BMJ Open

How to distinguish promotion, prevention and treatment trials in public mental health? Study protocol for the development of the VErona-LUgano Tool (VELUT)

Marianna Purgato ,[1] Emiliano Albanese,[2] Davide Papola,[1,3] Eleonora Prina ,[1] Federico Tedeschi,[1] Alden Gross ,[4,5] Marit Sijbrandij,[6] Ceren Acarturk,[7] Anna Maria Annoni,[2] Manuela Silva,[8] Mark J D Jordans,[9,10,11] Crick Lund,[9,12] Wietse A Tol,[13,14] Pim Cuijpers ,[6] Corrado Barbui[1]

To cite: Purgato M, Albanese E, Papola D, *et al*. How to distinguish promotion, prevention and treatment trials in public mental health? Study protocol for the development of the VErona-LUgano Tool (VELUT). *BMJ Open* 2024;**14**:e082652. doi:10.1136/bmjopen-2023-082652

► Prepublication history and additional supplemental material for this paper are available online. To view these files, please visit the journal online (https://doi.org/10.1136/bmjopen-2023-082652).

MP and EA are joint first authors.

Received 29 November 2023
Accepted 23 July 2024

Check for updates

© Author(s) (or their employer(s)) 2024. 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
Professor Marianna Purgato; marianna.purgato@univr.it

ABSTRACT

Introduction Promoting mental health, preventing and treating mental disorders are critically important in public health, and many randomised controlled trials (RCTs) evaluate intervention strategies for these objectives. However, distinguishing promotion from prevention and from treatment RCTs is challenging. A tool to place studies along the promotion-to-treatment continuum in mental health research does not exist, leaving it to researchers and policymakers to decide on how to classify individual RCTs, which hinders evidence synthesis.

Methods and analysis We present a protocol for the development of a new tool to assist researchers in distinguishing RCTs along the promotion-to-treatment continuum. We will establish a Tool Development Group, and use the Population, Intervention, Comparison and Outcome framework to define constructs. We will generate, define, categorise and reduce the items in the tool using qualitative methods, including cognitive interviews and a Delphi exercise. Psychometric evaluation—including unidimensionality, local independence, monotonicity and item homogeneity—will include data collection, scoring, internal consistency checks and factor analysis of the tool's indicators for available RCTs. We will use standard Cohen's kappa statistics to assess the reliability of the tool.

Ethics and dissemination This study involves data collection from the already published literature. However, this protocol has been approved by the ethics committee of the Università della Svizzera Italiana (CE 2024 04). The results of the present project will be disseminated in peer-reviewed journals and at international and national scientific meetings. Training materials for the application of the tool will also be developed and disseminated to the scientific community. The tool and all related implementation materials will be published on a website and will be freely accessible to the public.

STRENGTHS AND LIMITATIONS OF THIS STUDY

⇒ We will devise, select and combine relevant items into a new measurement tool to position randomised trials in the promotion to treatment continuum, which—to the best of our knowledge—does not exist.
⇒ Devised items will be organised according to a Population, Intervention, Comparison and Outcome structure, mindful of their potential to inform the promotive, preventive or treatment nature of the primary study domains.
⇒ We will apply techniques adapted from the research prioritisation methodology to consolidate and combine (and remove redundant) items.
⇒ The tool will be preliminarily validated in a large sample of randomised trials evaluating the effectiveness of promotion, prevention and treatment interventions and included in Cochrane reviews.

INTRODUCTION

Prevention of mental disorders is a rapidly growing area of research with substantial potential benefits for population health. Prevention includes universal and/or selective and/or indicated, consistent with the Institute of Medicine Framework,[1–3] and may encompass early detection, diagnosis and the reduction of the impact of disease on functionality and quality of life.[4 5] Health promotion is about empowering people to improve their health. It could be considered as a separate intervention strategy, but it may also be concerned with the spectrum of prevention modalities.[6] Several randomised controlled trials (RCTs) and systematic reviews (SRs) have been conducted to evaluate the efficacy and effectiveness of a variety of promotion and prevention interventions across different population groups.[7–12] Different in scope from treatment RCTs, these studies aimed broadly to test interventions for empowering

BMJ Open: first published as 10.1136/bmjopen-2023-082652 on 13 August 2024. Downloaded from http://bmjopen.bmj.com/ on December 27, 2024 at Koc Universitesi. Protected by copyright.

people to strengthen their mental health (promotion) and/or averting, reducing or delaying the onset of overt mental disorders in general and at-risk populations (prevention).

