BMJ Open

# Assessing the impact of mental health difficulties on young people's daily lives: protocol for a scoping umbrella review of measurement instruments

Karolin Rose Krause [1,2] Sophie Chung,[3] Terri Rodak,[4] Kristin Cleverley [5,6] Nancy J Butcher [7,8] Peter Szatmari[1,8]

**To cite:** Krause KR, Chung S, Rodak T, *et al*. Assessing the impact of mental health difficulties on young people's daily lives: protocol for a scoping umbrella review of measurement instruments. *BMJ Open* 2022;**12**:e054679. doi:10.1136/bmjopen-2021-054679

► Prepublication history and additional supplemental material for this paper are available online. To view these files, please visit the journal online (http://dx.doi.org/10.1136/bmjopen-2021-054679).

Received 20 June 2021
Accepted 10 April 2022

Check for updates

© Author(s) (or their employer(s)) 2022. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

For numbered affiliations see end of article.

**Correspondence to**
Dr Peter Szatmari;
peter.szatmari@utoronto.ca

## ABSTRACT

**Introduction** An important consideration for determining the severity of mental health symptoms is their impact on youth's daily lives. Those wishing to assess 'life impact' face several challenges: First, various measurement instruments are available, including of global functioning, health-related quality of life and well-being. Existing reviews have tended to focus on one of these domains; consequently, a comprehensive overview is lacking. Second, the extent to which such instruments truly capture distinct concepts is unclear. Third, many available scales conflate symptoms and their impact, thus undermining much needed analyses of associations between the two.

**Methods and analysis** A scoping umbrella review will examine existing reviews of life impact measures for use with children and youth aged 6–24 years in the context of mental health and well-being research. We will systematically search six bibliographic databases (MEDLINE, Embase, APA PsycINFO, CINAHL, Web of Science, and the COSMIN database of systematic reviews of outcome measurement instruments), and conduct systematic record screening, data extraction and charting based on methodological guidance by the Joanna Briggs Institute. Data synthesis will involve the tabulation of scale characteristics, feasibility and measurement properties, and the use of summary statistics to synthesise how these instruments operationalise life impact. The protocol was registered prospectively with the Open Science Framework (osf.io/ers48).

**Ethics and dissemination** This study will provide a comprehensive road map for researchers and clinicians seeking to assess life impact in youth mental health, providing guidance in navigating available measurement options. We will seek to publish the findings in a leading peer-reviewed journal in the field. Formal research ethics approval will not be required.

## Strengths and limitations of this study

- ► Umbrella review methodology will enable a higher-level synthesis of existing review efforts, thus generating a comprehensive map of available measurement instruments, and their properties.
- ► Our methodological approach is based on the Joanna Briggs Institute guidelines for scoping reviews, umbrella reviews and psychometric reviews.
- ► This review is based on a rigorous systematic search developed and executed by a health science librarian.
- ► We will only include studies published in the English language since 1990.
- ► As this is a scoping umbrella review designed to map available measurement instruments, the quality and risk of bias of included review articles will not be systematically assessed.

## INTRODUCTION

A key consideration for determining the severity of mental health difficulties is the extent to which these difficulties impact on a young person's daily life. The Diagnostic and Statistical Manual of Mental Disorders-5[1] determines 'clinical significance' in relation to two criteria: individuals must display specific symptoms, and those symptoms must cause considerable distress or impairment in daily life.[2] Impaired daily functioning has been shown to influence help-seeking and health providers' decisions about the type of care an individual should receive.[3 4] Assessing life impact can also help contextualise changes in symptom severity scores when assessing treatment efficacy and effectiveness.[5–10] From a public health perspective, consideration of life impact has moved common mental health conditions like depression to the fore of public health agendas, by showcasing that the associated burden of disease is comparable to that of cardiovascular or respiratory diseases.[2]

In child and youth mental health (hereafter we will use the terms 'youth mental health' and 'young people'/'youth' for brevity to refer to those aged 6–24 years, in line with definitions of middle childhood, adolescence and young adulthood by the National Institute of Child Health and Human Development

(Paediatric) Terminology[11] and the United Nations[12]), life impact has typically been approached through the lens of functional impairment.[13–15] Functioning describes a young person's ability 'to adapt to varying demands of home, school, peer group or neighbourhood' in line with age-specific expectations and cultural norms (p. 1060).[13] On a continuum of functioning, impairment marks one end of the spectrum, while high levels of adaptation and competency (eg, thriving, flourishing) mark the other end.

The Children's Global Assessment Scale (CGAS)[16] is a commonly used single-item measure that provides an overall rating of a young person's functioning, based on clinician report. Other instruments take a more fine-grained approach by assessing functioning in specific areas of life. For example, the Social Adjustment Scale[17] generates separate subscale scores for social functioning with friends, family, at school and in dating contexts. In addition, measures of symptom-specific or condition-specific impairment, focus on the extent to which psychopathological 'symptoms interfere with and reduce adequate performance of important and desired aspects of a child's life' (p. 455).[18] For example, the Strengths and Difficulties Questionnaire (Impact Supplement),[19 20] and diagnostic interviews like the Kiddie Schedule for Affective Disorders and Schizophrenia[21] enquire about functional impairment caused by psychopathology symptoms indicated during earlier parts of the respective assessments.

