BMJ Open

Digital self-report instruments for repeated measurement of mental health in the general adult population: a protocol for a systematic review

Zhao Hui Koh ,[1] Jason Skues,[2] Greg Murray[1]

To cite: Koh ZH, Skues J, Murray G. Digital self-report instruments for repeated measurement of mental health in the general adult population: a protocol for a systematic review. *BMJ Open* 2023;**13**:e065162. doi:10.1136/bmjopen-2022-065162

► 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-2022-065162).

Received 27 May 2022
Accepted 10 January 2023

[1]Centre for Mental Health, Swinburne University of Technology, Hawthorn, Victoria, Australia
[2]Department of Psychological Sciences, Swinburne University of Technology, Hawthorn, Victoria, Australia

Correspondence to
Zhao Hui Koh;
zhao.koh@gmail.com

ABSTRACT

Introduction Digital technologies present tremendous opportunities for enabling long-term measurement of mental health in the general population. Emerging studies have established preliminary efficacy of collecting self-report data digitally. However, a key challenge when developing a new self-report instrument is navigating the abundance of existing instruments to select relevant constructs for measurements. This review is a precursor to developing a novel future integrated digital instrument for repeated measurements. We interrogate the literature as the first step towards optimal measurement of the multifaceted mental health concept, in the context of digital repeated measurement. This review aims to identify (1) digital self-report instruments administered repeatedly to measure the mental health of the general adult population; (2) their structure and format; (3) their psychometric properties; (4) their usage in empirical studies; and (5) the constructs these instruments were designed to measure (as characterised in the original publication), and the constructs the instruments have been used to measure in the identified empirical studies.

Methods and analysis Five major electronic databases will be searched. Studies administering mental health instruments (in English) repeatedly to community dwellers in the general adult population are eligible. A reviewer will preliminarily screen for eligible studies. Then, two reviewers will independently screen the full text of the eligible articles and extract data. Both reviewers will resolve any disagreement through discussion or with a third reviewer. After the data extraction, a reviewer will manually search for the structure, format, psychometric properties and the original constructs these instruments were developed to measure. This review will synthesise the results in a narrative approach. The reporting in this review will be guided by the Preferred Reporting Items for Systematic Reviews and Meta-Analyses checklist.

Ethics and dissemination Ethical approval is not required as no data will be collected. Findings of the systematic review will be disseminated through peer-reviewed publications and conference presentations.

PROSPERO registration number CRD42022306547

STRENGTHS AND LIMITATIONS OF THIS STUDY

⇒ This review and its search strategy are guided by a broad definition of mental health informed by a well-recognised conceptual framework, potentially improving the relevance of the review results towards public mental health.
⇒ This review protocol establishes a defensible framework referencing prominent frameworks and taxonomies to develop mid-level terms to guide the search strategy and eligibility criteria in this review.
⇒ This review is guided by the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) checklist,[68] which will improve the quality, transparency and comprehensiveness of the review.
⇒ Exclusion of feasibility and pilot studies may reduce the likelihood of detecting emerging instruments in the review results.
⇒ The restriction to English-only instruments may limit the generalisability of the results.

INTRODUCTION

Mental health is a multifaceted concept with fuzzy boundaries.[1–10] Constructs subsumed by 'mental health' include 'mental illness', 'mental well-being',[11 12] 'psychological well-being'[13] and 'mental wellness'.[14 15] In the context of this conceptual complexity, a vast array of literature reports on empirical studies aiming to measure aspects of mental health using various self-report instruments. Several critics have noted that there is a remarkable lack of argument concerning *the choice of self-report measurement instruments for the wide variety of study aims*.[16–21] The relatively new domain of digital data collection (collecting such self-report data through digital technologies such as online websites, mobile applications and self-service health kiosks) is not exempt from this criticism,[22 23] with many studies apparently treating selection of a self-report mental health measure as uncontentious.

The overarching aim of the present review is to advance understanding of the measurement of the multifaceted construct 'mental health' by conducting (to our knowledge) the first audit of instruments focusing on