However, promotion, prevention and (early) treatment trials in mental health are often difficult to distinguish. First, designing and implementing true prevention studies, focused on the onset of new mental disorders rather than reducing the symptoms of existing ones, is complex and requires resources.[13 14] Ascertainment of incident mental disorders is hard to measure with certainty because baseline exclusion of prevalent cases can be challenging, and onset dates can be unclear. Large samples and constant monitoring of disease onset over time are necessary to demonstrate prevention. However, measures of the prospective worsening of psychological symptoms are commonly used as proxy outcome, which may blur evidence on the true efficacy/effectiveness of interventions for the prevention of mental disorders.[15] The incidence of mental disorders at follow-up is a better outcome measure of preventive studies. It implies the administration of sophisticated diagnostic interviews from trained clinical staff. This is costly and time-consuming in large clinical trials and in clinical practice and potentially unfeasible in low-resourced settings.[16] Furthermore, it may imply dichotomising continuous measures when the diagnosis is made according to a cut-off on a symptom scale, both at baseline (aiming to exclude those scoring above the scale threshold), and at study endpoints (aiming to calculate the proportion of participants developing a mental disorder according to the same scale threshold rule).[7 15 17–20] However, mental health is not just a matter of being 'mentally healthy' or 'mentally ill'. The dimensional approach to the diagnosis of mental disorders posits that mental health is best conceived and measured along a continuum of signs and symptoms of severity and intensity, with a range of states between the hypothetical two categories of absence and presence of disease.[1–3 16] Therefore, study designs that operationalize the concept of mental health as a binary or all-or-nothing condition may be seen as an oversimplification of a complex phenomenon. This conundrum is still far from being resolved both conceptually and practically. Existing evidence is often hard to reconcile and consolidate. Several studies have both typical features of prevention trials and others of treatment trials.[21] Many of these, including large RCTs evaluating promotive and preventive strategies, have been completed,[22] including in humanitarian settings in low-income and middle-income countries.[23] These studies contributed to advance knowledge and understanding of intervention mechanisms and practice and to inform public health decisions and action.[24]

Second, and relatedly, although many researchers have focused on identifying biomarkers of mental disorders, current scientific evidence is still insufficient on the putative underlying processes, which cannot be targeted with preventive interventions. An SR of 780 studies encompassing biochemical, genetic, neuroimaging, neurophysiological and neuropsychological measures failed to identify candidate diagnostic biomarkers for detecting or confirming the presence of neurodevelopmental disorders.[25] The lack of valid, reliable and widely available biomarkers is a reason for the use of psychological distress or early disorder-specific symptoms as a marker of risk for mental disorders.[26] And even in the case when biomarkers are identified, there remains a need to rely on clinical interviews for the diagnosis of mental and behavioural disorders.

Third, for several psychological interventions tested in RCTs the boundaries between prevention and treatment are blurred. For example, in humanitarian settings, a broad range of psychological and social interventions are implemented under the composite term 'mental health and psychosocial support' (MHPSS). MHPSS refers to 'any type of local or outside support that aims to protect or promote psychosocial well-being and/or prevent or treat mental disorders'. In this multilayered framework, interventions have been depicted by the Inter-Agency Standing Committee as a pyramid of supports.[27 28] The four tiers of the pyramid are: the interventions located at the top levels are the most specialised (eg, the fourth layer is psychotherapies, the third is focused psychosocial interventions), at the basis sit basic interventions (second layer) and general social support (first layer).[27] Nevertheless, the contents of MHPSS interventions are seldom explicitly described in experimental studies (see below). Moreover, while there are different conceptual frameworks included under the label prevention, interventions are delivered without describing a clear theory of change.[29] Another example of lack of clarity regarding the distinction between prevention and early treatment comes from the field of traumatic stress. The time elapsed since trauma plays a crucial role in determining whether an intervention qualifies as preventive or treatment.[30] The DSM-5 (Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition) stipulates that a diagnosis of post-traumatic stress disorder cannot be established during the first month following trauma.[31] Consequently, all interventions provided within the initial month following the trauma are categorised as secondary prevention, even when administered to individuals experiencing severe mental health symptoms.[26]

Intervention manuals, including administration modalities and settings, are often not reported, insufficiently described or not publicly available. Whether the intervention was conceived and developed for promotion, prevention or treatment is often not clearly stated in the intervention manuals. Inferential reasoning is difficult because of the uncertainty of both internal and external validity of studies, and the indirectness of evidence. The applicability and use of interventions to diverse populations and to their use for both prevention and treatment of mental disorders may widen. For example, the WHO did not specify whether its Self-Help Plus (SH+) intervention for stress reduction (a pivotal MHPSS) was

BMJ Open: first published as 10.1136/bmjopen-2023-082652 on 13 August 2024. Downloaded from http://bmjopen.bmj.com/ on December 27, 2024 at Koc Universitesi. Protected by copyright.

conceived for prevention, treatment of mental disorders, or both.[32] Consequently, experimental studies have been conducted to test the efficacy of SH+ in different populations for both prevention[24 33 34] and treatment of mental disorders.[35]

Fourth, the reporting of many RCTs is still suboptimal. As said, details of the interventions are often insufficient. Sampling procedures are poorly reported, and inclusion and exclusion criteria for participants unclear.[36] This makes it difficult to draw coherent lines between populations and samples at-risk and with mental disorders.[37] As mentioned above, the same intervention given to the former may be conceived of as preventive, and to the latter as therapeutic, and the design of the trial would change accordingly.

Promotion, prevention and treatment studies must be clearly discerned to facilitate a better understanding of research findings in the field of public mental health,[38] and to inform best practice, including delivery precision, resources allocation and ultimately the effectiveness of interventions. Therefore, the characterisation of mental health trials along the promotion-to-treatment continuum is crucial for effectively identifying what works for whom in different contexts, including where resources are limited, and intervention responses vary.

Next, a clear distinction can help researchers optimise the choice of outcomes, pinpoint research gaps, allocate efforts and resources where needed most, avoiding redundancies. This ensures that interventions meet populations' mental health needs.[39 40]

However, a tool to place RCTs along the promotion-to-treatment continuum is lacking. A scoping review of existing measures of the *promotiveness* to *treatment-ness* of studies in mental health interventions (mainly RCTs) returned zero eligible studies and/or reports (as of 3 October 2023, details in the online supplemental appendix 1).