In physical health contexts and some population-based research, the impact of a particular health condition or of a person's overall health status on their daily life is often conceptualised as health-related quality of life (HRQoL). Quality of life, has been described as 'the overall positivity with which individuals view their state and circumstances' (p. 455),[18] and is thought to span physical, mental and social well-being.[18 22] HRQoL refers more specifically to quality of life in a health or medical context.[23] Relevant instruments include, for example, the brief EuroQol 5D-youth that is commonly used in economic evaluations[24]; the 52-item KIDSCREEN that was developed for the measurement of HRQoL in the general paediatric population,[25] or the PROMIS item bank for paediatric global health, designed to assess overall perceptions of health in youth with chronic health conditions.[26]

Well-being is another domain that researchers may consider when assessing the life impact of mental health conditions. While a consensus definition is lacking, this domain has been described as 'a combination of positive emotions, engagement, meaningful relationships and a sense of accomplishment, or as flourishing in aspects of feeling and functioning, thus reflecting the positive aspects of mental health' (p. 771).[27] For example, the Warwick-Edinburgh Mental Well-Being Scale[28] is a self-report instrument validated in adolescents that exclusively assesses positive aspects of mental health.

The conceptual domains of functioning, HRQoL and well-being have different theoretical roots, yet it has been suggested that these terms are often used interchangeably.[18] All three might be considered as avenues for assessing the life impact of mental health difficulties in children and youth. For example, a recently developed core outcome set for child and youth anxiety and depression recommends assessing functioning via three measures: the CGAS as a measure of clinician-rated global functioning; a self-report scale of condition-specific impairment; and the KIDSCREEN as a HRQoL measure.[9] More generally, it is not clear whether scales purported to assess life impact via these domains are truly conceptually distinct, or whether they merely focus on different ends of the functioning continuum[29]. A systematic review of measurement instruments that examines degrees of overlap and complementarity is lacking.

Researchers wishing to assess the life impact of mental health difficulties in young people further face the challenge of selecting the most appropriate instrument. A recent scoping review identified 14 different measures of global functioning, three measures of condition-specific impairment, and 14 measures of HRQoL, across 257 treatment studies for child and youth anxiety, depression, obsessive–compulsive disorder and post-traumatic stress disorder.[9] Several reviews provide overviews of available instruments,[30–36] their measurement properties and feasibility characteristics, but these have tended to be domain-specific (eg, focusing only on HRQoL); consequently, a comprehensive overview of life impact measures is lacking. On the other hand, broader reviews of mental health assessment tools [37–39] have not typically been exhaustive in their coverage of life impact measures, and have not tended to examine methodological questions specific to life impact assessment.

A third challenge to the measurement of life impact in mental health is that many available instruments conflate items that assess symptom severity with items assessing the life impact of such symptoms. For example, the CGAS's description of 'superior functioning' includes 'no symptoms' as a criterion.[16] Similarly, the Health of the Nation Outcome Rating Scale[40] is a 13-item measure that includes 7 symptom-focused items alongside five functional items (covering school functioning, self-care and relationships with peers and at home). The conflation of symptom severity and life impact items in a single scale hinders analyses of cross-sectional and longitudinal associations between the two domains.[18]

Finally, an important fourth challenge is that many available instruments have been developed in Western high-income countries, and may not have cross-cultural validity or measurement invariance in lower-income or middle-income contexts or in specific cultural communities[41]. As functioning is defined in relation to age-specific and culture-specific expectations and norms,[13] life impact measures that are not culturally sensitive and appropriate may yield misleading data. Even in the contexts where measurement instruments were originally developed, youth may not always have been involved in their creation, which may weaken their content validity.[42]

### Objectives and research questions

This scoping umbrella review will examine how functioning, impairment, HRQoL and well-being have been conceptualised and measured in the youth mental health literature. It will seek to provide an overview of the design characteristics, feasibility aspects and measurement properties of available instruments, by considering existing individual reviews as primary studies. We seek to answer the following research questions:

RQ 1. What child-reported, parent-reported and clinician-reported measurement scales are available for assessing life impact in children and youth aged 6–24 years in the context of mental health and well-being research?

RQ 2. What information is available from existing reviews about the design characteristics (eg, target construct, target age range and use context, intended informant), feasibility (ie, length, cost and accessibility, language version availability) and measurement properties (ie, validity, reliability, responsiveness) of these instruments?

RQ 3. What populations and use contexts were these instruments originally designed for, according to their initial development study? Which cultural contexts were the instruments validated in?

RQ 4. According to an instrument's original development study, were young people consulted as part of the measure development process?

RQ 5. Do measures of functioning, HRQoL and well-being appear to capture distinct conceptual domains, as opposed to assessing the same domain at different ends of the functioning continuum, based on subscale and item content?

