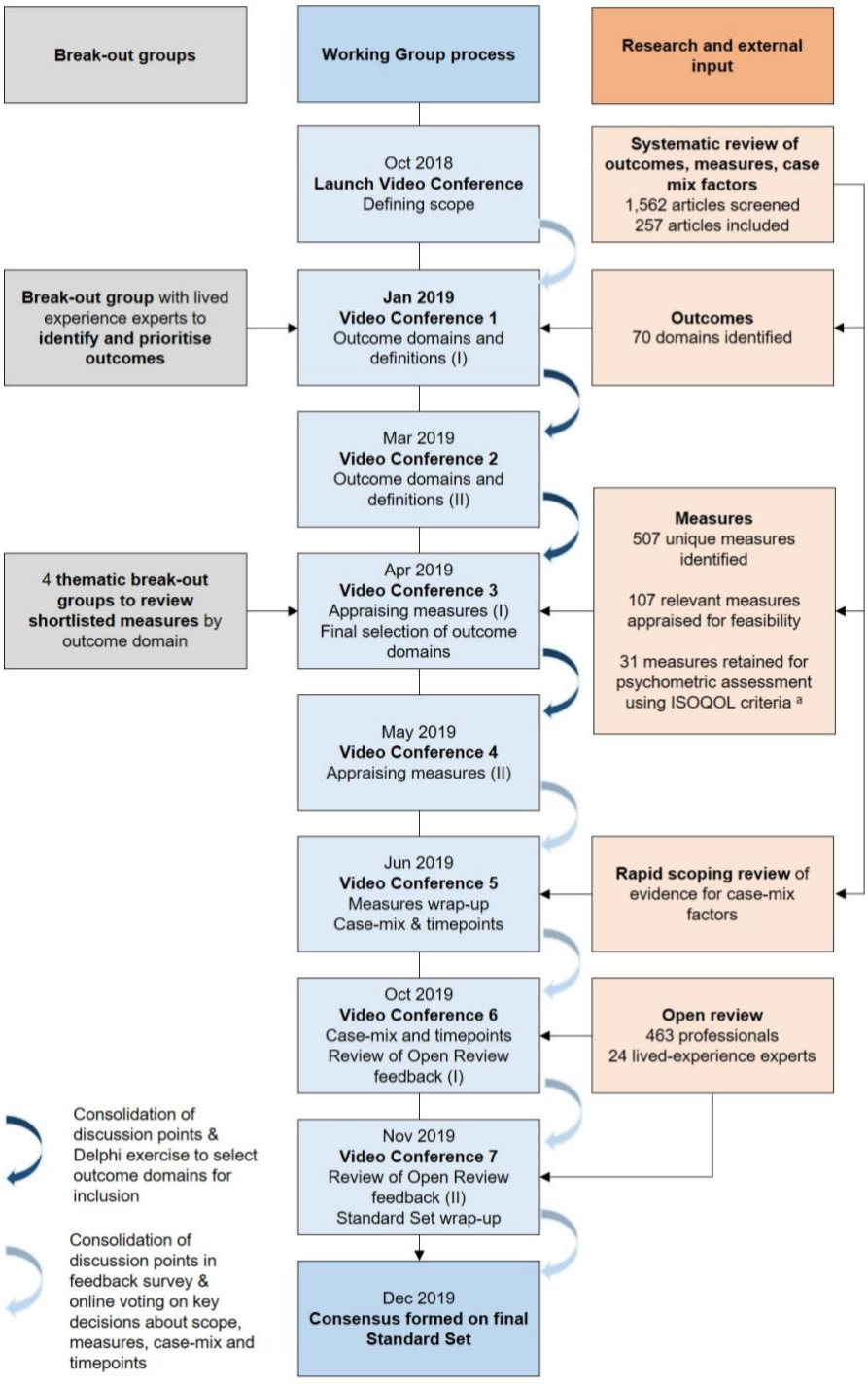
The Working Group encountered challenges with identifying suitable measures of functioning, with common issues including overlap in item coverage between symptoms and functioning, a perceived overemphasis on bodily functions as opposed to psychosocial functioning, lack of validation across the full age span or in 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, and in different real-world clinical settings (e.g. primary versus specialist care), and how acceptable the full question battery is 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 CYP 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 from 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 vital to ensure that this Standard Set makes a viable recommendation that meaningfully captures change for CYP across contexts.