Scope and objectives

Against this background, the main aim of this project is to produce a measurement tool, the VErona-LUgano Tool (VELUT) to position RCTs of MHPSS interventions along the promotion-to-treatment continuum. In this paper, we describe the methodology to develop and validate the tool.

METHODS

The methods, process and procedures described here are adapted from the practical guide for the development of health measurements scales by Streiner *et al*,[41] which we combined with the Child Health and Nutrition Research Initiative method.[42–44]

The VELUT will be outcome based and will comprise items that may include statements or questions with appropriate answer options. Items will be devised, drafted, selected and tested. In the conceptualisation step, we will establish a Tool Development Group (TDG) of international experts in public and global mental health and related disciplines, who are coauthors of the present paper. In the item devising step, the TDG will generate a first pool of items related to, and collectively reflecting, our conceptualised constructs. Next, in the item selection phase, the TDG will group, reduce and refine items. Figure 1 summarises the steps for developing the tool, which are described in detail below.

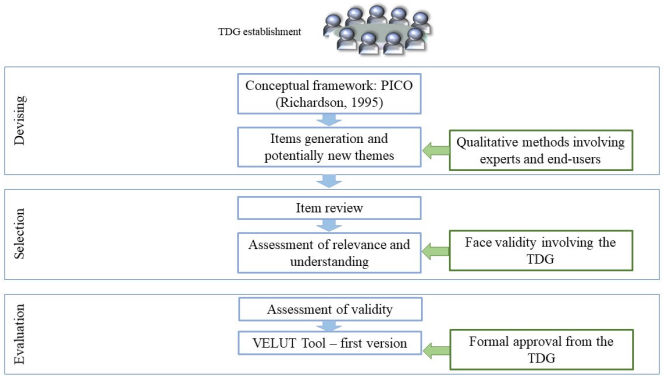


Figure 1 Flow-diagram for the development of the VELUT. PICO, Population, Intervention, Comparison and Outcome; TDG, Tool Development Group; VELUT, VErona-LUgano Tool.

Devising of items

The TDG will adopt an iterative process to define the construct(s) that we aim to assess and concur to use the Population, Intervention, Comparison and Outcome (PICO) and IOM framework for devising, reviewing and selecting items.[1 45] The size of the TDG is n=8/10, which is similar to the size of the coordinating/editing groups developing WHO mental health guidelines.[46]

We will combine two approaches for item generation. First, we will attempt to retrieve items from existing measures (if any), searching PubMed, Epistemonikos, Ovid PsycInfo and the Equator platform. Second, the TDG will integrate various sources to devise scale items, including theory, research expertise and expert opinions, and collaboration with researchers and practitioners. Based on an already performed search strategy and screening, which identified zero studies reporting on similar tools (see online supplemental appendix 1), we anticipate we will rely mostly on the latter approach. For this, we will use qualitative methods with systematic reviewers, trialists and experts (being mindful of gender, cultural, age heterogeneity). As this is a qualitative study, a sample size calculation was not performed. We plan to collect data from 10 to 12 participants. Qualitative methods include: (1) focus groups with the tool's intended end users. These focus groups will elicit general themes and insights regarding the tool's usability and relevance. Participants will be encouraged to share their perspectives, preferences and suggestions for improvement; (2) interviews until saturation of themes is reached about the key methodological features of primary studies that must be critically appraised to gauge the study's promotive, preventive and treatment nature. These

BMJ Open: first published as 10.1136/bmjopen-2023-082652 on 13 August 2024. Downloaded from http://bmjopen.bmj.com/ on December 27, 2024 at Koc Universitesi. Protected by copyright.

interviews will focus on exploring the key methodological features of primary studies that are critically appraised to evaluate the study's promotive, preventive and treatment nature. Interview topics and questions may include but are not limited to: the types of study populations, recruitment and screening of study participants, types of interventions, comparator, outcome measures and setting; (3) Collective opinions to generate a comprehensive list of items for the tool. Through facilitated discussions and consensus-building exercises, the group will prioritise and refine the identified items to ensure they capture the key dimensions of interest.

The principle of thematic saturation will guide the determination of a sufficient number of interviews. This means that data collection will continue until no new themes or insights emerge from the interviews, indicating a comprehensive understanding of the topic has been achieved.

Experts may also hint at and mention additional relevant themes and domains, which will be integrated with public health theory of prevention and treatment MHPSS. This will serve the purpose of anchoring our work to a heuristic model, from which to derive items.

Selection of items and Delphi exercise

All steps will be tracked and documented and discussed with the TDG.

First, all items devised will be stored in a preliminary 'bucket' and organised according to a standard PICO structure to preliminarily allocate/attribute items to the promotion-to-treatment continuum. Therefore, two TDG coordinators (MP and EA) in consultation with TDG members will categorise the devised items according to the PICO elements, being mindful of the potential of the item to inform the promotive, preventive or treatment nature of the primary study domains.

Second, two researchers with different professional backgrounds (MP and EA) will apply techniques adapted from the research prioritisation methodology to consolidate and combine (and remove redundant) items.[42] This means eliminating duplicates, consolidating and combining similar items and maintaining an overall balance between granularity and overall salient features of the construct.