RQ 6. To what extent do available measures of life impact conflate the measurement of psychopathology symptoms with the measurement of life impact?

## METHODS AND ANALYSIS

### Study design

The proposed study is a scoping umbrella review. An umbrella review considers existing review articles as its principal source of evidence and aims to compare, contrast or synthesise their findings.[43] While systematic reviews (including umbrella reviews) typically seek to answer clearly defined questions (eg, 'Which measure of global functioning provides the highest degree of validity and reliability'), scoping reviews often seek to answer broader questions about the state of the evidence or about predominant methodological approaches in a given area.[44] A scoping umbrella review is an appropriate approach for this study because several existing reviews can be synthesised to provide a comprehensive mapping of available instruments and their properties. We will follow the methodological guidance provided by the Joanna Briggs Institute (JBI) for the conduct of umbrella reviews, scoping reviews and reviews of measurement properties.[45–47]

### Protocol

This review protocol complies with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) for Protocols reporting guidelines (online supplemental appendix 1).[48] The final review will follow the PRISMA for Scoping Reviews.[44] The protocol was registered prospectively with the Open Science Framework (OSF) on 26 May 2021.[49] On registration and submission of the protocol, title and abstract screening was complete, but full text screening had not begun. Any important amendments to this protocol will be documented on the OSF registration page.

### Inclusion criteria

This scoping umbrella review will consider systematic reviews, scoping reviews, rapid reviews and narrative reviews that seek to provide an overview of available measurement scales to assess functioning, impairment, HRQoL or well-being. These may be reviews that systematically identify a range of measurement instruments, or reviews that synthesise the available literature for a single instrument. Narrative reviews, rapid reviews and scoping reviews will be included in addition to systematic reviews because this scoping umbrella review aims to map the landscape of available life impact measures as comprehensively as possible, rather than to identify the most systematic or robust evidence pertaining to these instruments. Inclusion criteria for reviews are defined to match the PICO components for systematic reviews of measurement properties (ie, Population, Instruments, Construct, Outcome) in line with JBI guidelines.[47] The PICO criteria are summarised in table 1.

#### Population (P)

Review articles must have an explicit focus on measurement in middle childhood (defined here as starting at age 6 in line with proposed age group standards[11]), adolescence, and/or young adulthood (defined here as ending at age 24, in line with the United Nations' definition of 'youth'[12]). Studies with a majority focus on adults will not be considered, unless they include a separate appraisal of tools for a relevant paediatric age group. We will exclude reviews focused on early childhood (ie, ages 0–5 years), where tailored assessment approaches are likely needed.

We will include reviews that examine the measurement of life impact in populations with a primary mental health or substance use concern, or in the context of assessing mental health and well-being in the general population or in non-specific health contexts. We will exclude reviews that focus on youth with physical health conditions. Instruments identified in such reviews may place a considerable focus on physical body functions that may be less relevant in a mental health context. We will also exclude reviews focused on youth with intellectual disabilities, neurological conditions (eg, epilepsy, cerebral palsy) or autism spectrum disorder (ASD). These profiles may require specialised assessments of life impact, and the conceptual separation of symptoms from functioning may

Table 1 PICO statement for scoping umbrella review

| | P (Population) | I (Instruments) | C (Constructs) | O (Outcomes) |
|---|---|---|---|---|
| Included | Children and adolescents aged 6–24 years, with a primary mental health condition/concern, subject to mental health assessment in general population, or in the context of assessing life impact in health contexts broadly speaking. | Youth, parent, clinician or external rater report; initial assessment or outcome measure | Global functioning<br>Social functioning<br>Functional impairment<br>HRQoL<br>Well-being<br>Flourishing | Construct domain<br>Target age group<br>Reporter<br>Target use context<br>Length<br>Accessibility and cost<br>Measurement properties |
| Excluded | Ages 0–5 or 24+<br>Children and youth with intellectual disabilities or where mental health is a secondary concern to a primary physical condition; pure physical health context | Performance test; biometric assessment | Language ability<br>Cognitive ability<br>Executive functioning<br>Symptom severity | |

HRQoL, health-related quality of life.

be particularly complicated (eg, with social functioning constituting a symptom of ASD). A separate review may be warranted to cover life impact measures for children and youth with these conditions.

### Instruments (I)

We will consider scales deigned for completion by clinicians, external raters, parents or carers, and young people. These may be assessment or outcome measures but must focus on an eligible domain of life impact (ie, see below) rather than symptom severity or psychopathology. We will exclude reviews focused on diagnostic tools or on the assessment of specific mental health conditions, unless the article's abstract explicitly states that measures of an eligible life impact construct were considered alongside symptom severity measures. We will further exclude performance tests, cognitive tests, language assessments, biometric tests, school-based functional behavioural assessments[50] and population-level composite indices of well-being or HRQoL.