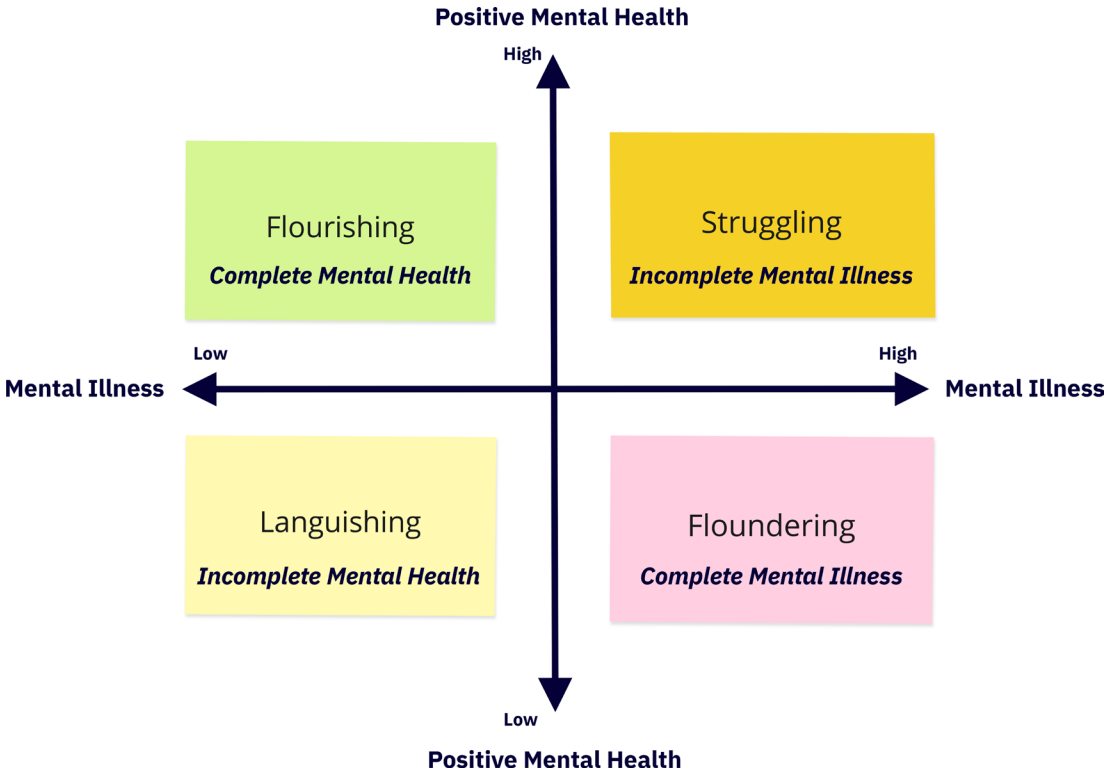

Figure 1 Complete state model of mental health (Keyes[69]). Source: adapted from Teng *et al.*[70]

digital delivery and repeated measurement of mental health in the general adult population. Such instruments are becoming more common and more important across the landscape of mental health and well-being in digital monitoring,[24] digital assessments,[25] digital phenotyping,[26] self-management[27] and so on in research and practice. The outcomes from the present review will expand on and update previous reviews conducted on mental health instruments (eg, Breedvelt *et al* and Beidas *et al*[28 29]) and ultimately be used to inform the development of a new digital mental health instrument for monitoring purposes in the general population. This new digital instrument will be developed using item response theory[30 31] and adopting the Patient-Reported Outcomes Measurement Information System (PROMIS) instrument development and validation scientific standards.[32] Once the digital instrument is developed, it will be integrated into our commercial partner's health kiosk and online portal,[33] complementing their physical well-being measurements. The present review is guided by the question: What is the optimal operationalisation of mental health in the context of *digital delivery of assessment* and *repeated measurement* in the general population?

This review will take a broad definition of 'mental health' to target the general adult population, which covers the entire spectrum of mental health phenomena. Informed by the complete state model of mental health[2 34] (see figure 1), this broad definition of mental health encompasses two correlated but separate dimensions: *positive mental health* (eg, flourishing, satisfaction of life, hedonic/emotional well-being, psychological well-being, social well-being) and *mental illness* (eg, anxiety, depression, post-traumatic stress disorder, psychosis, schizophrenia). We recognise that there are subgroups within the general population with different health characteristics such as individuals who have depression and individuals who are living with chronic illnesses (eg, diabetes, cancer). Depression is prevalent in the general adult population in Australia,[35] and we expect our new instrument will measure it. By adopting the broad definition of mental health in guiding the search strategy, this review can improve the relevancy of the systematic review and increase the ecological validity of the new digital instrument for the general population, including subgroups with different health characteristics.

One consequence of the present study focusing on repeated measure instruments is a slightly expanded approach to the investigation of instruments' format (eg, survey length, response options) and psychometric properties. Indeed, there are considerable differences between the desirable properties of instruments that are used for longitudinal measures (repeated) and instruments that are used for cross-sectional (one-off) measures.[36] For example, format-wise, a small number of items (between 1 and 40 items) and multiple response options (eg, 7-point Likert scales) are recommended for an instrument used for repeated measures. In contrast, a large number of items (between 20 and 100 items) and a smaller number of response options (eg, binary or at most three options) are recommended for an instrument used for cross-sectional (one-off) measures.[36] Similarly, as an instrument is repeatedly administered in a longitudinal study, this

instrument's ability to detect meaningful changes (shifts) over time is important compared with an instrument used for one-off measurement.[23 37] Furthermore, as most of the general public in Australia (the target population of our new instrument) are not diagnosed with mental disorders,[38] the existence of floor and ceiling effects in an instrument could jeopardise its overall utility.

The focus on the digital delivery of instruments in this review also raises novel questions about the influence of digital format on the instruments' psychometric properties. Specifically, the historical development of an instrument, that is, whether an instrument is developed with digital delivery in mind or adapted from an original instrument previously designed for other modalities such as pen-and-paper or face-to-face/telephone interview. While past literature has established the feasibility and acceptability of using digital measurements as part of the monitoring routines in psychiatric treatments,[25 39] complemented by an abundance of mobile apps that enable symptoms monitoring,[40] questions remain on how the digital delivery/modality may influence the psychometric properties of an instrument compared with their equivalent conventional counterparts (eg, pen-and-paper). Although the comparison between different modalities is out of scope in this review, the sole focus on digital instruments in this review will provide comparable insights into their structure, format, psychometric properties and usage and measurement aims.

