Tool to assess risk of bias in studies estimating the prevalence of mental health disorders (RoB-PrevMH)

Thomy Tonia,1 Diana Buitrago-Garcia,1,2 Natalie Luise Peter,3 Cristina Mesa-Vieira,1 Tianjing Li,4 Toshi A Furukawa,5 Andrea Cipriani,6,7,8 Stefan Leucht,9 Nicola Low,1,2 Georgia Salanti6

ABSTRACT
Objective There is no standard tool for assessing risk of bias (RoB) in prevalence studies. For the purposes of a living systematic review during the COVID-19 pandemic, we developed a tool to evaluate RoB in studies measuring the prevalence of mental health disorders (RoB-PrevMH) and tested inter-rater reliability.

Methods We decided on items and signalling questions to include in RoB-PrevMH through iterative discussions. We tested the reliability of assessments by different users with two sets of prevalence studies. The first set included a random sample of 50 studies from our living systematic review. The second set included 33 studies from a systematic review of the prevalence of post-traumatic stress disorders, major depression and generalised anxiety disorder. We assessed the inter-rater agreement by calculating the proportion of agreement and Kappa statistic for each item.

Results RoB-PrevMH consists of three items that address selection bias and information bias. Introductory and signalling questions guide the application of the tool to the review question. The inter-rater agreement for the three items was 83%, 90% and 93%. The weighted kappa scores were 0.63 (95% CI 0.54 to 0.73), 0.71 (95% CI 0.67 to 0.85) and 0.32 (95% CI −0.04 to 0.63), respectively.

Conclusions RoB-PrevMH is a brief, user-friendly and adaptable tool for assessing RoB in studies on prevalence of mental health disorders. Initial results for inter-rater agreement were fair to substantial. The tool’s validity, reliability and applicability should be assessed in future projects.

BACKGROUND
Studies of prevalence provide essential information for estimating the burden of mental health conditions, which can inform research and policymaking.1 The pandemic of COVID-19, a disease first described in 2020,2 rapidly generated a large volume of literature, about studies on the prevalence of a wide range of conditions, including those related to mental health. Increased levels of anxiety, depression, psychological distress, as well as an increase in violent behaviour, alcohol and substance use, among others have been described in association with fear of infection and the effects of contamination measures.3 4 Temporary relief from obligations at school or work, or the need to commute, on the other hand, might alleviate stress for some populations.1

A systematic review provides a structured way to gather, assess and synthesise evidence from prevalence studies. One essential step in performing a systematic review is the assessment of risk of bias (RoB) of the included studies5 because the potential biases affect how certain we are about the included evidence and its interpretation.6 There is no agreement on how to assess RoB in prevalence studies,7 despite a 10-fold increase in systematic reviews of prevalence studies in the last decade.7 8 Substantial variability exists in how RoB in prevalence studies have been assessed with more than 30 tools identified and several judged to be inappropriate.9 Notably, some questions/items in existing tools focus on the quality of reporting which makes not possible to assess the biases present in prevalence studies.

To overcome the shortcomings of previous tools, such as distinguishing between RoB and quality of reporting and being adaptable to different questions, the purpose of this paper is to present a RoB tool developed to evaluate RoB in studies measuring the prevalence of mental health disorders (RoB-PrevMH). We describe the steps for developing this tool, its items, and the results of inter-rater agreement obtained by applying the tool to two sets of prevalence studies on mental health disorders.

METHODS
RoB-PrevMH was developed within the MHCOVID project (https://mhcovid.ispm.unibe.ch/), a living systematic review assessing the effect of the COVID-19 pandemic and the containment measures on mental health of the population.1 4 10 MHCOVID involves many volunteers recruited through crowdsourcing to help with data extraction and RoB assessment of a large volume of literature (referred to as the MHCOVID Crowd). We prioritised brevity and ease of application in developing the tool, owning to the different backgrounds and levels of experience and expertise of MHCOVID Crowd members in the assessment of RoB.