### Constructs (C)

We will consider instruments designed to assess life impact through the measurement of global functioning, social functioning/adaptation, general or condition-specific impairment, HRQoL, well-being (including flourishing), and life satisfaction. Constructs that are not eligible include symptoms of psychopathology, language ability, cognitive ability, executive functioning and motor functioning. Instruments that cover any of these constructs at an item level as part of measuring a broader eligible construct (eg, HRQoL) may be included.

### Outcomes (O)

We will include articles that state an intent to review, appraise or map relevant measurement instruments and that provide a structured discussion or a tabulated overview of the instruments identified.

### Publication type

Reviews must have been published from January 1990 onwards. We will limit the language of publication to English to accommodate languages spoken within the review team; to ensure that all records can be screened by at least two review team members; and because item content will be an important aspect to consider and must therefore be accessible to the majority of the review team.

We will include review articles published in peer-reviewed journals, assessment handbooks (if accessible online) and conference proceedings (including workshop summaries and conference papers, but not including conference abstracts). We will further include reviews that were published as grey literature (eg, as reports on organisational websites) and otherwise meet the inclusion criteria.

## Search strategy

The development of the search strategy is led by a health science librarian (TR) in collaboration with other members of the review team (KKR, PS). The search strategy combines search terms describing the population (eg, "child*" OR "youth" AND "depress* OR "anxiety disorder*" OR "externalizing problem*") and domains of interest (eg, "function* OR "HRQOL") with search terms limiting the results to reviews (eg, "systematic review" OR "scoping review") of measurement instruments or outcome measurement approaches (eg, "psychometr*" OR "measurement instrument*"). Our tailored search syntax is informed by existing search filters for measurement instruments that were developed by the University of Oxford's Patient-Reported Outcome Measurement Group[51] and by the COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN) initiative.[52] Pilot searches informed the final search strategy (see online supplemental appendix 2).

The final search will be performed by the review team's health science librarian (TR) in Medline, Embase, APA PsycINFO, Cumulative Index to Nursing & Allied

Health Literature (CINAHL) and Web of Science, and by a member of the review team (KRK) in the COSMIN database of systematic reviews of outcome measurement instruments. Retrieved records will be deduplicated using Covidence systematic review software.[53]

We will ask a group of subject matter experts to review the list of articles identified through the automatic search, and to suggest additional reviews that may have been missed. We will also conduct a targeted grey literature search via specific databases and websites identified as relevant by the team's health science librarian (TR). In addition, we will handsearch the reference lists of included reviews to identify and retrieve the original development papers associated with eligible instruments, as well as copies of the instruments themselves, as available. We will consider supplemental searches if key information about a measure's design characteristics is not available from the identified reviews or the instruments' original development studies. Due to resource constraints, we will not conduct supplemental searches for a measure's feasibility characteristics or measurement properties, and will base our reporting for these aspects on the information available from existing reviews.

We will review clearinghouses of measurement instruments[37] for any additional scales that were missed by the included reviews, and will also make a note of any additional instruments identified while screening for eligibility. These additional instruments will not be subject to a systematic appraisal, but will be listed in the final report.

## Study selection

Eligibility will be assessed via a two-stage screening process. For the title and abstract screening, 20% of all identified records will be screened independently and in duplicate by two reviewers (KRK and SC). A kappa coefficient exceeding 0.7 will indicate substantial inter-rater agreement.[54] We will discuss any discrepancies and agree a final inclusion or exclusion rating. A single reviewer (KRK) will then screen the remaining titles and abstracts. All records retained for full-text screening will be checked for eligibility independently and in duplicate by two raters. Disagreements will be discussed and decisions about inclusion will be made with the help of a third reviewer as needed. Articles that do not meet inclusion criteria will be coded for exclusion in the Covidence software environment with the first exclusion criterion that becomes apparent. Eligible review articles will progress to data charting.

## Data extraction and charting

Data will be extracted and charted using tailored adaptations of the JBI data extraction templates for systematic reviews of measurement properties[47] and for umbrella reviews.[46] The adapted matrices will be piloted to ensure an appropriate level of detail is charted. Data extraction will be conducted by one review author, and spot checks for comprehensiveness and accuracy will be conducted on at least 20% of the included reviews by a second reviewer.

**Table 2** Overview of data to extract and chart

| Information category | Detailed information to extract |
|---|---|
| Publication identifiers | Journal, year, first author |
| Review characteristics | Type of publication, type of review, objective of the review, population and setting considered, number and names of databases searched, date range of search, language/geographical restrictions, number of studies included |
| Instrument design characteristics | Instrument name, domain measured, number of items, number and names of subscales, target age group, target population group (clinical vs non-clinical), target use context (screening, diagnosis, outcome measurement), reporter(s), response scale, recall period, involvement of youth in instrument development, cultural context of development and validation studies |
| Feasibility characteristics | Length, cost and accessibility, available language versions |
| Measurement properties | Summary findings relating to validity, reliability, responsiveness |

Any disagreements will be resolved through discussion. Based on the extent of disagreement identified, the two reviewers will consider extending the spot checks to a larger subset of studies.