A key methodological challenge this review must plan for is the abundance of self-report mental health instruments in common use (research and practice) due to the multifaceted nature of the mental health concept and its heterogeneity. To address this challenge, this review will reference prominent conceptual frameworks and taxonomies in the mental health domain to guide the search strategy and eligible criteria (see Methods and Analysis section). The broad mental health constructs measured by current instruments could span from symptoms of mental disorders[28 41] to mental well-being.[15]

Instruments that measure the symptoms of mental disorders are commonly based on two widely accepted taxonomies—the Diagnostic and Statistical Manual of Mental Disorders (DSM-5[42]) and the International Classification of Diseases (ICD-11[43]), such as the 9-item Patient Health Questionnaire[44] measuring the severity of depression symptoms and the 7-item General Anxiety Disorder Questionnaire[45] screening symptoms of the Generalised Anxiety Disorder. In contrast, instruments that measure mental well-being vary depending on the constructs,[28] ranging from mental states (eg, happiness, emotional well-being, psychological well-being, social well-being), cognitive evaluation (life satisfaction, life meaning), protective factors (eg, resilience, optimism, hope, compassion) and risk factors (eg, stress, sleep quality, traumatic experience). For example, the 5-item Satisfaction with Life Scale[46] measures global evaluation of one's life and the 8-item Flourishing Scale[47] measures perceived success in one's life areas (eg, self-esteem, purpose and optimism). The abundance of mental health self-report instruments also reveals another phenomenon, in which one or more instruments can measure a construct. For example, at least 11 self-report instruments are available to assess the severity of depression with varying degrees of measurement precision, range and target population.[48] Similarly, at least 92 self-report instruments are available to measure anxiety.[49] To navigate and manage the complexity among this large pool of mental health instruments and constructs, this protocol will put together a defensible framework (described in the 'Methods and analysis' section), referencing prominent conceptual frameworks and taxonomies to guide the search strategy and eligibility criteria.

This review also aims to generate insights into the relationship between mental health measures and mental health constructs, as exemplified by the studies identified here. Specifically, we will, where possible, extract from text in empirical studies the constructs authors were intending to measure with a given instrument, and compare them to the constructs the instrument developers were intending to capture (as described in the text of the original validation article). This comparison may highlight potential validity mismatches between the phenomena that instruments were developed to measure, the phenomena measured by empirical studies and participants' views of the constructs.[17] Although studying this mismatch is out of scope for the present review, we will extract constructs that instruments *were used to measure empirically* in identified studies, as well as constructs that the instruments were originally *designed to measure*. The preliminary analysis and data extracted about this phenomenon in this review could provide the groundwork for future reviews and potentially increase the awareness of instrument selection for future studies.

Importance of this review

Previous and ongoing systematic review efforts have investigated instruments used in non-clinical adult populations for mental disorder diagnosis,[41] symptoms screening[28] and mental well-being or similar constructs (eg, subjective well-being)[21 50–53] grounded in different conceptual frameworks. A preliminary search on PROSPERO (conducted 5 November 2021) also revealed several similar reviews of public mental health instruments (completed or ongoing reviews), such as[54] on public mental health outcome measures in the UK and[15] on mental wellness of adolescents.

To the best of our knowledge, there is no review (or protocol) that shares the specific focus of the present review, that is, digital self-report mental health instruments for repeated measurements in the general adult population.

Research objectives and review questions

The main objective of this review is to systematically identify past empirical studies in which mental health in the general population was measured (1) using digital

self-report instruments, and (2) in repeated measures designs. We will extract from identified empirical studies information about (1) the structure, (2) format, (3) psychometric properties, (4) frequency of use and (5) the construct instruments were used to *measure* in these studies, as well as the constructs the instruments were *designed* to measure in the original publication. The review will study these instruments from different perspectives through five research questions:

1. What digital self-report mental health instruments are used in repeated measures designs (more than one time point, either within-person or within-group) in the general adult population?
2. What is the structure (eg, dimensions, subscales, etc) and format (eg, number of question items, response format, instructions, etc) of the instruments identified in Question 1?
3. What are the psychometric properties (eg, reliability, validity, responsiveness, etc, and norms used) of the instruments identified in Question 1 (defined in the original publication and other relevant studies)?
4. What is the frequency of use of the instruments identified in Question 1 among the selected studies? The usage of the instrument in this review is operationalised as the frequency of the instrument being administered in identified empirical studies over the number of years since the release of this instrument, bounded by the time frame of this review.
5. Which mental health construct(s) are the instruments identified in Question 1 intended to measure in the identified empirical study (as described in the empirical study), and which mental health construct(s) were the instruments originally developed to measure?

METHODS AND ANALYSIS

This protocol adheres to the Preferred Reporting Items for Systematic Review and Meta-analysis Protocols (PRISMA-P) 2015 guidelines.[55 56]

Patient and public involvement

No patient involved.

Inclusion and exclusion criteria

This review will include studies that match the inclusion criteria in the PICOT format[57] defined below.

Population

All studies involving community dwellers in the general adult population will be included. Adulthood will be defined as 18 and above, and studies with samples including adults aged 18 or above, for example, a study that recruits participants (young adults) aged between 16 and 25 will be included. In addition, studies involving subpopulation groups, for example, grouped by occupational or sociodemographical, will also be included. Studies of clinical populations, including those with physical and mental health conditions will be included, as long as the sample population is living in the community.

Studies that exclusively target infants, children, adolescents and individuals not residing in the community (eg, inpatients, prisoners, military personnel during deployment) will be excluded.

Intervention of interest

Interventions are not a focus of this review. All peer-reviewed empirical studies reporting on the administration of digital (online websites, mobile apps, health kiosks) self-report and self-administered mental health instruments in English at more than one time point will be included. This covers empirical studies measuring mental health constructs of the same individuals (within-person) or the same groups (within-group) over time. Study designs are likely to include population-based longitudinal studies, repeated cross-sectional studies, multiwaves surveys, cohort studies (retrospective/prospective), case-control studies, mixed-method studies, scale evaluation studies, quantitative randomised and non-randomised controlled trials (pre-post measurements after any intervention/treatment, eg, psychological, medication).