Third, MP, EA and CB first discuss on the wording and clarity of items. TDG members will first participate in a survey to provide quantitative feedback on the items (pertinence with respect to the constructs being measured and clarity). Then, TDG members will participate in structured discussions that ensure that all TDG members interpret the items (and responses) the same way. The scope is not to validate the VELUT in users (ie, systematic reviewers) but rather to confirm the conceptual consistencies within TDG members as the items are being developed. The TDG members will discuss the relevance and understanding of items and use a Likert scale (1=irrelevant to 5=extremely suitable) to assess the face validity of each item, and of the provisional items collection. Face validity concerns whether and to what extent the items measure what they are set out to measure and their salience with respect to the construct of interest. In addition, two researchers (MP and EA) will conduct cognitive interviews in a hybrid format (face to face and digital) with the TDG members to explore their understanding of the items and of the meaning of responses. Item wording and phrasing will be improved accordingly and through a consolidation step based on iterative discussion in a dedicated session of TDG members.

The TDG will conduct an adapted Delphi exercise, setting up an online survey (in REDCap) among global mental health fellows and experts identified by the TDG members, with the following purposes:

1. Soliciting general feedback on the preliminary list of items prepared by the TDG.
2. Suggesting additional items (to complete the list of items) according to the PICO structure mentioned above and suggesting removal of redundant or non-pertinent items.
3. Asking fellows and experts to score each item on a Likert scale of the informativeness and relevance of each item to appraise the promotive, preventive or treatment nature of the study (RCT) and the face and content validity of the collection of items, and open-ended questions. Th Delphi exercise will be coordinated by the TDG, implemented in REDCap,[47] and the list will be pilot-tested in REDCap before use.
4. The TDG will use the set of information of this step to compile the final list of items that compose the new measure.

Experts, fellows and stakeholders for both the qualitative process (first phase—devising items) and the Delphi exercise (second phase—selection of items) will be identified according to a mapping exercise using standard mapping methods. This will involve the WHO Collaborating Centers of the Universities of Verona and Geneva, and the Institute of Public Health of the Università della Svizzera Italiana to which the leading authors (MP, CB and EA) belong. We aim to include experts in global mental health across our research networks based on their level of influence, interest or relevance to the project and to the demonstrated competence in the topic based on their work and publications and the research they had designed and conducted in past years. Then, experts, fellows and stakeholders will be recruited through snowballing techniques by the TDG members.

Patient and public involvement

Considering that this study collects data from already published randomised trials, it is not possible and/or appropriate to involve patients or the public in the design of this study.

BMJ Open: first published as 10.1136/bmjopen-2023-082652 on 13 August 2024. Downloaded from http://bmjopen.bmj.com/ on December 27, 2024 at Koc Universitesi. Protected by copyright.

Preliminary evaluation of the psychometric properties of the VELUT (statistical analysis)

Data collection

The aim of the anticipated preliminary formal explorations of the psychometric properties of our tool is to inform decisions about the combination of individual items into a scale and then derive a final score along the range from *promotiveness* and *preventiveness* to *treatmentness*. To attain this goal, we plan to collect data of the application of the VELUT in real-life conditions. Data collection consists in the use of the tool on published primary randomised studies of psychosocial interventions. The assessors (at least two) will be experienced systematic reviewers (ie, the tool users). The primary studies to be assessed were previously searched, selected and are stored in a repository at the University of Verona.[9 10 48 49] We plan to assess approximately 200 primary studies. This is deemed a large enough sample size for our Confirmatory Factor Analysis (CFA) and Items Response Theory (IRT) models (described below).[50 51]

We will compute an overall score for each primary study using the data obtained with the application of the VELUT. These scores will allow to quantify the *promotiveness*, *preventiveness* and *treatmentness* of the primary studies along the continuum of the score range. We will also attempt to establish cut-off points (either using distribution-based or anchor-based strategies) to classify the primary studies closer to the promotive or treatment boundaries.[52]

Analysis

With the standardised version (Guttman lambda 6) of Cronbach alpha we will seek to assess the internal consistency of the scale, the averaged between-item correlation and the signal-to-noise ratio.[53]

We will employ statistical techniques to identify redundant or overlapping items. Additionally, we plan to perform a preliminary formal psychometric evaluation to assess reliability and validity. This includes assessing internal consistency, test–retest reliability and factor structure. This process helps streamline the scale and improve its efficiency without compromising its psychometric properties.[53] In addition, we perform an alpha testing of the provisional list of items to identify unreliable items exploring ceiling and floor effects, that is items endorsed by everyone or no one, respectively.

We will use CFA, consistent with IRT, to confirm the items fit with the anticipated domain structure of the scale (ie, PICO framework). IRT models are used to relate the responses to the scale items to the underlying construct of interest (a continuum of *promotiveness, preventiveness* to *treatmentness* of interventions tested in RCTs).[54] This relationship is measured and graphically displayed using item characteristic curves (ICC) for each item, a graph of the probability of endorsing an item answer by the latent trait level, a probability that increases with the latent trait level following a cumulative logistic distribution. The interpretation of the ICC is based on the two main parameters: difficulty and discrimination.[55 56]

We anticipate the use of two-parameter logistic models (2PL).[57] In 2PL models, both difficulty and discrimination can vary across items. The latter is the slope of the ICC for each item and depicts the item ability to distinguish between neighbouring levels of the latent trait. Thus, an item with a steep ICC is expected to be endorsed (answered) differently also when the latent trait levels vary only slightly. We plan to run both standard 2PL IRT models and an alternative model specification, the BiDimensional Two-parameter Logistic model Multiple IRT (2D 2L MIRT (Multidimensional Item Response Theory)), with formal comparisons of goodness of fit based on likelihood ratio tests and additional statistics including Akaike information criteria, the Bayesian information criteria and the M2 statistics. More advanced models, including the root mean square residual, the Tucker-Lewis index[58] and the comparative fit index[59] may also be used as needed. Our analysis of the item fit will be based on a signed χ^2 statistic and the RMSEA (Root Mean Square Error of Approximation.