The information to be charted is shown in table 2. We will refer to the original development studies as needed, to extract whether or not youth or families were involved in measure development. We will also extract in which contexts an instrument has been validated. Where possible, we may review each measure's item content to indicate whether items cover symptoms of psychopathology as well as life impact, and to examine the extent of overlap between measures purported to assess different life impact domains. Depending on the number and accessibility of the instruments identified, we may seek to undertake a systematic item-level mapping of content.[55 56]

## Risk of bias assessment

Scoping-type reviews do not seek to generate critically appraised and summative responses to specific research questions, but instead aim to map the available evidence on a given topic. Therefore, risk of bias assessments are not typically conducted as part of scoping reviews,[57] and are not planned for this scoping umbrella review.

## Strategy for data synthesis

We will synthesise the findings of existing reviews in relation to instrument design characteristics, feasibility characteristics, and measurement properties by applying the five-step data synthesis process recommended by Miles and Huberman[58] and Whittemore and Knafl.[59] This process consists of (1) data reduction; (2) data display; (3) data comparison; (4) conclusion drawing and (5) verification. During verification, we will review the original development studies associated with each instrument to ensure that the information about key instrument characteristics compiled during the scoping umbrella review is accurate.

We will present the characteristics of the included review articles, as well as the characteristics of the

identified measurement instruments in tabular format. We will report high-level quantitative summary statistics (ie, counts or frequencies) to describe the reviews and instruments identified (eg, number of instruments overall; number of instruments per life impact domain; number of instruments by type of reporter). We will further generate summary statistics and visualisations to report on the domains and subdomains of life impact covered by the identified instruments, and the extent to which these instruments appear to conflate items measuring symptoms of psychopathology with items measuring life impact, based on an examination of item or subscale content. We will also specifically indicate whether a measure was validated in a population with a mental health concern, or whether it was validated exclusively in community samples.

### Patient and public involvement

The Centre for Addiction and Mental Health implements a Youth Engagement Initiative that brings the voices of youth aged 14–29 years with lived experience of mental health challenges into research and service design.[60–62] We will collaborate with a designated youth research partner in conducting specific aspects of this review that require the qualitative interpretation of measure content. In addition, we will present draft review findings to a virtual focus group including between four and eight youth advisors to solicit their feedback, and incorporate this into our interpretation and contextualisation of the review findings prior to finalising the study manuscript.

### Ethics and dissemination

#### Ethics

Formal approval by a research ethics board will not be required, as the proposed project is a scoping umbrella review of existing data. We will consult youth as research partners rather than research subjects and will not collect or report any individual-level participant data.

#### Dissemination

We will seek to publish the findings from this scoping umbrella review in a leading peer-reviewed journal in the field of child and adolescent mental health. We will also seek to disseminate findings at national and international conferences, and will consider submitting the final review to the COSMIN database of systematic reviews of measurement properties. We will equally consider additional channels of dissemination, such as blog posts or podcasts.

## DISCUSSION

Historically, outcome measurement in youth mental health research has focused on symptom severity.[63–65] Yet, many common symptom scales are not immediately interpretable with regard to how a score change translates into real-world changes in a young person's life. Assessing life impact in a structured way through use of suitable measurement scales can provide important complementary information to data collected via diagnostic tools and symptom severity measures.[5–7 66 67] Within an empathetic, person-centred framework of care, it is important that clinicians pay attention to youth and family members' unstructured, narrative accounts of how a mental health condition affects daily life, and that clinicians consider these narratives when making care decisions together with service users.[68] In addition, however, the administration of suitable structured measurement instruments can help ensure that life impact is assessed systematically and reliably, so that comparisons can be made between individuals and over time. Similarly, the use of structured assessment tools can enable the systematic consideration of different perspectives (eg, when combining youth-rated, carer-rated and clinician-rated instruments),[69] and the inclusion of self-reported datapoints that convey the youth's perspective directly without mediation by the clinician. Two recent initiatives have highlighted functioning as a core outcome to track when evaluating clinical care for paediatric anxiety and depression,[9] and when measuring youth mental health outcomes in population surveys.[70] Yet, difficulties have been reported with identifying a gold-standard measure.[9]

This scoping umbrella review does not aim to comprehensively identify all life impact measures available worldwide, but will focus on instruments that have been reviewed in English language publications. This review also does not seek to yield an authoritative summary of which instruments provide the best measurement properties. This would require an in-depth assessment of the methodological quality of the psychometric evidence underpinning each instrument in line with COSMIN guidelines,[71] which in turn would constitute a study in its own right for each instrument identified.[72] Instead, this review will examine a range of design, feasibility and measurement properties to facilitate the preselection of candidates for future in-depth psychometric appraisals. It further aims to identify gaps with regards to age groups or use contexts covered, and examine the degree of conceptual overlap between instruments designated to assess different outcome domains (eg, functioning vs HRQoL). As such, it aims to take stock of current measurement practice, to inform discussions about suitable ways forward and to provide a road map to researchers and clinicians seeking to appraise which tool or combination of tools may be appropriate for a given population and context.