Studies that administered a mental health instrument that is non-English (including translated instruments from English) or at one time point only (eg, single-wave cross-sectional survey, screening participants) will be excluded. Secondary analyses of previously collected surveys, feasibility, pilot, proof-of-concept, exploratory studies, qualitative studies, case studies and protocol for research studies or review protocols will be excluded.

Included empirical studies must be published in English, in the format of peer-reviewed journal articles. Review articles (systematic review, literature review, scoping review, integrative review, meta-analyses) will be excluded. Studies that were not peer-reviewed, not published in English or published as preprints, case reports, opinions, conceptual or theoretical discussion articles will also be excluded.

Included studies must have at least one self-report digital instrument in English measuring mental health at more than one time point (except studies that administered a single wave of a national survey that had been administered in the past). Screening instruments are eligible. Studies that used self-report instruments as third-party observations such as proxy-report instruments (eg, parent's self-report on a child's behaviours) will be excluded.

The mental health constructs of interest in this review are guided by frameworks defined in the section 'Conceptual Frameworks and Taxonomies' (see figure 2 and online supplemental appendix 1). As shown in figure 2, syndromes (mental disorders) defined in Forbes *et al*[58] guided by HiTOP are included. Personality disorders were excluded because they are closely related to personality, which is a trait that is generally considered stable over time in adults.[59] In this review, we are interested in state-like constructs and their level of change across time. Furthermore, personality disorders could manifest as symptoms of other psychopathology such as anxiety

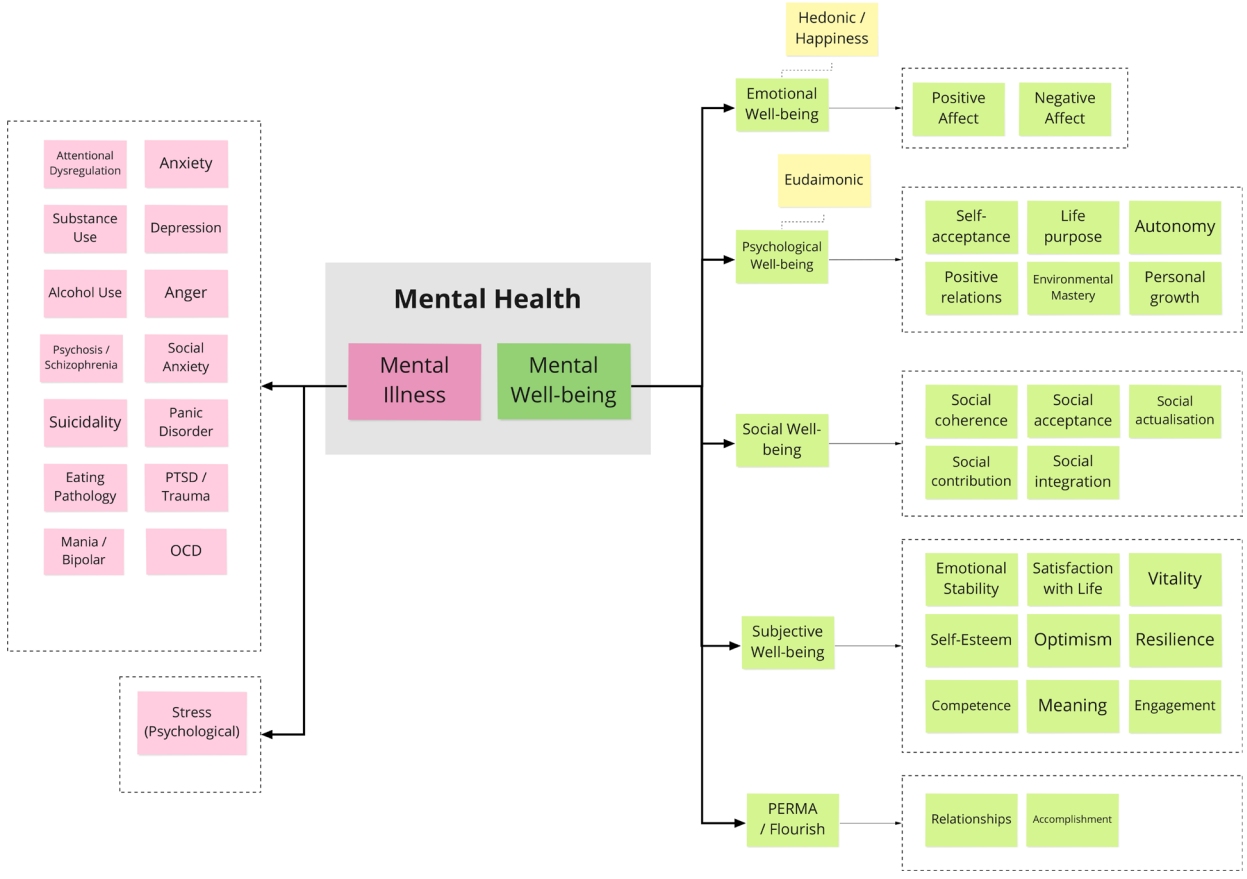

Figure 2 The conceptualisation of mental health in this review. This diagram depicts our conceptualisation of mental health used in this protocol to guide our search strategy and formulation of search terms.

and depression, which will be included in this review.[58 60] We also included psychological stress because stress is commonly being recognised to precipitate anxiety and depression and it has also been found as a separate factor while analysing anxiety and depression scales.[61] All these syndromes will form the search terms in this review.