Figure 1. Working Group Process



[a] 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).

Table 1. Overview of Recommended Outcome Measures and their Evaluation Criteria

		Relevance	**Feasibility & Acceptability**				**Psychometric Performance** [a]		
		Domain coverage	**Short**	**Free**	**Validated**	**Translated**	**Reliability**	**Validity**	**Sensitivity to change**
Outcome	Measure	Satisfactory domain coverage at subscale or item level	< 60 items / < 20 min	No licencing costs / restrictions	At least one validation in a CYP sample	>1 language version available	Test-retest reliability (TRT) or inter-rater reliability (IRR) ≥ 0.70	Internal consistency (e.g. Cronbach Alpha ≥ 0.70	Any evidence of sensitivity to change
Symptoms									
Anxiety & depression	RCADS-25	GAD; MDD; OCD; PD; SAD; SP	25 items (< 10 min)	Yes	Ages 6-18 cl. & non-cl.	4 /16 [b]	TRT (71)	Yes (26,73)	No evidence
OCD	OCI-CV	Doubting/checking; obsessing; hoarding; washing; ordering; neutralising	21 items (< 10 min)	Yes	Ages 6-18 cl. & non-cl.	3+	TRT (29,74–76)	Yes (29,74–78)	Some evidence (29)
PTSD	CRIES-8	Intrusion; avoidance (hyperarousal)	8 /13 items (< 5 min)	Yes	Ages 7-18 cl. & non-cl.	27+	TRT (79)	Yes (79–81)	Some evidence (82) [c]
Suicidal thoughts and behaviour	C-SSRS	Severity of ideation; behaviour (attempts)	3 or 6 items [d] (< 5 min)	Yes	Ages 12-18 b cl. & non-cl	100+	IRR (83) [e]	Yes (31) [e]	Some evidence (31) [e]
Functioning									
Global	KIDSCREEN-10	Physical activity & energy; emotions; leisure time and participation; relationships with parents & peers; cognition; school	10 items (< 5 min)	Yes	Ages 8-18 non-cl.	22+	TRT (34,84)	Yes (34,84,85)	No evidence
Global	CGAS	Global functioning	1 item (< 5 min)	Yes	Ages 4-18	2+	IRR (34,86–88) TRT (35,87)	N.A.	Some evidence (89)
Impact	CALIS	Enjoyable activities; relationships with siblings, parents, friends, peers; sports; schoolwork; distress	9/10 items (< 5 min)	Yes	Ages 6-17 cl. & non-cl.	7	TRT (36)	Yes (36,90,91)	Some evidence (36)