Development of the tool
We searched Medline and Embase (Ovid) from inception to September 2020 to identify published tools or checklists designed to assess the quality, RoB, and quality of reporting in prevalence studies (online supplemental appendix 1). In addition, we searched the Equator network website (https://www.equator-network.org/). The online supplemental appendix 2 contains a list of tools and checklists identified for the quality and reporting of prevalence studies.
www.equator-network.org/) and a database of systematic reviews of prevalence studies. One researcher (DBG) screened the search results to identify relevant tools that assessed RoB in prevalence studies.

We extracted the items from each tool selected for inclusion and grouped them under the domains of selection bias and information bias. For selection bias, items from the existing tools were separated into those referring to population representativeness or to ‘the proportion of respondents’. For information bias, items from the existing tools were separated into those referring to observer bias, recall bias or misclassification bias. Items not related to the named biases were tagged as ‘other bias’ or ‘reporting’.

Five researchers (DBG, NL, NLP, GS and TT) individually went through the list of questions in each included tool, excluded duplicated questions, and marked those that were most relevant for prevalence studies for mental health disorders. They then discussed their assessments and reached consensus prior to drafting the first version of the tool and the signalling questions. Figure 1 illustrates the process of developing RoB-PrevMH.

**Testing and finalisation of the tool**

Four members of the team (SL, NLP, GS and TT) pilot tested the first version of the tool and drafted a guidance document. Subsequently, these four researchers and four volunteers from the MHCOVID Crowd (who were not involved in the development of the tool) further tested the tool in a total of eight studies. Based on feedback from this exercise, the guidance document was updated accordingly, including examples and practical advice.

**Inter-rater reliability**

We tested the reliability of assessments by different users of RoB-PrevMH with two sets of prevalence studies. The first set included 50 prevalence studies (two sets of 25) randomly selected from those identified as potentially relevant for the MHCOVID project during the abstract screening stage. Two pairs of researchers independently applied RoB-PrevMH (team A, 25 studies: CMV and TT; team B 25 studies: DBG and NLP). The second set included 33 studies from a systematic review of the prevalence of post-traumatic stress disorders, major depressive disorder and generalised anxiety disorder in migrants with premigration exposure to armed conflict. By using this second set of studies, we examined how RoB-PrevMH performed in a research question that was different from the one it was originally developed for. Two researchers (team C: DBG and CMV) independently applied RoB-PrevMH in this set of studies.

To assess reproducibility, we calculated the unweighted and weighted kappa statistic (with 95% CI). For weighted kappa, the observed and expected proportions of agreement are modified to measure the agreement among the ordered levels of bias (low, unclear, high) by assigning a weight of 0 to complete disagreement (rating low vs high RoB), 1 to perfect agreement and 0.5 for partial disagreement (ratings low vs unclear or high vs unclear). We also calculated the percentage of agreement between raters (number of agreements/number of assessments x 100). The analysis was conducted in STATA V.15.1. We followed the interpretation of the kappa statistic proposed by Landis and Koch (1977) and described in the STATA manual where the values below the cut points 0.00, 0.20, 0.40, 0.60, 0.80 and 1.00 approximately define poor, slight, fair, moderate substantial and almost perfect agreement.

**RESULTS**

**Description of RoB-PrevMH tool**

We identified 10 tools that assess RoB in prevalence studies, summarised in table 1. Following the process mentioned above, we developed RoB-PrevMH which consists of one introductory question and three items (table 2). It also includes signalling questions aimed to help the user reach a judgement; after completing our study we improved and refined the questions associated with two items and these are presented in table 2 alongside the original questions. The elaboration and guidance document is presented in online supplemental appendix 2. RoB
for each item can be judged as ‘high’, ‘low’ or ‘unclear’. We instructed users to avoid judging any of the questions as unclear, whenever possible. This recommendation was based on the guidelines to assess the risk bias for Systematic Reviews on Interventions, which states that ‘unclear’ should be only used when the information about the domain is truly unknown.23 The tool does not allow a summary RoB assessment because some aspects defined? By ‘target population’ we refer to the entire population for which we are interested to draw inference. In the first set of studies from the MHCOVID project, the target population of the systematic review was defined as ‘the general population’ or any age or gender-based subgroups of the general population (eg, children only, or men only, or elderly, see online supplemental appendix 2). In the second set of studies, the target population of the systematic review was migrants exposed to armed conflict.25