While the syndrome will guide the search strategy, the symptoms of each syndrome will be used to guide the inclusion criteria of instruments. For example, symptoms of psychopathology may include sleep difficulties, worry, obsessive thoughts, fear, nervousness, fear of losing control and so on.[62] Instruments that measure some of these symptoms in the eligible studies will be included in this review if the symptoms are part of the included syndromes. For example, if the search term 'anxiety' retrieves an empirical study that administered the Locus of Control Scale,[63] this instrument will be included in this review because it is associated with the symptom 'fear of losing control'. To reduce the complexities of the search terms, we intentionally exclude these symptoms from the search strategy. In this review, we want to be guided by a defensible framework that we could refer to when developing search terms due to the complexity of the mental health concept.

Similarly, mental well-being constructs and subconstructs defined in figure 2 will be used as search terms. The definition of these subconstructs based on the original theoretical publications or existing scale items (if any) will be used to determine the eligibility of an instrument.

Given the heterogeneous nature of mental health constructs, we anticipate there will be ambiguous scenarios ('grey areas'), where reviewers will be unable to decide on the eligibility of an article or instrument based on the framework above. If such a situation should arise, it will be considered on a case-by-case basis through discussions among reviewers and a third researcher (GM). Through this process, we will discover, learn and report on these ambiguous cases, which could clarify or generate further hypotheses about the boundaries of mental health.

Comparison intervention

Comparison intervention is not applicable in this systematic review as we are only interested in the mental health instruments administered in empirical studies.

Primary outcome

The primary outcome of this systematic review is a comprehensive list of self-report digital mental health instruments that have been used in repeated measures designs. Instruments will be characterised in terms of their structure and format, psychometric properties, their relative usage in the literature and their measurement aims (as described in text of the identified empirical study, and the instrument's original validation paper).

Timeframe

Only articles published between 1 January 2010 and 31 December 2021 (inclusive) will be included.

Search strategy

We will search five major electronic databases systematically including Scopus, Web of Science, PubMed, PsychInfo and Psychology & Behavioral Sciences collection (via EBSCOhost) for studies to answer Question 1. For Questions 2, 3 and 5, additional targeted manual searches will be performed to elicit the structure, format, psychometric properties and measurement aims of instruments identified from Question 1. Similar to another ongoing PROSPERO review (https://www.crd.york.ac.uk/prospero/display_record.php?RecordID=186218), we will manage our searches in multiple stages to answer the research questions.

Stage 1: identify relevant articles

The search strategy is developed based on the eligibility criteria defined above and will be peer-reviewed by a librarian from Swinburne University of Technology who has expertise in developing searches for systematic reviews. The search strategy consists of four high-level concepts: mental health, instruments, repeated measures and digital/online (see online supplemental appendix 2 for a sample of the search terms for each concept). The search terms will target the title, abstract and keywords and combine with Boolean operators. If appropriate, snowballing strategies and a custom selection of articles will be used and documented.

Stage 2: identify structure and format

The structure and format that each instrument was developed to measure will be identified through the original paper that documents the development of the instrument. This stage will be performed mainly using the citations (reference list) from the publications of the selected empirical studies identified in Stage 1. If the reference is unavailable, a targeted manual search will be performed to retrieve the original paper.

Stage 3: identify psychometric properties and relevant norms

The psychometric properties (reliability, validity, responsiveness, etc) of each identified instrument will be reviewed through the original and subsequent validation studies (including reviews and meta-analyses). A targeted manual search will be performed to retrieve these studies that evaluated the instrument targeting the general adult population.

Stage 4: identify measurement aims

Similar to Stage 2, Stage 4 will use similar mechanism to identify the constructs of mental health of each identified instruments from the original study.

The search protocols will be published in a transparent and reproducible manner. Once the search strategy is finalised, it will be adapted and run across the designated databases.

Study selection process

Summary information on the matching articles (including the abstract) will be extracted from the search result and imported into Microsoft Excel for preliminary screening. Duplicate items will be manually removed. Preliminary screening of eligible articles will be performed in two passes by one reviewer (ZHK). Under the supervision of JS and GM, ZHK will screen titles and abstracts for all studies extracted by labelling them as either 'relevant' or 'not relevant'. ZHK will record rejection reasons for articles that are marked 'not relevant'.

After the initial screening above, two reviewers (ZHK and one other researcher) will independently extract and screen the full text of the eligible studies. If it is still eligible, the reviewers will mark it as 'relevant' and it will be ready for data extraction. Reasons for rejecting studies will be recorded. If anything is unclear about the full-text articles, the corresponding authors will be contacted for further details. If there is a disagreement between the two reviewers, it will be resolved by consulting a third party (GM). The inclusion and exclusion processes will be reported in a PRISMA flow chart.

Data extraction process

Before commencing the data extraction process, a data extraction spreadsheet will be developed (guided by the Cochrane Consumers and Communication Review Group's data extraction template; only relevant sections to this review are included) and used by the two reviewers. Both reviewers will test the process using 10 random included articles and then refine the extraction sheet accordingly. In the test, one reviewer will extract data items, and the other reviewer will check the data. Disagreements will be resolved if they arise. If the disagreements cannot be resolved among the two reviewers, a third party will be consulted (GM). The reviewers will contact the authors of included studies if there is anything unclear.