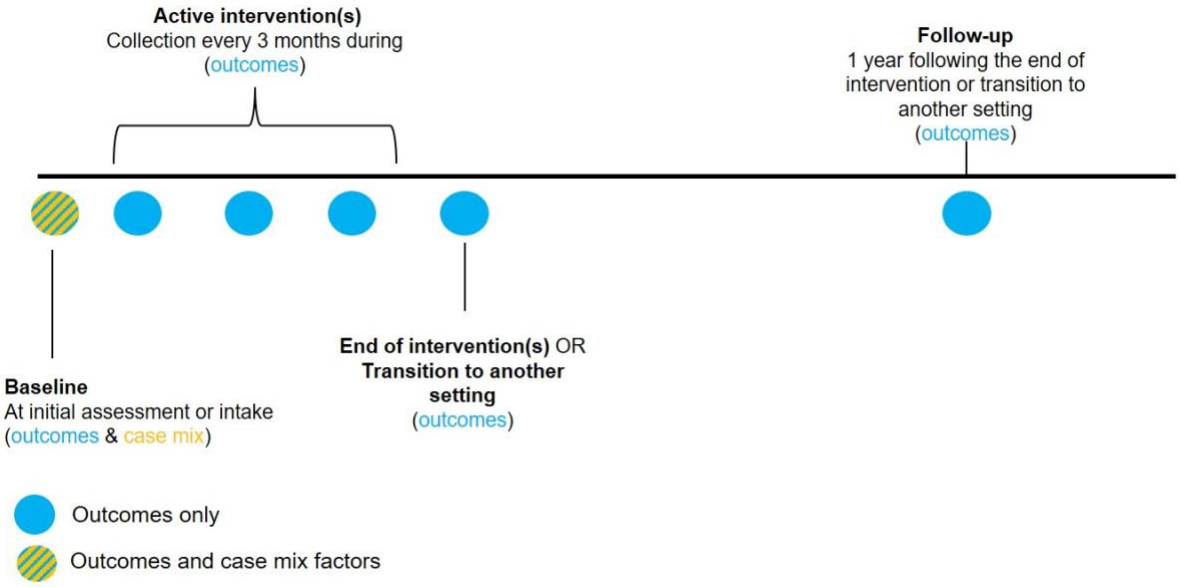
Note. GAD = Generalized Anxiety Disorder; MDD = Major Depressive Disorder (MDD); OCD = Obsessive Compulsive Disorder; PD = Panic disorder; SAD = Separation Anxiety Disorder; SP = Social Phobia. [a] These thresholds are based on the ISOQOL minimum standards for patient-reported outcome measures (25). [b] The RCADS-25 is currently available in four languages on the source website (see Table S12, Appendix pg. 22), but the 47-item long form is available in 16 languages, so that item-level translations are available in more than four languages for those items included in the RCADS-25. [b] These validation studies tested a parent-report version only. [c] These validation studies tested a parent-report version only. [d] 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. [e] To the authors' knowledge, the C-SSRS self-report screener had not been subject to a validation study in CYP at the time of writing. 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 2. Case-Mix Factors in the Standard Set for Pediatric Anxiety, Depression, OCD, and PTSD

	Case-mix factor	Measure	Reporter
Demographic factors	Age [a]	Year of birth	CYP / parent
	Sex [a]	Sex assigned at birth	CYP / parent
	Gender Identity	"Do you think of yourself as...?"	CYP
	Parent or carer education [a]	Highest level of education completed by any of the CYP's parents or carers (ISCED Standards)	CYP / parent
	Ethnicity	Do you consider yourself to be in an ethnic minority where you live?	CYP / parent
	Marginalised group status	Do you consider yourself to be a member of a marginalised group where you live?	CYP / parent
	Living situation	Which of the following people live with you [your child] at your [their] home?	CYP / parent
Clinical factors	Diagnoses and co-morbidities	Measured via the provisional Problems' list of the Current View Tool	Clinician
	Duration of symptoms [a]	For how many months have you [your child] been experiencing [specific condition] symptoms?	CYP / parent
	Prior 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, (c) other.	CYP / 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, substance abuse (for example, alcohol or drugs), schizophrenic disorder, personality disorder, somatoform disorder (unexplained physical symptoms), other)?	Parent
	Parental service use	Have you ever sought help for your mental health?	Parent
	Education/work difficulties	Measured via the 'Contextual Problems' of the Current View Tool	Clinician
Intervention variables	Intervention focus (i)	Who is actively involved in the intervention? Select all that apply.	Clinician
	Intervention focus (ii)	Is the intervention delivered to an individual child / family, or to a group of children / 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 (i)	Does this intervention involve an overnight stay at an institution providing mental health support?	Clinician
	Intervention setting (ii)	Does this intervention involve the use of a digital platform?	Clinician

[a] Variables defined and operationalised as per an existing ICHOM standard. More detailed descriptions of each variable, as well as response options can be consulted in the reference guide for the Standard Set, at https://www.ichom.org/portfolio/anxiety-depression-ocd-and-ptsd-in-children-and-young-people/.

Figure 2. Timeline for data collection



References

1. World Health Organization. Depression and Other Common Mental Disorders: Global Health Estimates [Internet]. Geneva; 2017. Available from: https://apps.who.int/iris/bitstream/handle/10665/254610/WHO-MSD-MER-2017.2-eng.pdf [10/01/2020].