Next, for all items, we will calculate the item information function, which refers to the precision of the item in measuring the latent trait. Item information sums to a test information function, which is used to depict the combined coverage and precision of the scale items relative to the *promotiveness, preventiveness* or *treatmentness* latent trait.

We will run formal testing to confirm the main assumptions of IRT models are met, namely unidimensionality and local/conditional independence. For the former, this is done by contrasting the 2D 2PL model with its unidimensional specification. For the latter, we will use a local dependence statistic between pair of items based on signed χ^2 values and Cramer's V, its relative standardised version. Items that violate these assumptions will be flagged for potential exclusion, which will be discussed by the TDG. In addition, we will evaluate the number of dimensions covered and perform reliability and validity assessment. For all the RCTs from the Cochrane reviews mentioned above, inter-rater and test–retest reliability for each item will be measured by Cohen's kappa coefficient, which compensates and corrects for the proportion of agreement that might occur by chance.

The agreement indicated by the kappa coefficient can be poor (<0.21), fair (0.21–0.40), moderate (0.41–0.60), good (0.61–0.80) or very good (0.81–1.0).[60] Usually, a value of kappa>0.70 is considered adequate agreement. Because Kappa is affected by the presence of bias between observers or times, we plan to apply a test of symmetry of the off-diagonal cells. Analyses will be carried out using Stata for Windows.[61] We anticipate that further construct validation studies will follow and may be independently conducted by different research groups once the tool will be available. These studies will be crucial to provide empirical support for the construct validity of our tool, demonstrating its ability to capture the latent traits we

BMJ Open: first published as 10.1136/bmjopen-2023-082652 on 13 August 2024. Downloaded from http://bmjopen.bmj.com/ on December 27, 2024 at Koc Universitesi. Protected by copyright.

seek to measure. These studies are best conducted after the items are consolidated, the scale is implemented, and may inform further improvements of the scale (including bettering of items).

Expected results

The main outcome of the presented project is the VELUT issue in global mental health RCTs. We expect to publish the results of this project and to make the tool available for use at the beginning of 2025. The tool will come with instructions, including the importance of focusing on one outcome at a time when appraising the promotive treatment of a study. The tool and all related implementation materials will be published on a website and will be freely accessible.

We have adapted our methods and procedures from robust and consolidated psychometric theory and practice, combined with structured consensus-based decision-making processes used in global health. Our approach can be further adapted and re-used in similar exercises and to craft tools of relevance in global health.

Our tool holds the promise to provide important advantages for researchers and clinicians in the field of mental health. First, it will orient on the *promotiveness, preventiveness,* or *treatmentness* nature of trials. It may be also used during the design stage of an experimental study to assure that the study design and methods suit at best the intended goal of testing either promotion, prevention or treatment efficacy of MHPSS interventions. An immediate advantage of the new tool is also the stratification of primary studies robustly, transparently and in a replicable manner. Our tool holds aims to keep subjectivity and selection bias at bay in SRs and meta-analysis. Moreover, the classification and stratification of primary studies along the promotive to treatment continuum can factually contribute to reduce indirectness issues. It is best to use the evidence that matters and is pertinent to inform action.

Second, our tool can have important implications for the design and conduction of SRs and meta-analyses. It can be used to define and better specify inclusion and exclusion criteria of primary studies based on an objectively defined and quantified measure of *promotiveness, preventiveness* or *treatmentness* of the study. Moreover, the classification of primary studies can inform stratified and subgroup analyses, and meta-regression analyses aimed at exploring sources of clinical heterogeneity attributable to methodological differences between studies. Nonetheless, similar to the Cochrane risk of bias tool,[62] the items of our new tool may apply to each outcome separately. At the dissemination stage, our tool can contribute to improve quality of reporting standards of both primary studies and SRs and meta-analysis. For example, our tool may be used to complement the Consolidated Standards of Reporting Trials, Standard Protocol Items: Recommendations for Interventional Trials, Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) and PRISMA Protocols guidelines, respectively.[63] This will also presumably contribute to align terminology of titles, abstracts and indexing of studies with the existing international lexicon and nomenclature in public health, that is, not limited to mental health interventions.

Finally, there are potentially immediate and direct implications for implementers and decisions makers. Our tool will allow to classify and distinguish between trials based on their use: promotion, prevention or treatment, and inform decision and policy action about interventions, populations, and intended benefits accordingly.

ETHICS AND DISSEMINATION

This study involves data collection from the already published literature, there is no requirement for ethical approval. However, this protocol has been approved by the ethics committee of the Università della Svizzera Italiana (CE 2024 04). The results of the present project will be disseminated in peer-reviewed journals and at international and national scientific meetings. Training materials for the application of the tool will also be developed and disseminated to the scientific community. The tool and all related implementation materials will be published on a website and will be freely accessible to the public.