The introductory question is ‘Was the target population clearly defined?’ By ‘target population’ we refer to the entire population for which we are interested to draw inference. In the first set of studies from the MHCOVID project, the target population of the systematic review was defined as ‘the general population’ or any age or gender-based subgroups of the general population (eg, children only, or men only, or elderly, see online supplemental appendix 2). In the second set of studies, the target population of the systematic review was migrants exposed to armed conflict.25

<table>
<thead>
<tr>
<th>Table 1</th>
<th>RoB tools considered for developing RoB-PrevMH</th>
</tr>
</thead>
<tbody>
<tr>
<td>ID</td>
<td>Tool description</td>
</tr>
<tr>
<td>1</td>
<td>A tool designed to assess the quality of prevalence studies on low back pain.</td>
</tr>
<tr>
<td>2</td>
<td>A critical appraisal tool designed to assess the methodological strengths, results and relevance of articles on prevalence of a health problem.</td>
</tr>
<tr>
<td>3</td>
<td>A guideline to critically appraise prevalence studies on psychiatric disorders, both in the general population and in specific clinical settings.</td>
</tr>
<tr>
<td>4</td>
<td>A tool to assess the usefulness of prevalence studies in the context of surveillance activities.</td>
</tr>
<tr>
<td>5</td>
<td>A tool for evaluating the quality of studies that examine the prevalence of chronic conditions or risk factors.</td>
</tr>
<tr>
<td>6</td>
<td>A risk of bias tool for prevalence studies based on Leboeuf-Yde and Lauritzen1995.</td>
</tr>
<tr>
<td>7</td>
<td>An instrument for the qualitative assessment of the methodology of prevalence studies.</td>
</tr>
<tr>
<td>8</td>
<td>A critical appraisal tool for assessing studies included in systematic reviews of prevalence.</td>
</tr>
<tr>
<td>9</td>
<td>A tool to assess the methodological quality of a prevalence study and the possibility of bias.</td>
</tr>
<tr>
<td>10</td>
<td>A tool for assessing the risk of bias in prevalence studies of exposure to occupational risk factors.</td>
</tr>
</tbody>
</table>

for which we are interested to draw inference. In the first set of studies from the MHCOVID project, the target population of the systematic review was defined as ‘the general population’ or any age or gender-based subgroup of the general population (eg, children only, or men only, or elderly, see online supplemental appendix 2). In the second set of studies, the target population of the systematic review was migrants exposed to armed conflict.25

This introductory question had two response options; ‘yes’ or ‘no’ and has implications for the evaluation of the first RoB item: if the answer is ‘no’, the first item of the tool is automatically assigned an ‘unclear’ risk.

<table>
<thead>
<tr>
<th>Table 2</th>
<th>Items included in RoB-PrevMH, suggested rephrasing and guidance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Item</td>
<td>Question</td>
</tr>
<tr>
<td>1</td>
<td>Representativeness of the sampling frame</td>
</tr>
<tr>
<td>2</td>
<td>Representativeness of the responders</td>
</tr>
<tr>
<td>3</td>
<td>Measurement of the condition</td>
</tr>
</tbody>
</table>

Each of the three items can be given a judgement high, low or unclear.

*This was a requirement for inclusion in the MHCOVID study and hence not included in the current version of the tool.*
Item 1 selection bias: representativeness of the sampling frame
The first RoB item is related to the representativeness of the sample invited with respect to the target population by asking ‘Was the sample invited to participate in the study a true or close representation of the target population?’ The signalling question for this item asked about the method for recruitment of participants and, based on the response, the instructions guided the user to reach the corresponding RoB judgement (eg, low risk when the total or a randomly selected sample of the target population was invited; high risk for open calls for participation online or quota sampling; and unclear risk when the method to invite participants and the specific context of the sampling was not specified or when the target population was not defined; for more details, see the instructions in online supplemental appendix 2.