2. Kessler RC, Amminger GP, Aguilar-Gaxiola S, Alonso J, Lee S, Üstün TB. Age of onset of mental disorders: A review of recent literature. Curr Opin Psychiatry. 2007;20(4):359–64.

3. Albano AM, Chorpita BF, Barlow DH. Childhood anxiety disorders. In: Mash EJ, Barkley RA, editors. Child psychopathology. 2nd ed. New York, NY, US: The Guilford Press; 2003. p. 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(1):72–9.

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(1):90–9.

6. Boswell JF, Kraus DR, Miller SD, Lambert MJ. Implementing routine outcome monitoring in clinical practice: Benefits, challenges, and solutions. Psychother Res. 2015;25(1):6–19.

7. Scott K, Lewis CC. Using measurement-based care to enhance any treatment. Cogn Behav Pract. 2015 22(1):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(10):1114–9.

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 [published online ahead of print, 2019 Dec 24]. J Am Acad Child Adolesc Psychiatry. 2019;S0890-8567(19)32234-8.

11. Clarke M, Williamson PR. Core outcome sets and systematic reviews. Syst Rev. 2016;5(1):11.

12. Obbarius A, van Maasakkers L, Baer L, Clark DM, Crocker AG, de Beurs E, 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(12):3211–25.

13. Porter ME. What Is Value in Health Care? N Engl J Med. 2010;363(26):2477–81.

14. Kazdin AE. The Meanings and Measurement of Clinical Significance. J Consult Clin Psychol. 1999;67(3):332–9.

15. Spitzer RL, Kroenke K, Williams JBW, Löwe B. A brief measure for assessing generalized anxiety disorder: The GAD-7. Arch Intern Med. 2006 May 22;166(10):1092–7.

16. Spitzer RL, Kroenke K, Williams JBW. Validation and utility of a self-report version of PRIME-MD: The PHQ Primary Care Study. J Am Med Assoc. 1999;282(18):1737–44.

17. Üstün T, Chatterji S, Kostanjsek N, Rehm J, Kennedy C, Epping-Jordan J, et al. Developing the World Health Organization disability assessment schedule 2.0. Bull World Health Organ. 2010;88(11):815–23.

18. Regier DA, Kuhl EA, Kupfer DJ. The DSM-5: Classification and criteria changes. World Psychiatry. 2013;12(2):92–8.

19. De Los Reyes A, Augenstein TM, Wang M, Thomas SA, Drabick DAG, Burgers DE, et al. The validity of the multi-informant approach to assessing child and adolescent mental health. Psychol Bull. 2015;141(4):858–900.

20. Weisz JR, Kuppens S, Ng MY, Eckshtain D, Ugueto AM, Vaughn-Coaxum R, 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. Vol. 72, American Psychologist. American Psychological Association; 2017. p. 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 Allied Discip. 2020; 61(6): 628-643.

22. Linstone HA, Turoff M. The Delphi Method: Techniques and Applications. Reading, MA: Adison Wesley; 1975.

23. Prinsen CAC, Vohra S, Rose MR, Boers M, Tugwell P, Clarke M, et al. How to select outcome measurement instruments for outcomes included in a "Core Outcome Set" - a practical guideline. Trials. 2016;17(1):1–10.

24. Williamson PR, Altman DG, Bagley H, Barnes KL, Blazeby JM, Brookes ST, et al. The COMET Handbook: Version 1.0. Trials. 2017;18(Suppl 3):1–50.

25. Reeve BB, Wyrwich KW, Wu AW, Velikova G, Terwee CB, Snyder CF, 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(8):1889–905.

26. Ebesutani C, Reise SP, Chorpita BF, Ale C, Regan J, Young J, 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(4):833–45.

27. 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(8):835–55.

28. 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.

29. Foa EB, Coles M, Huppert JD, Pasupuleti R V., Franklin ME, March J. Development and Validation of a Child Version of the Obsessive Compulsive Inventory. Behav Ther. 2010;41(1):121–32.

30. Yule W. Anxiety, Depression and Post-Traumatic Stress in Childhood. In: Sclare I, editor. Child Psychology Portfolio. Windsor, England: NFER-Nelson; 1997.