Separate Excel spreadsheets will be used for data extraction. Each person will review the studies independently with a separate copy of the spreadsheet and maintain review records through Microsoft Excel. The following information will be extracted from eligible articles:

- Data source (eg, electronic database names or custom selection)
- Study information:
  i. Title
  ii. Authors
  iii. Year of publication
  iv. Journal
  v. Study type (eg, cross-sectional, longitudinal prospective, retrospective).
  vi. Study design
  vii. Data collection periods and survey administration frequencies
- Sample information:
  i. Target population, sample size.
  ii. Demographics (sex, age range, country, region).

iii. Settings (rural, urban, community, home).
- Instrument information:
  i. Administration mode (web, tablet or mobile app).
  ii. Structure (factors, dimensions and any subscale).
  iii. Format (number of items and response format) including any modification done to the original scale.
  iv. The version used (short-form, long-form or version number).
  v. Terms of use (free, paid, ask for permissions, non-commercial).
  vi. The construct(s) that the instrument was used to measure (in the identified study).
  vii. The rationale for choosing the instruments to measure mental health (if any).
  viii. Theoretical/conceptual framework of the instrument (if mentioned in the article).

After data extraction is completed, all instruments will be consolidated. A usage score will be calculated for each instrument based on the operationalisation defined above. To answer Question 4, the distribution of the usage score of these instruments will be summarised and discussed. To answer Questions 2, 3 and 5, targeted searches of the instruments or citations about the instruments will be performed. The following information will be extracted for each instrument:

- The structure, format of the instrument (as defined in the original scale development publication)
- The psychometric properties (eg, reliability, validity, responsiveness) of the instrument.
- The construct(s) that the instrument was designed to measure (as defined in the original scale development publication)

Risk of bias assessment

Like Zamperoni *et al*[64] and Breedvelt *et al*[28] this review is concerned with the digital instruments that have been used in past studies that repeatedly measure mental health, rather than the execution of the studies themselves. The risk of bias and quality assessment of each instrument identified will be assessed by the psychometric properties and the norm population available in answering research Question 3. Besides the reliability (eg, internal consistency, test-retest reliability), validity (eg, content validity, criterion validity, construct validity) and responsiveness, we may also consider quality criteria suggested in Terwee *et al*[65] and Reyes *et al*,[66] such as the normative populations, longitudinal validity, floor and ceiling effects and interpretability of each instrument (if the information is available in the literature).

Data synthesis

We will summarise findings among the mental health instruments identified pertinent to the research questions in this review using a narrative synthesis approach.[67] For each instrument identified (Question 1), targeted manual searches will be conducted to elicit the structure, format (Question 2) and psychometric properties of the instrument (Question 3) before the information is synthesised through the original and subsequent validation studies (including reviews and meta-analyses). The usage of the instrument (Question 4) will be presented as descriptive statistics. Finally, the mental health constructs (Question 5) that the instrument was used to measure in eligible studies will be compared with original mental health construct that it was developed based on the original publication that describes the development methodology.

Twitter Zhao Hui Koh @zhaohuik

Acknowledgements We would like to acknowledge librarian Mr David Bradley from Swinburne University of Technology for his valuable input in formulating the search strategy for this systematic review protocol.

Contributors GM is the guarantor. ZHK drafted the initial manuscript, revised and reviewed by GM and JS. All authors contributed to the research questions, selection criteria, the risk of bias assessment strategy and data extraction criteria. ZHK developed the search strategy with the librarian (acknowledged above). GM and JS reviewed the search strategy. All authors read, provided feedback and approved the final manuscript.

Funding Digital Health CRC Limited and SiSU Health co-fund this project through a PhD scholarship (Reference Number DHCRC-0049) and support the project costs for data collection, data management and analyses. Digital Health CRC Limited is funded under the Commonwealth Government's Cooperative Research Centres Program. Digital Health CRC Limited and SiSU Health are not involved in any other aspects of the project and have no input on the interpretation or publication of the study results.

Competing interests None declared.

Patient and public involvement Patients and/or the public were not involved in the design, or conduct, or reporting or dissemination plans of this research.

Patient consent for publication Not applicable.

Provenance and peer review Not commissioned; externally peer reviewed.

ORCID iD

Zhao Hui Koh http://orcid.org/0000-0001-9533-6495

REFERENCES

1 World Health Organization. *Promoting mental health: concepts, emerging evidence, practice: A report of the World Health Organization, Department of Mental Health and Substance Abuse in collaboration with the Victorian Health Promotion Foundation and the University of Melbourne*. 2005.
2 Keyes CLM. Mental illness and/or mental health? Investigating axioms of the complete state model of health. *J Consult Clin Psychol* 2005;73:539–48.
3 Manderscheid RW, Ryff CD, Freeman EJ, *et al*. Evolving definitions of mental illness and wellness. *Prev Chronic Dis* 2010;7:A19. Available: https://www.ncbi.nlm.nih.gov/pubmed/20040234