Author affiliations
[1]WHO Collaborating Centre for Research and Training in Mental Health and Service Evaluation, Department of Neuroscience, Biomedicine and Movement Sciences, Section of Psychiatry, University of Verona, Verona, Italy
[2]Institute of Public Health, Faculty of Biomedical Sciences, Università della Svizzera Italiana, Lugano, Switzerland
[3]Department of Global Health and Social Medicine, Harvard Medical School, Boston, Massachusetts, USA
[4]Johns Hopkins University Center on Aging and Health, Baltimore, Maryland, USA
[5]Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland, USA
[6]Department of Clinical, Neuro- and Developmental Psychology, VU University Amsterdam, Amsterdam, The Netherlands
[7]Department of Psychology, Koc University, Istanbul, Türkiye
[8]Lisbon Institute of Global Mental Health, NOVA University of Lisbon, Lisboa, Portugal
[9]Health Service and Population Research Department, Centre for Global Mental Health, Institute of Psychiatry, Psychology and Neuroscience, King's College, London, UK
[10]Research and Development Department, War Child, Amsterdam, the Netherlands
[11]Amsterdam Institute for Social Science Research, University of Amsterdam, Amsterdam, the Netherlands
[12]Department of Psychiatry and Mental Health, University of Cape Town, Cape Town, South Africa
[13]Global Health Section, University of Copenhagen, Kobenhavn, Denmark
[14]Athena Research Institute, Vrije Universiteit Amsterdam, Amsterdam, the Netherlands

X Alden Gross @alden.gross

Contributors Conceptualisation: MP, EA and CB. Methodology: MP, AMA, EA and CB. Critical conceptual discussion and decisions on the main methodological steps: MP, EA, CB, PC, MJDJ, CL, DP and WAT. Planning and being responsible for data analysis: EA, AG, FT, AMA, MP and CB. Being active part of the Tool Development Group: MP, EA, CB, CA, AG, PC, MJDJ, MSilva, AMA, MSijbrandij, CL, DP, FT, EP and WAT. Writing, critically reviewing, editing: All authors. MP and CB are the guarantors and responsible for this work, its conduction and will have full access to the data.

Funding This research is partially funded by a grant for research internationalisation of the University of Verona (Bando PIA-2023). DP was funded by the European Union's Horizon-MSCA-2021-PF01 research program under grant

BMJ Open: first published as 10.1136/bmjopen-2023-082652 on 13 August 2024. Downloaded from http://bmjopen.bmj.com/ on December 27, 2024 at Koc Universitesi. Protected by copyright.

agreement N 101061648. The funder had no role in the design and conduct of the study; collection, management, analysis and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

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.

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
Marianna Purgato http://orcid.org/0000-0002-3783-8195
Eleonora Prina http://orcid.org/0000-0002-4005-1967
Alden Gross http://orcid.org/0000-0002-4678-0562
Pim Cuijpers http://orcid.org/0000-0001-5497-2743