31. Posner K, Brown GK, Stanley B, Brent DA, Yershova K V., Oquendo MA, 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(12):1266–77.

32. Hoagwood KE, 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(8):1055–63.

33. Rapee RM, Bögels SM, Van Der Sluis CM, Craske MG, Ollendick T. Annual research review: Conceptualising functional impairment in children and adolescents. J Child Psychol Psychiatry Allied Discip. 2012;53(5):454–68.

34. Ravens-Sieberer U, Erhart M, Rajmil L, Herdman M, Auquier P, Bruil J, 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(10):1487–500.

35. Shaffer D, Gould MS, Brasic J, Fisher P, Aluwahlia S, Bird H. A Children's Global Assessment Scale (CGAS). Arch Gen Psychiatry. 1983;40(11):1228–31.

36. Lyneham HJ, Sburlati ES, Abbott MJ, Rapee RM, Hudson JL, Tolin DF, et al. Psychometric properties of the Child Anxiety Life Interference Scale (CALIS). J Anxiety Disord. 2013;27(7):711–9.

37. UNESCO Institute of Statistics. International Standard Classification of Education: ISCED 2011 [Internet].

Montreal, PQ, Canada; 2012. Available from: http://uis.unesco.org/sites/default/files/documents/international-standard-classification-of-education-isced-2011-en.pdf [18/06/2020].

38. 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(3):422–36.

39. 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(4):713–31.

40. 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(5):683–94.

41. Emslie GJ, Kennard BD, Mayes TL. Predictors of treatment response in adolescent depression. Psychiatr Ann. 2011;41(4):212–9.

42. 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(8):524–35.

43. Brent DA, Kolko DJ, Birmaher B, Baugher M, Bridge J, Roth C, 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(9):906–14.

44. 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(1):47–56.

45. Storch EA, Merlo LJ, Larson MJ, Geffken GR, Lehmkuhl HD, Jacob ML, et al. Impact of comorbidity on cognitive-behavioral therapy response in pediatric obsessive-compulsive disorder. J Am Acad Child Adolesc Psychiatry. 2008;47(5):583–92.

46. Skriner LC, Chu BC, Kaplan M, Bodden DHM, Bögels SM, Kendall PC, et al. Trajectories and predictors of response in youth anxiety CBT: Integrative data analysis. J Consult Clin Psychol. 2019;87(2):198–211.

47. Hudson JL, Keers R, Roberts S, Coleman JRI, Breen G, Arendt K, 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(6):454–63.

48. Jones M, Hopkins K, Kyrke-Smith R, Davies R, Vostanis P, Wolpert M. Current view tool: Completion guide. London, United Kingdom: CAMHS Press; 2013.

49. World Health Organization. International Classification of Diseases for Mortality and Morbidity Statistics (11th Revision) [Internet]. 2018 [cited 2019 Oct 15]. Available from: https://icd.who.int/browse11/l-m/en [18/06/2020].

50. Bodden DHM, Bögels SM, Nauta MH, De Haan E, Ringrose J, Appelboom C, 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(12):1384–94.

51. 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(3):245–262

52. Legerstee JS, Huizink AC, van Gastel W, Liber JM, Treffers PDA, Verhulst FC, et al. Maternal anxiety predicts favourable treatment outcomes in anxiety-disordered adolescents. Acta Psychiatr Scand. 2008;117(4):289–98.

53. 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(2):141–51.

54. Batty MJ, Moldavsky M, Foroushani PS, Pass S, Marriott M, Sayal K, et al. Implementing routine outcome measures in child and adolescent mental health services: From present to future practice. Child Adolesc Ment Health. 2013;18(2):82–7.

55. 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 Jan;1–8.

56. Davis K, Gorst SL, Harman N, Smith V, Gargon E, Altman DG, 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(2):1–14.

57. Kieling C, Rohde LA. Child and Adolescent Mental Health Research Across the Globe. J Am Acad Child Adolesc Psychiatry. 2012;51(9):945–7.

58. Trauer T. Outcome Measurement in Mental Health: Theory and Practice. Cambridge, England: Cambridge University Press; 2010.