4 Mechanic D. *The murky challenge of mental health. in: title: the truth about health care: why reform is not working in America*. New Brunswick, N.J: Rutgers University Press, 2006: 67–83.
5 Huber M, Knottnerus JA, Green L, *et al*. How should we define health? *BMJ* 2011;343:d4163bmj.d4163.
6 Galderisi S, Heinz A, Kastrup M, *et al*. A proposed new definition of mental health [propozycja nowej definicji zdrowia psychicznego]. *Psychiatr Pol* 2017;51:74145:407–11:.
7 Manwell LA, Barbic SP, Roberts K, *et al*. What is mental health? Evidence towards a new definition from a mixed methods multidisciplinary international survey. *BMJ Open* 2015;5:e007079.
8 Cyranka K. Controversial issues in current definitions of mental health. *Arch Psych Psych* 2020;22:7–11.
9 Granlund M, Imms C, King G, *et al*. Definitions and operationalization of mental health problems, wellbeing and participation constructs in children with NDD: distinctions and clarifications. *Int J Environ Res Public Health* 2021;18:1656.
10 Lundqvist C, Andersson G. Let's talk about mental health and mental disorders in elite sports: a narrative review of theoretical perspectives. *Front Psychol* 2021;12:700829.
11 Hides L, Quinn C, Stoyanov S, *et al*. Testing the interrelationship between mental well-being and mental distress in young people. *The Journal of Positive Psychology* 2020;15:314–24.
12 Hutcheson C, Ferguson H, Nish G, *et al*. Promoting mental wellbeing through activity in a mental health hospital. *British Journal of Occupational Therapy* 2010;73:121–8.
13 Kawada T. Feeling refreshed by sleep can predict psychological wellbeing assessed using the General Health Questionnaire in male workers: a 3-year follow-up study. *Psychiatry Investig* 2012;9:418–21.
14 Zoorob D, Shah S, La Saevig D, *et al*. Insight into resident burnout, mental wellness, and coping mechanisms early in the COVID-19 pandemic. *PLoS One* 2021;16:e0250104.
15 Orth Z, van Wyk B. Adolescent mental wellness: A systematic review protocol of instruments measuring general mental health and well-being. *BMJ Open* 2020;10:e037237.
16 Al Sayah F, Jin X, Johnson JA. Selection of patient-reported outcome measures (PROMs) for use in health systems. *J Patient Rep Outcomes* 2021;5(Suppl 2):99):99.:.
17 Records K, Keller C, Ainsworth B, *et al*. Instrument selection for randomized controlled trials: why this and not that? *Contemp Clin Trials* 2012;33:143–50.
18 Dodd AL, Priestley M, Tyrrell K, *et al*. University student well-being in the United Kingdom: a scoping review of its conceptualisation and measurement. *Journal of Mental Health* 2021;30:375–87.
19 Etkin RG, Shimshoni Y, Lebowitz ER, *et al*. Using evaluative criteria to review youth anxiety measures, part I: self-report. *J Clin Child Adolesc Psychol* 2021;50:58–76.
20 Silverman WK, Ollendick TH. Evidence-based assessment of anxiety and its disorders in children and adolescents. *J Clin Child Adolesc Psychol* 2005;34:380–411.
21 Linton M-J, Dieppe P, Medina-Lara A. Review of 99 self-report measures for assessing well-being in adults: exploring dimensions of well-being and developments over time. *BMJ Open* 2016;6:e010641.
22 Alfonsson S, Maathz P, Hursti T. Interformat reliability of digital psychiatric self-report questionnaires: a systematic review. *J Med Internet Res* 2014;16:e268.
23 Kroenke K, Miksch TA, Spaulding AC, *et al*. Choosing and using patient-reported outcome measures in clinical practice. *Arch Phys Med Rehabil* 2022;103:S0003-9993(21)00227-6:S108–17:.
24 Alqahtani F, Al Khalifah G, Oyebode O, *et al*. Apps for mental health: an evaluation of behavior change strategies and recommendations for future development. *Front Artif Intell* 2019;2:30:30:.
25 Richter MF, Storck M, Blitz R, *et al*. Repeated digitized assessment of risk and symptom profiles during inpatient treatment of affective disorder: observational study. *JMIR Ment Health* 2020;7:e24066.
26 Insel TR. Digital phenotyping: a global tool for psychiatry. *World Psychiatry* 2018;17:276–7.
27 Lupton D. The digitally engaged patient: self-monitoring and self-care in the digital health era. *Soc Theory Health* 2013;11:256–70.
28 Breedvelt JJF, Zamperoni V, South E, *et al*. A systematic review of mental health measurement scales for evaluating the effects of mental health prevention interventions. *Eur J Public Health* 2020;30:539–45.
29 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.
30 Embretson SE, Reise SP. *Item response theory for psychologists*. Psychology Press, 2014.
31 Lord FMN. *Statistical theories of mental test scores*. Addison-Wesley, 1966.
32 PROMIS. PROMIS instrument development and validation scientific standards. 2013. Available: https://www.healthmeasures.net/images/PROMIS/PROMISStandards_Vers2.0_Final.pdf
33 SiSU Health. SiSU Health website. 2020. Available: https://sisuhealthgroup.com/
34 Keyes C. Mental health as a complete state: how the salutogenic perspective completes the picture. In: Bauer GF, Hämmig O, eds. *Bridging occupational, organizational and public health*. Dordrecht: Springer Netherlands, 2014: 179–92.
35 Australian Institute of Health and Welfare. Mental health. 2020. Available: https://www.aihw.gov.au/reports/australias-health/mental-health
36 Hyland ME. A brief guide to the selection of quality of life instrument. *Health Qual Life Outcomes* 2003;1:24.
37 Guyatt G, Walter S, Norman G. Measuring change over time: assessing the usefulness of evaluative instruments. *J Chronic Dis* 1987;40:171–8.
38 Department of Health. National primary health networks guidance: initial assessment and referral for mental healthcare. 2020. Available: https://www.health.gov.au/sites/default/files/documents/2021/04/national-phn-guidance-initial-assessment-and-referral-for-mental-health-care.pdf
39 Faurholt-Jepsen M, Vinberg M, Frost M, *et al*. Smartphone data as an electronic biomarker of illness activity in bipolar disorder. *Bipolar Disord* 2015;17:715–28.
40 Sucala M, Cuijpers P, Muench F, *et al*. Anxiety: there is an app for that. A systematic review of anxiety apps. *Depress Anxiety* 2017;34:518–25.
41 Newson JJ, Hunter D, Thiagarajan TC. The heterogeneity of mental health assessment. *Front Psychiatry* 2020;11:76:76:.
42 American Psychiatric Association. *Diagnostic and statistical manual of mental disorders*. Washington, DC: American Psychiatric Association, 22 May 2013.
43 World Health Organization. World health organization. ICD-11: international classification of diseases. 2019. Available: https://icd.who.int/
44 Kroenke K, Spitzer RL. The PHQ-9: a new depression diagnostic and severity measure. *Psychiatric Annals* 2002;32:509–15.
45 Spitzer RL, Kroenke K, Williams JBW, *et al*. A brief measure for assessing generalized anxiety disorder: the GAD-7. *Arch Intern Med* 2006;166:1092–7.
46 Diener E, Emmons RA, Larsen RJ, *et al*. The satisfaction with life scale. *J Pers Assess* 1985;49:71–5.
47 Diener E, Wirtz D, Tov W, *et al*. New well-being measures: short scales to assess flourishing and positive and negative feelings. *Soc Indic Res* 2010;97:143–56.
48 Wahl I, Löwe B, Bjorner JB, *et al*. Standardization of depression measurement: a common metric was developed for 11 self-report depression measures. *J Clin Epidemiol* 2014;67:S0895-4356(13)00316-8:73–86:.
49 McHugh RK, Rasmussen JL, Otto MW. Comprehension of self-report evidence-based measures of anxiety. *Depress Anxiety* 2011;28:607–14.
50 Rose T, Joe S, Williams A, *et al*. Measuring mental wellbeing among adolescents: a systematic review of instruments. *J Child Fam Stud* 2017;26:2349–62.
51 Martín-María N, Lara E, Cresswell-Smith J, *et al*. Instruments to evaluate mental well-being in old age: a systematic review. *Aging Ment Health* 2021;25:1191–205.
52 Lindert J, Bain PA, Kubzansky LD, *et al*. Well-Being measurement and the WHO health policy health 2010: systematic review of measurement scales. *Eur J Public Health* 2015;25:731–40.
53 Ong ZX, Dowthwaite L, Perez Vallejos E, *et al*. Measuring online wellbeing: a scoping review of subjective wellbeing measures. *Front Psychol* 2021;12:616637:616637:.
54 Hulin J, Weich S, Mukuria C, *et al*. *A systematic review of public mental health outcome measures*. 2019. Available: https://www.crd.york.ac.uk/prospero/display_record.php?RecordID=186218
55 Moher D, Shamseer L, Clarke M, *et al*. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Syst Rev* 2015;4:1.
56 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;350:g7647.
57 Riva JJ, Malik KMP, Burnie SJ, *et al*. What is your research question? An introduction to the PICOT format for clinicians. *J Can Chiropr Assoc* 2012;56:167–71.