**Author affiliations**
[1]Cundill Centre for Child and Youth Depression, Centre for Addiction and Mental Health, Toronto, Ontario, Canada
[2]Research Department for Clinical, Educational and Health Psychology, University College London, London, UK
[3]Independent Health Researcher, London, UK
[4]CAMH Library, Department of Education, Centre for Addiction and Mental Health, Toronto, Ontario, Canada
[5]Lawrence Bloomberg Faculty of Nursing and Department of Psychiatry, University of Toronto, Toronto, Ontario, Canada
[6]The Margaret and Wallace McCain Centre for Child, Youth & Family Mental Health, Centre for Addiction and Mental Health, Toronto, Ontario, Canada

[7]Child Health Evaluative Sciences, Hospital for Sick Children Research Institute, Toronto, Ontario, Canada
[8]Department of Psychiatry, University of Toronto, Toronto, Ontario, Canada

**Twitter** Karolin Rose Krause @KrauseKarolin, Kristin Cleverley @ClevKristin and Nancy J Butcher @NancyJButcher

**Contributors** KRK, SC, KC, NJB and PS were responsible for study conception and design. KRK drafted the manuscript. TR designed the search strategy with input from KRK and PS, and led the database search. All authors provided feedback on the study design and manuscript, and approved the final manuscript prior to its submission.

**Funding** This research is supported by the Cundill Centre for Child and Youth Depression. This research did not receive any specific grant from any funding agency in the public, commercial or not-for-profit sectors.

**Competing interests** None declared.

**Patient and public involvement** Patients and/or the public will be involved in the conduct of this research. Refer to the Methods section for further details.

**Patient consent for publication** Not applicable.

**Provenance and peer review** Not commissioned; externally peer reviewed.

**Supplemental material** This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

**Open access** This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/.

**ORCID iDs**
Karolin Rose Krause http://orcid.org/0000-0003-3914-7272
Kristin Cleverley http://orcid.org/0000-0002-2822-2129
Nancy J Butcher http://orcid.org/0000-0002-5152-0108