Item 2 selection bias: representativeness of the respondents
The second item requires a judgement as to whether those who declined the invitation, in relation to those who participated in the study, would introduce bias in the prevalence estimate, ‘How would you rate the risk of non-response bias?’ The reasons for non-participation are instrumental in forming a judgement about RoB. However, these are rarely reported, if ever. We assumed that in our context the decision not to participate is associated, directly or indirectly with the mental health of the persons invited to the study. The signalling question for this item therefore inquires only about the participants providing data as a proportion of the number of people invited to participate. RoB judgement is based on the response.

Item 3 information bias: measurement of the condition
The third item assesses the likelihood of misclassification due to the methods used to measure the target condition, ‘How do you judge the risk of information bias?’ We provided guidance for judging this question for the MHCovid project (online supplemental appendix 2); for instance, if the tool/method used to measure the condition was not applied properly across time points or across groups of participants, the risk of bias for this item was judged as high.

Inter-rater agreement
Table 3 shows the results of the inter-rater agreement for each item of RoB-PrevMH, including both weighted and unweighted kappa for the 83 included studies. For item 1, the inter-rater agreement was substantial (weighted kappa 0.63, 95% CI 0.54 to 0.73) and overall agreement 83%. For item 2, the agreement was substantial (weighted kappa 0.71, 95% CI 0.67 to 0.83) and overall agreement 90%. For item 3, the weighted kappa was 0.32 (95% CI −0.04 to 0.63; overall agreement 93%), classifying inter-rater agreement as fair.

There was a total of 45 disagreements out of 249 paired assessments among 83 studies. Most of the disagreements (n = 35) were between ‘unclear’ and either ‘high’ or ‘low’. Ten disagreements were between ‘high’ versus ‘low’ assessments.

Table 3 Results of inter-rater agreement testing

<table>
<thead>
<tr>
<th>Item</th>
<th>Unweighted kappa (95% CI)</th>
<th>% agreement</th>
<th>Weighted kappa (95% CI)</th>
<th>% agreement</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Representativeness of the sampling frame</td>
<td>0.60 (0.48 to 0.68)</td>
<td>74.7</td>
<td>0.63 (0.54 to 0.73)</td>
<td>83.1</td>
</tr>
<tr>
<td>2. Representativeness of the respondents</td>
<td>0.69 (0.59 to 0.70)</td>
<td>81.9</td>
<td>0.71 (0.67 to 0.85)</td>
<td>90.3</td>
</tr>
<tr>
<td>3. Measurement of the condition</td>
<td>0.28 (0.10 to 0.73)</td>
<td>89.2</td>
<td>0.32 (−0.04 to 0.63)</td>
<td>93.4</td>
</tr>
</tbody>
</table>

DISCUSSION
Summary of findings
We developed RoB-PrevMH, a concise RoB tool for prevalence studies in mental health that was designed with the intention to be adaptable to different systematic reviews and consisting of three items: representativeness of the sample, non-response bias and information bias. Our tool showed fair to substantial inter-rater reliability when applied to studies included in two systematic reviews of prevalence studies. All three items from RoB-PrevMH have been considered or included in existing tools.14 18 21 RoB-PrevMH does not contain any item on reporting and does not require an assessment of the overall RoB in a study. For each item, three assessments of RoB are possible (high, unclear and low).

Strengths and limitations
The strengths of RoB-PrevMH include the fact that it was created after a comprehensive review of items identified in previous tools as well as a consensus between researchers. Second, the feedback we received from the MH-COVID Crowd who used the tool suggests that the tool is concise and easy to use. Third, it focuses on RoB only and avoids questions that assess reporting. Fourth, the tool was tested by three pairs of extractors in two sets of studies with different aims. The inter-rater reliability was rated from fair to substantial. Finally, the tool has the potential to be tailored to other research questions.