58 Forbes MK, Sunderland M, Rapee RM, *et al*. A detailed hierarchical model of psychopathology: from individual symptoms up to the general factor of psychopathology. *Clin Psychol Sci* 2021;9:139–68.
59 Cobb-Clark DA, Schurer S. The stability of big-five personality traits. *Economics Letters* 2012;115:11–5.
60 Forbes MK, Kotov R, Ruggero CJ, *et al*. Delineating the joint hierarchical structure of clinical and personality disorders in an outpatient psychiatric sample. *Compr Psychiatry* 2017;79:S0010-440X(17)30051-2:19–30:.
61 Lovibond SH, Lovibond PF. *Manual for the depression anxiety stress scales*. Psychology Foundation of Australia, 1996.
62 Forbes F-J, Whisenhunt BL, Citterio C, *et al*. Making mental health a priority on college campuses: implementing large scale screening and follow-up in a high enrollment gateway course. *J Am Coll Health* 2021;69:275–82.
63 Rotter JB. Generalized expectancies for internal versus external control of reinforcement. *Psychol Monogr* 1966;80:1–28.
64 Zamperoni V, Breedvelt J, Kousoulis A, *et al*. A systematic review of mental health measurement scales for the assessment of the effects of mental health prevention interventions. *PROSPERO* 2018; Available: https://www.crd.york.ac.uk/prospero/display_record.php?RecordID=95519
65 Terwee CB, Bot SDM, de Boer MR, *et al*. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol* 2007;60:34–42.
66 Reyes ADL, Langer DA. Assessment and the journal of clinical child and adolescent psychology's evidence base updates series: evaluating the tools for gathering evidence. *J Clin Child Adolesc Psychol* 2018;47:357–65.
67 Popay J, Roberts H, Sowden A, *et al*. Guidance on the conduct of narrative synthesis in systematic reviews. In: *A product from the ESRC methods programme Version 2006;1:b92*.
68 Page MJ, McKenzie JE, Bossuyt PM, *et al*. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ* 2021;372:n71.
69 Keyes CLM. The mental health continuum: from languishing to flourishing in life. *J Health Soc Behav* 2002;43:207–22.
70 Teng E, Venning A, Winefield H, *et al*. Half full or half empty: the measurement of mental health and mental illness in emerging Australian adults. *SIIW* 2015;1:1.