## REFERENCES

1 American Psychiatric Association. *Diagnostic and statistical manual of mental disorders. 5th ed*. Arlington, VA: American Psychiatric Association, 2013.
2 Ustün B, Kennedy C. What is "functional impairment"? Disentangling disability from clinical significance. *World Psychiatry* 2009;8:82–5.
3 Hodges K, Doucette-Gates A, Kim CS. Predicting service utilization with the child and adolescent functional assessment scale in a sample of youths with serious emotional disturbance served by center for mental health services-funded demonstrations. *J Behav Health Serv Res* 2000;27:47–59.
4 Striley CW, Stiffman AR, Spitznagel EL. Functioning mediates between symptoms and provider assessment. *Ment Health Serv Res* 2003;5:155–71.
5 Hill CE, Chui H, Baumann E. Revisiting and reenvisioning the outcome problem in psychotherapy: an argument to include individualized and qualitative measurement. *Psychotherapy* 2013;50:68–76.
6 Kazdin AE. The meanings and measurement of clinical significance. *J Consult Clin Psychol* 1999;67:332–9.
7 Kazdin AE. Arbitrary metrics: implications for identifying evidence-based treatments. *Am Psychol* 2006;61:42–9.
8 Zimmerman M, McGlinchey JB. Why don't psychiatrists use scales to measure outcome when treating depressed patients? *J Clin Psychiatry* 2008;69:1916–9.
9 Krause KR, Chung S, Adewuya AO, *et al*. International consensus on a standard set of outcome measures for child and youth anxiety, depression, obsessive-compulsive disorder, and post-traumatic stress disorder. *Lancet Psychiatry* 2021;8:76–86.
10 McKnight PE, Kashdan TB. The importance of functional impairment to mental health outcomes: a case for reassessing our goals in depression treatment research. *Clin Psychol Rev* 2009;29:243–59.
11 Williams K, Thomson D, Seto I, *et al*. Standard 6: age groups for pediatric trials. *Pediatrics* 2012;129 Suppl 3:S153–60.
12 United Nations. Youth, 2020
13 Hoagwood K, Jensen PS, Petti T, *et al*. Outcomes of mental health care for children and adolescents: I. A comprehensive conceptual model. *J Am Acad Child Adolesc Psychiatry* 1996;35:1055–63.
14 Bird HR, Shaffer D, Lucas CP. The assessment of functional impairment. In: *Diagnostic assessment in child and adolescent psychopathology*. New York, NY: The Guilford Press, 1999: 209–29.
15 Bird HR, Yager TJ, Staghezza B, *et al*. Impairment in the epidemiological measurement of childhood psychopathology in the community. *J Am Acad Child Adolesc Psychiatry* 1990;29:796–803.
16 Shaffer D, Gould MS, Brasic J, *et al*. A children's global assessment scale (CGAS). *Arch Gen Psychiatry* 1983;40:1228–31.
17 Weissman MM, Bothwell S. Assessment of social adjustment by patient self-report. *Arch Gen Psychiatry* 1976;33:1111–5.
18 Rapee RM, Bőgels SM, van der Sluis CM, *et al*. Annual research review: conceptualising functional impairment in children and adolescents. *J Child Psychol Psychiatry* 2012;53:454–68.
19 Goodman R. The strengths and difficulties questionnaire: a research note. *J Child Psychol Psychiatry* 1997;38:581–6.
20 Goodman R. The extended version of the strengths and difficulties questionnaire as a guide to child psychiatric caseness and consequent burden. *J Child Psychol Psychiatry* 1999;40:791–9.
21 Kaufman J, Birmaher B, Brent D, *et al*. Schedule for affective disorders and schizophrenia for school-age children-present and lifetime version (K-SADS-PL): initial reliability and validity data. *J Am Acad Child Adolesc Psychiatry* 1997;36:980–8.
22 World Health Organization. Preamble to the constitution of the world health organization. *Off Rec World Heal Organ* 1948;2:100.
23 Fayers PM, Machin D. Developing and validating instruments for assessing quality of life and patient-reported outcomes. In: *Quality of life the assessment, analysis, and reporting of patient-reported outcomes*, 2016: 3–33.
24 Wille N, Badia X, Bonsel G, *et al*. Development of the EQ-5D-Y: a child-friendly version of the EQ-5D. *Qual Life Res* 2010;19:875–86.
25 The Kidscreen Group Europe. The KIDSCREEN questionnaires: quality of life questionnaires for children and adolescents. In: *Handbook*, 2006.
26 Forrest CB, Bevans KB, Pratiwadi R, *et al*. Development of the PROMIS ® pediatric global health (PGH-7) measure. *Qual Life Res* 2014;23:1221–31.
27 Patalay P, Fitzsimons E. Correlates of mental illness and wellbeing in children: are they the same? results from the UK millennium cohort study. *J Am Acad Child Adolesc Psychiatry* 2016;55:771–83.
28 Tennant R, Hiller L, Fishwick R, *et al*. The Warwick-Edinburgh mental well-being scale (WEMWBS): development and UK validation. *Health Qual Life Outcomes* 2007;5:1–13.
29 Böhnke JR, Croudace TJ. Calibrating well-being, quality of life and common mental disorder items: psychometric epidemiology in public mental health research. *Br J Psychiatry* 2016;209:162–8.
30 Canino G, Costello EJ, Angold A. Assessing functional impairment and social adaptation for child mental health services research: a review of measures. *Ment Health Serv Res* 1999;1:93–108.
31 Winters NC, Collett BR, Myers KM. Ten-Year review of rating scales, VII: scales assessing functional impairment. *J Am Acad Child Adolesc Psychiatry* 2005;44:309–38.
32 John K. Measuring children's social functioning. *Child psychol. psychiatr. rev.* 2001;6:181–8.
33 Crowe LM, Beauchamp MH, Catroppa C, *et al*. Social function assessment tools for children and adolescents: a systematic review from 1988 to 2010. *Clin Psychol Rev* 2011;31:767–85.
34 Solans M, Pane S, Estrada M-D, *et al*. Health-Related quality of life measurement in children and adolescents: a systematic review of generic and disease-specific instruments. *Value Health* 2008;11:742–64.
35 Rajmil L, Roizen M, Psy AU, *et al*. Health-Related quality of life measurement in children and adolescents in Ibero-American countries, 2000 to 2010. *Value Health* 2012;15:312–22.
36 Bentley N, Hartley S, Bucci S. Systematic review of self-report measures of general mental health and wellbeing in adolescent mental health. *Clin Child Fam Psychol Rev* 2019;22:225–52.
37 Becker-Haimes EM, Tabachnick AR, Last BS, *et al*. Evidence base update for brief, free, and accessible youth mental health measures. *J Clin Child Adolesc Psychol* 2020;49:1–17.
38 Beidas RS, Stewart RE, Walsh L, *et al*. Free, brief, and validated: standardized instruments for low-resource mental health settings. *Cogn Behav Pract* 2015;22:5–19.