REFERENCES

1 Institute of Medicine (US) Committee on Prevention of Mental Disorders. In: Mrazek PJ, Haggerty RJ, eds. *Reducing risks for mental disorders: frontiers for preventive intervention research*. National Academies Press (US), 1994.
2 National Research Council (US), Institute of Medicine Committee (US) on the Prevention of Mental Disorders and Substance Abuse Among Children, Youth, and Young Adults: Research Advances and Promising Interventions; The National Academies Collection: Reports funded by National Institutes of Health. In: O'Connell ME, Boat T, Warner KE, eds. *Preventing mental, emotional, and behavioral disorders among young people: progress and possibilities*. Washington(DC): National Academies Press (US) Copyright © 2009, National Academy of Sciences, 2009.
3 National Academies of Sciences, Engineering, and Medicine; Division of Behavioral and Social Sciences and Education; Board on Children, Youth and Families et al. *Fostering healthy mental, emotional, and behavioral development in children and youth: a national agenda*. Washington (DC): National Academies Press (US) Copyright © 2019, National Academy of Sciences, 2019.
4 Tol WA, Purgato M, Bass JK, *et al*. Mental health and psychosocial support in humanitarian settings: a public mental health perspective. *Epidemiol Psychiatr Sci* 2015;24:484–94.
5 Eaton WW. *Public mental health*. Oxford University Press, 2019.
6 Naidoo J, Jane W. *Foundations for health promotion*. 4th edn. 2016.
7 Buntrock C, Berking M, Smit F, *et al*. Preventing depression in adults with subthreshold depression: health-economic evaluation alongside a pragmatic randomized controlled trial of a web-based intervention. *J Med Internet Res* 2017;19:e5.
8 van Zoonen K, Buntrock C, Ebert DD, *et al*. Preventing the onset of major depressive disorder: a meta-analytic review of psychological interventions. *Int J Epidemiol* 2014;43:318–29.
9 Papola D, Purgato M, Gastaldon C, *et al*. Psychological and social interventions for the prevention of mental disorders in people living in low- and middle-income countries affected by humanitarian crises. *Cochrane Database Syst Rev* 2020;9:CD012417.
10 Purgato M, Prina E, Ceccarelli C, *et al*. Primary-level and community worker interventions for the prevention of mental disorders and the promotion of well-being in low- and middle-income countries. *Cochrane Database Syst Rev* 2023;10:CD014722.
11 Please provide ref 11.
12 Buntrock C, Ebert DD, Lehr D, *et al*. Effect of a web-based guided self-help intervention for prevention of major depression in adults with subthreshold depressionn: a randomized clinical trial. *JAMA* 2016;315:1854.
13 Cuijpers P, van Straten A. Preventing the onset of mental disorders: an emerging field. *Clin Psychol Rev* 2021;90:102097.
14 Cuijpers P, Smit F, Furukawa TA. Most at-risk individuals will not develop a mental disorder: the limited predictive strength of risk factors. World J Psychiatry. 2021;20:224–5.
15 Cuijpers P. Why primary prevention often is no prevention at all. *Eur Neuropsychopharmacol* 2022;58:1–3.
16 Patel V, Saxena S, Lund C, *et al*. The lancet commission on global mental health and sustainable development. *Lancet* 2018;392:1553–98.
17 Cuijpers P, van Straten A, Smit F, *et al*. Preventing the onset of depressive disorders: a meta-analytic review of psychological interventions. Am J Psychiatry. 2008;165:1272–80.
18 Cuijpers P. Examining the effects of prevention programs on the incidence of new cases of mental disorders: the lack of statistical power. *Am J Psychiatry* 2003;160:1385–91.
19 Cuijpers P. Prevention: an achievable goal in personalized medicine. *Dialogues Clin Neurosci* 2009;11:447–54.
20 Cuijpers P, Beekman ATF. [Personalised prevention of mental disorder]. *Tijdschr Psychiatr* 2018;60:189–93.
21 Mendelson T, Eaton WW. Recent advances in the prevention of mental disorders. *Soc Psychiatry Psychiatr Epidemiol* 2018;53:325–39.
22 Cuijpers P, Pineda BS, Quero S, *et al*. Psychological interventions to prevent the onset of depressive disorders: a meta-analysis of randomized controlled trials. *Clin Psychol Rev* 2021;83:101955.
23 Cuijpers P. Preventing the onset of depressive disorders in low- and middle-income countries: an overview. *Glob ment health* 2023;10:e28.
24 Acarturk C, Uygun E, Ilkkursun Z, *et al*. Effectiveness of a WHO self-help psychological intervention for preventing mental disorders among Syrian refugees in Turkey: a randomized controlled trial. *World Psychiatry* 2022;21:88–95.
25 Cortese S, Solmi M, Michelini G, *et al*. Candidate diagnostic biomarkers for neurodevelopmental disorders in children and adolescents: a systematic review. *World Psychiatry* 2023;22:129–49.
26 Bisson JI, Wright LA, Jones KA, *et al*. Preventing the onset of post traumatic stress disorder. *Clin Psychol Rev* 2021;86:102004.
27 IASC. IASC reference group for mental health and psychosocial support in emergency settings. In: *IASC guidelines on the mental health and psychosocial support in emergency settings*. 2007.
28 Tol WA. Stemming the tide: promoting mental health and preventing mental disorders in low- and middle-income countries. *Glob ment health* 2015;2:e11.
29 Miller KE, Jordans MJD, Tol WA, *et al*. A call for greater conceptual clarity in the field of mental health and psychosocial support in humanitarian settings. *Epidemiol Psychiatr Sci* 2021;30:e5.
30 Astill Wright L, Horstmann L, Holmes EA, *et al*. Consolidation/reconsolidation therapies for the prevention and treatment of PTSD and re-experiencing: a systematic review and meta-analysis. *Transl Psychiatry* 2021;11:453.
31 American Psychiatric Association. *The diagnostic and statistical manual of mental disorders (DSM-5)*. 5th edn. Washington DC: American Psychiatric Association, 2013.
32 Epping-Jordan JE, Harris R, Brown FL, *et al*. Self-help plus (SH+): a new WHO stress management package. *World Psychiatry* 2016;15:295–6.
33 Purgato M, Carswell K, Acarturk C, *et al*. Effectiveness and cost-effectiveness of self-help plus (SH+) for preventing mental disorders in refugees and asylum seekers in Europe and Turkey: study protocols for two randomised controlled trials. *BMJ Open* 2019;9:e030259.
34 Purgato M, Carswell K, Tedeschi F, *et al*. Effectiveness of self-help plus in preventing mental disorders in refugees and asylum seekers in Western Europe: a multinational randomized controlled trial. *Psychother Psychosom* 2021;90:403–14.
35 Tol WA, Leku MR, Lakin DP, *et al*. Guided self-help to reduce psychological distress in South Sudanese female refugees in Uganda: a cluster randomised trial. *Lancet Glob Health* 2020;8:e254–63.
36 Miller KE, Rasmussen A, Jordans MJD. Strategies to improve the quality and usefulness of mental health trials in humanitarian settings. *Lancet Psychiatry* 2023;10:974–80.
37 Cuijpers P, Cristea IA. How to prove that your therapy is effective, even when it is not: a guideline. *Epidemiol Psychiatr Sci* 2016;25:428–35.

BMJ Open: first published as 10.1136/bmjopen-2023-082652 on 13 August 2024. Downloaded from http://bmjopen.bmj.com/ on December 27, 2024 at Koc Universitesi. Protected by copyright.