59. Delespaul PAEG. Routine outcome measurement in the Netherlands-A focus on benchmarking. Int Rev

Psychiatry. 2015;27(4):320–8.

60. Department of Health and Ageing. Mental Health National Outcomes and Casemix Collection: Overview of Clinician-Rated and Consumer Self-Report Measures V1.50 [Internet]. Canberra; 2003. Available from: https://www.amhocn.org/sites/default/files/publication_files/mhnocc_overview_doha.pdf [18/06/2020].

61. National Quality Forum. Measure Applications Partnership, 2019 [Internet]. 2019. Available from: https://www.qualityforum.org/Setting_Priorities/Partnership/Measure_Applications_Partnership.aspx [22/06/2020].

62. Monga S, Monsour A, Stallwood E, Desai R, Cleverley K, Courtney D, et al. Core Outcome Set Development for Adolescent Major Depressive Disorder Clinical Trials: A Registered Report. J Am Acad Child Adolesc Psychiatry. Forthcoming 2020.

63. UNICEF. Measurement of Mental Health Among Adolescents at the Population Level (MMAP): An Overview [Internet]. 2019. Available from: https://data.unicef.org/wp-content/uploads/2018/02/Mental-health-measurement_MMAP_overview_UNICEF.pdf [06/12/2019].

64. Myers K, Winters NC. Ten-Year Review of Rating Scales. II: Scales for Internalizing Disorders. J Am Acad Child Adolesc Psychiatry. 2002;41(6):634–59.

65. 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(4):309–38.

66. 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):1–17.

67. Beidas RS, Stewart RE, Walsh L, Lucas S, Downey MM, Jackson K, et al. Free, brief, and validated: Standardized instruments for low-resource mental health settings. Cogn Behav Pract. 2015;22(1):5–19.

68. Mash EJ, Hunsley J. Evidence-based assessment of child and adolescent disorders: Issues and challenges. J Clin Child Adolesc Psychol. 2005;34(3):362–79.

69. Reeve BB, Thissen D, DeWalt DA, Huang I-C, Liu Y, Magnus B, et al. Linkage between the PROMIS® pediatric and adult emotional distress measures. Qual Life Res. 2016;25(4):823–33.

70. Sales CMD. Seeing the person in the patient: Making the case for individualized PROMs in mental health care. Curr Psychiatry Rev. 2017;13(3):184–7.

71. 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(1):1–12.

72. L. Milfont T, Fischer R. Testing measurement invariance across groups: applications in cross-cultural research. Int J Psychol Res. 2010;3(1):111–30.

73. 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(3):829–40.

74. 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.

75. 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(2):174–9.

76. 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(8):1–12.

77. Jones AM, De Nadai AS, Arnold EB, McGuire JF, Lewin AB, Murphy TK, 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(1):137–51.

78. 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(4):262–9.

79. Verlinden E, van Meijel EPM, Opmeer BC, Beer R, de Roos C, Bicanic IAE, et al. Characteristics of the children's revised impact of event scale in a clinically referred dutch sample. J Trauma Stress. 2014;27(3):338–44.

80. Magalhães S de S, 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. 2018;45(2):27–32.

81. 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(6):605–11.

82. 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. 2019;12(1):23–9.

83. 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(2):115–32.

84. 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(2):407–20.

85. 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(7):1729–37.

86. Weissman MM, Warner V, Fendrich M. Applying Impairment Criteria to Children's Psychiatric Diagnosis. J Am Acad Child Adolesc Psychiatry. 1990;29(5):789–95.

87. Bird HR. Further Measures of the Psychometric Properties of the Children's Global Assessment Scale. Arch Gen Psychiatry. 1987;44(9):821.

88. Dyrborg J, Warborg Larsen F, Nielsen S, Byman J, Buhl Nielsen B, 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(3):195–201.

89. Sourander A, Helenius H, Piha J. Child psychiatric short-term inpatient treatment: CGAS as follow-up measure. Child Psychiatry Hum Dev. 1996;27(2):93–104.

90. 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(5):756–63.

91. 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 e Crit. 2015;28(3):454–62.