Our tool also has limitations. First, the team of methodologists and investigators involved in development and testing was small. The tool would have benefited by a wider consultation strategy that involved more mental health experts and investigators who have designed and undertaken prevalence studies, as well as more methodologists. Second, the brevity of the tool could also be considered a limitation. For example, the MHCovid project only includes studies that used validated tools for measuring mental health outcomes, so we did not include specific items for recall bias and observer bias, which might be important for other questions. Third, even if we assume that RoB-PrevMH would likely be quicker to complete than other tools, we did not formally assess the time required for completion in comparison with other tools. Fourth, the need to tailor the tool for each project and create training material for the people who will apply it might require more time than other tools at the start of a project. Moreover, the inter-rater reliability varied between the three items, with kappa values ranging from 0.32 to 0.71.

An important part of the evaluation of any RoB tool is the assessment of its validity. This is often done indirectly, by contrasting findings from studies judged at low versus high RoB in each domain. For example, randomised trials at high RoB from poor allocation concealment show, on average, larger effects than studies with low RoB.22 Prevalence studies are characterised by large heterogeneity, and it is expected that some of this heterogeneity might be associated with differences in RoB.23 However, RoB-PrevMH was not found to be associated with different study findings in a meta-analysis of the changes of symptoms of depression, anxiety and psychological distress during the...
pandemic, possibly because other design and population-related factors played a more important role in heterogeneity. A large-scale evaluation of the validity of RoB-PrevMH is needed to understand which design and analysis features impact most on the estimation of prevalence.

When we compare our tool’s performance with the available instruments, only the tool proposed by Hoy et al tested the inter-rater agreement and calculated the kappa value with a considerable number of studies on the prevalence of low back and neck pain. Even though representativeness of the target population might be difficult to judge objectively, the inter-rater agreement for this item was substantial while in the 5+4 studies assessed by Hoy et al the inter-rater agreement achieved was higher. For the second item on non-response, inter-rater agreement was substantial, but lower than similar items in the Hoy tool. The third item on misclassification had the lowest kappa statistic but the highest agreement between raters. In classification tables with great imbalance in the marginal probabilities and a high underlying correct classification rate kappa can be paradoxically low, as was the case of kappa for information bias. We did not make an overall RoB assessment for each study, which the Hoy tool does because of the problems with this approach.

**Application of RoB-PrevMH in future projects**

The design of prevalence studies differs substantially depending on the question they intend to answer; as a result, having a universal tool for all types of prevalence studies, like we have for RCTs and some observational studies, might not be realistic; instead, we need tools that can be tailored to specific research questions.

Future projects applying RoB-PrevMH might need to improve the questions, and provide a more complete list of signalling questions and considerations to choose from, depending on the context and the nature of the measured prevalence. RoB-PrevMH was conceptualised and developed for the MHCOVID project, which required the use of a validated assessment tool. Additional questions about information bias might be needed for projects in which there are no validated diagnostic tools for a condition (eg, cognitive deficits in post-COVID-19) or the project does not impose inclusion criteria. Another example comes from the MHCOVID project itself. In this project we decided to rate RoB for the second and third item at every follow-up time point instead of following the original instructions to give one global rating for each study. Other projects might consider the idea of not having an arbitrary threshold for the proportion of respondents and instead extract the reported proportion and analyse the data by conducting prespecified subgroup analyses, based on this continuum of response rate with meta-regression. Moreover, our chosen arbitrary threshold for response rate might be inappropriate for other studies, as we included studies on the general population, during a pandemic and mostly done online; in other settings a ‘good’ response rate might be higher than 70%.

Evaluating the risk of information bias in prevalence studies of mental health problems requires special attention. The most reliable way to measure the presence of a condition is a diagnostic interview with a trained mental health professional; yet most studies use self-administered screening tools. These are questionnaires aiming to measure symptoms of the condition and the resulting score is used to infer about the presence or not of the condition. However, this has been shown to overestimate the true prevalence. Consequently, care is needed in the interpretation of the prevalence estimated from such studies: the meta-analysis summary result cannot be interpreted as true prevalence of the condition, but rather as the prevalence of symptoms scores above the studied threshold. Training for the tool should be tailored to a specific project and include relevant examples. For instance, for the MHCOVID project, we developed an educational video and provided online training for the volunteers of the project who extracted data from included studies and conducted RoB assessment (https://mh covid.ispm.unibe.ch/crowd.html).