38 Francis-Oliviero F, Cambon L, Wittwer J, *et al*. Theoretical and practical challenges of proportionate universalism: a review. *Rev Panam Salud Publica* 2020;44:e110.
39 Purgato M, Singh R, Acarturk C, *et al*. Moving beyond a 'one-size-fits-all' rationale in global mental health: prospects of a precision psychology paradigm. *Epidemiol Psychiatr Sci* 2021;30:e63.
40 Cuijpers P, Ebert DD, Acarturk C, *et al*. Personalized psychotherapy for adult depression: a meta-analytic review. *Behav Ther* 2016;47:966–80.
41 Streiner DL, Norman GR, Cairney J. *Health measurement scales: a practical guide to their development and use*. 5th edn. Oxford University Press, 2015.
42 Shah H, Albanese E, Duggan C, *et al*. Research priorities to reduce the global burden of dementia by 2025. *Lancet Neurol* 2016;15:1285–94.
43 Rudan I. Setting health research priorities using the CHNRI method: IV. Key conceptual advances. *J Glob Health* 2016;6:010501.
44 Rudan I, Yoshida S, Chan KY, *et al*. Setting health research priorities using the CHNRI method: VII. A review of the first 50 applications of the CHNRI method. *J Glob Health* 2017;7:011004.
45 Richardson WS, Wilson MC, Nishikawa J, *et al*. The well-built clinical question: a key to evidence-based decisions. *ACP J Club* 1995;123:A12–3.
46 World Health Organization. WHO mhGAP guideline update. 2015. Available: https://www.who.int/publications/i/item/9789241549417
47 Harris PA, Taylor R, Thielke R, *et al*. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. *J Biomed Inform* 2009;42:377–81.
48 van Ginneken N, Chin WY, Lim YC, *et al*. Primary-level worker interventions for the care of people living with mental disorders and distress in low- and middle-income countries. *Cochrane Database Syst Rev* 2021;8:CD009149.
49 Purgato M, Gastaldon C, Papola D, *et al*. Psychological therapies for the treatment of mental disorders in low- and middle-income countries affected by humanitarian crises. *Cochrane Database Syst Rev* 2018;7:CD011849.
50 MacCallum RC, Browne MW, Cai L. Testing differences between nested covariance structure models: power analysis and null hypotheses. *Psychol Methods* 2006;11:19–35.
51 Packer T, Kephart G, Audulv Å, *et al*. Protocol for development, calibration and validation of the patient-reported inventory of self-management of chronic conditions (PRISM-CC). *BMJ Open* 2020;10:e036776.
52 Purgato M, Barbui C. Dichotomizing rating scale scores in psychiatry: a bad idea? *Epidemiol Psychiatr Sci* 2013;22:17–9.
53 Albanese E, Bütikofer L, Armijo-Olivo S, *et al*. Construct validity of the physiotherapy evidence database (PEDro) quality scale for randomized trials: item response theory and factor analyses. *Res Synth Methods* 2020;11:227–36.
54 Orlando M, Sherbourne CD, Thissen D. Summed-score linking using item response theory: application to depression measurement. *Psychol Assess* 2000;12:354–9.
55 van der Linden WJ. *Handbook of item response theory: three volume set*. CRC Press, 2018.
56 Gross AL, Sherva R, Mukherjee S, *et al*. Calibrating longitudinal cognition in Alzheimer's disease across diverse test batteries and datasets. *Neuroepidemiology* 2014;43:194–205.
57 Béguin AA, Glas CAW. MCMC estimation and some model-fit analysis of multidimensional IRT models. *Psychometrika* 2001;66:541–61.
58 Cai L, Chung SW, Lee T. Incremental model fit assessment in the case of categorical data: Tucker-Lewis index for item response theory modeling. *Prev Sci* 2023;24:455–66.
59 Bentler PM. Comparative fit indexes in structural models. *Psychol Bull* 1990;107:238–46.
60 Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977;33:159–74.
61 StataCorp. *Stata statistical software: release 17*. College Station, TX: StataCorp LLC, 2021.
62 Sterne JAC, Savović J, Page MJ, *et al*. RoB 2: a revised tool for assessing risk of bias in randomised trials. *BMJ* 2019;366:l4898.
63 Pandis N, Fedorowicz Z. The international EQUATOR network: enhancing the quality and transparency of health care research. *J Appl Oral Sci* 2011;19.

BMJ Open: first published as 10.1136/bmjopen-2023-082652 on 13 August 2024. Downloaded from http://bmjopen.bmj.com/ on December 27, 2024 at Koc Universitesi. Protected by copyright.

Search strategies

Date: 03-10-2023

Pubmed: 2885 records

((prevention[Title/Abstract] OR treatment[Title/Abstract] OR intervention[Title/Abstract]) AND (tool[Title/Abstract] OR rating scale[Title/Abstract] OR scale[Title/Abstract] OR strategy[Title/Abstract] OR check-list [Title/Abstract] OR checklist [Title/Abstract] OR list [Title/Abstract] OR distinguish[Title/Abstract]) AND (randomized controlled trial[Title/Abstract] OR controlled trial[Title/Abstract] OR clinical trial[Title/Abstract]) AND (global mental health[Title/Abstract] OR public mental health[Title/Abstract] OR mental health[Title/Abstract] OR psychosocial support[Title/Abstract] OR psychological intervention[Title/Abstract]))

Epistemonikos: 76 records

[Title/Abstract field]

(((prevention OR treatment OR intervention) AND (tool OR rating scale OR scale OR strategy OR check-list OR checklist OR list OR distinguish)) AND (randomized controlled trial OR controlled trial OR clinical trial)) AND (global mental health OR public mental health OR mental health OR psychosocial support OR psychological intervention)

Ovid PsycInfo: 1158 records

((prevention or treatment or intervention) and (tool or rating scale or scale or check-list or checklist or list or strategy or distinguish) and (randomized controlled trial or controlled trial or clinical trial) and (global mental health or public mental health or mental health or psychosocial support or psychological intervention)).ti,ab.

Additional sources: Equator list of reporting guidelines (n= 579), available at https://www.equator-network.org/