39 Kwan B, Rickwood DJ. A systematic review of mental health outcome measures for young people aged 12 to 25 years. *BMC Psychiatry* 2015;15:279.
40 Gowers SG, Bailey SJ. Outcomes in child and adolescent mental health services. *Curr Opin Psychiatry* 1999;12:439–43.
41 Williamson A, Andersen M, Redman S, *et al*. Measuring mental health in Indigenous young people: a review of the literature from 1998-2008. *Clin Child Psychol Psychiatry* 2014;19:260–72.
42 Terwee CB, Prinsen CAC, Chiarotto A, *et al*. COSMIN methodology for evaluating the content validity of patient-reported outcome measures: a Delphi study. *Qual Life Res* 2018;27:1159–70.
43 Aromataris E, Fernandez R, Godfrey CM, *et al*. Summarizing systematic reviews: methodological development, conduct and reporting of an umbrella review approach. *Int J Evid Based Healthc* 2015;13:132–40.
44 Tricco AC, Lillie E, Zarin W, *et al*. PRISMA extension for scoping reviews (PRISMA-ScR): checklist and explanation. *Ann Intern Med* 2018;169:467–73.
45 Peters M, Godfrey C, McInerney P. Chapter 11: Scoping reviews. In: *JBI manual for evidence synthesis*. JBI, 2020.
46 Aromataris E, Fernandez R, Godfrey C. Chapter 10: Umbrella reviews. In: Aromataris E, Munn Z, eds. *JBI manual for evidence synthesis*. JBI, 2020.
47 Stephenson M, Riitano D, Wilson S. Chapter 12: Systematic reviews of measurement properties. In: Aromataris E, Munn Z, eds. *JBI manual for evidence synthesis*. JBI, 2020.
48 Shamseer L, Moher D, Clarke M, *et al*. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: elaboration and explanation. *BMJ* 2015;349:g7647–25.
49 Butcher NJ, Cleverley K, Rodak T, *et al*. Assessing the impact of mental health difficulties on young people's daily lives: protocol for a scoping umbrella review of measurement instruments. protocol registration. open science framework, 2021. Available: https://osf.io/ers48 [Accessed 26 May 2021].
50 O'Neill RE, Albin RW, Storey K. Functional assessment and program development for problem behavior: a practical handbook. In: *Cengage learning*, 2014.
51 Mackintosh A, Comabella C, Hadi M. PROM Group construct & instrument type filters, 2010. Oxford, England. Available: https://cosmin.nl/wp-content/uploads/prom-search-filter-oxford-2010.pdf
52 Terwee CB, Jansma EP, Riphagen II, *et al*. Development of a methodological PubMed search filter for finding studies on measurement properties of measurement instruments. *Qual Life Res* 2009;18:1115–23.
53 Covidence systematic review software, 2021. Available: www.covidence.org
54 Cohen J. A coefficient of agreement for nominal scales. *Educ Psychol Meas* 1960;20:37–46.
55 Fried EI. The 52 symptoms of major depression: lack of content overlap among seven common depression scales. *J Affect Disord* 2017;208:191–7.
56 Newson JJ, Hunter D, Thiagarajan TC. The heterogeneity of mental health assessment. *Front Psychiatry* 2020;11:1–24.
57 Munn Z, Peters MDJ, Stern C, *et al*. Systematic review or scoping review? guidance for authors when choosing between a systematic or scoping review approach. *BMC Med Res Methodol* 2018;18:1–7.
58 Miles MB, Huberman AM. *Qualitative data analysis: an expanded sourcebook*. sage, 1994.
59 Whittemore R, Knafl K. The integrative review: updated methodology. *J Adv Nurs* 2005;52:546–53.
60 Heffernan OS, Herzog TM, Schiralli JE, *et al*. Implementation of a youth-adult partnership model in youth mental health systems research: challenges and successes. *Health Expect* 2017;20:1183–8.
61 Hawke LD, Darnay K, Relihan J, *et al*. Enhancing researcher capacity to engage youth in research: researchers' engagement experiences, barriers and capacity development priorities. *Health Expect* 2020;23:584–92.
62 Hawke LD, Relihan J, Miller J, *et al*. Engaging youth in research planning, design and execution: practical recommendations for researchers. *Health Expect* 2018;21:944–9.
63 Krause KR, Bear HA, Edbrooke-Childs J, *et al*. Review: what outcomes count? A review of outcomes measured for adolescent depression between 2007 and 2017. *J Am Acad Child Adolesc Psychiatry* 2019;58:61–71.
64 Jensen PS, Hoagwood K, Petti T. Outcomes of mental health care for children and adolescents: II. literature review and application of a comprehensive model. *J Am Acad Child Adolesc Psychiatry* 1996;35:1064–77.
65 Hoagwood KE, Jensen PS, Acri MC, *et al*. Outcome domains in child mental health research since 1996: have they changed and why does it matter? *J Am Acad Child Adolesc Psychiatry* 2012;51:1241–60.
66 Blanton H, Jaccard J. Arbitrary metrics in psychology. *Am Psychol* 2006;61:27–41.
67 Sechrest L, McKnight P, McKnight K. Calibration of measures for psychotherapy outcome studies. *Am Psychol* 1996;51:1065–71.
68 Charon R. Narrative medicine: a model for empathy, reflection, profession, and trust. *JAMA* 2001;286:1897–902.
69 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.
70 UNICEF. Measurement of mental health among adolescents at the population level (MMAP): an overview., 2019. Available: https://data.unicef.org/wp-content/uploads/2018/02/Mental-health-measurement_MMAP_overview_UNICEF.pdf [Accessed 5 Mar 2021].
71 Prinsen CAC, Mokkink LB, Bouter LM, *et al*. COSMIN guideline for systematic reviews of patient-reported outcome measures. *Qual Life Res* 2018;27:1147–57.
72 Stallwood E, Monsour A, Rodrigues C, *et al*. Systematic review: the measurement properties of the children's depression rating Scale-Revised in adolescents with major depressive disorder. *J Am Acad Child Adolesc Psychiatry* 2021;60:119–33.