Assessment of RoB in prevalence studies applies to any condition. The tools that have been published were mostly developed for specific situations, ranging from low back pain to exposure to occupational risk factors. The methods that we used to develop RoB-PrevMH follow recommended methods for the development of guidelines and should be used to further develop an RoB tool that can be applied to any systematic review question that aims to summarise the prevalence of a condition or risk factor. The MHCOVID project has provided the basis for building a network or experts with experience of RoB assessment and critical appraisal of prevalence studies to develop a generic framework for tools to assess RoB in prevalence studies.

**CONCLUSION**

RoB-PrevMH is a brief and adaptable tool for assessing RoB in studies on PrevMH disorders. Initial results for inter-rater agreement were fair to substantial. The validity, reliability and applicability of RoB-PrevMH should be further assessed in future projects.

**Author affiliations**

1Institute of Social and Preventive Medicine, University of Bern, Switzerland
2Graduate School of Health Sciences, University of Bern, Bern, Switzerland
3Department of Psychiatry and Psychotherapy, Klinikum rechts der Isar, School of Medicine and Health, Technical University of Munich, Munich, Germany
4Department of Ophthalmology, School of Medicine, University of Colorado Anschutz Medical Campus, Aurora, Colorado, USA
5Department of Health Promotion and Human Behavior, Kyoto University Graduate School of Medicine / School of Public Health, Kyoto, Japan
6Department of Psychiatry, University of Oxford, Oxford, UK
7Oxford Precision Psychiatry Lab, NIHR Oxford Health Biomedical Research Centre, Oxford, UK
8Oxford Health NHS Foundation Trust, Warneford Hospital, Oxford, UK
9Department of Psychiatry and Psychotherapy, Klinikum rechts der Isar, School of Medicine, Technical University of Munich, Freising, Germany
10NIHR Oxford Cognitive Health Clinical Research Facility, by an NIHR Research Professorship (grant RP-2017-08-ST2-006), by the NIHR Oxford and Thames Valley Applied Research Collaboration and by the NIHR Oxford Health Biomedical Research Centre (grant BRC-1215-20005). DBG is a recipient of the Swiss government excellence scholarship (grant number 2019.0774), the SSSH+ Global PhD Fellowship Programme in Public Health Sciences of the Swiss School of Public Health, and the Swiss National Science Foundation (project number 176233). NL received funding for the COVID-19 Open Access Project from the Swiss National Science Foundation (grant number 176233) and the European Union’s Horizon 2020 research and innovation programme—project Epise (Grant agreement number 10103688) and acknowledges the contributions of Dr. Leonie Heron and Ms. Hira Imeri. This work was funded by the National Research Programme 78 COVID-19 of the Swiss National Science Foundation (grant number 198418). AC is supported by the National Institute for Health Research (NIHR) Oxford Cognitive Health Clinical Research Facility, by an NIHR Research Professorship (grant RP-2017-08-ST2-006), by the NIHR Oxford and Thames Valley Applied Research Collaboration and by the NIHR Oxford Health Biomedical Research Centre (grant BRC-1215-20005). DBG is a recipient of the Swiss government excellence scholarship (grant number 2019.0774), the SSSH+ Global PhD Fellowship Programme in Public Health Sciences of the Swiss School of Public Health, and the Swiss National Science Foundation (project number 176233). NL received funding for the COVID-19 Open Access Project from the Swiss National Science Foundation (grant number 176233) and the European Union’s Horizon 2020 research and innovation programme—project Epise (Grant agreement number 10103688) and acknowledges the contributions of Dr. Leonie Heron and Ms. Hira Imeri. This work was funded by the National Research Programme 78 COVID-19 of the Swiss National Science Foundation (grant number 198418). AC is supported by the National Institute for Health Research (NIHR) Oxford Cognitive Health Clinical Research Facility, by an NIHR Research Professorship (grant RP-2017-08-ST2-006), by the NIHR Oxford and Thames Valley Applied Research Collaboration and by the NIHR Oxford Health Biomedical Research Centre (grant BRC-1215-20005). DBG is a recipient of the Swiss government excellence scholarship (grant number 2019.0774), the SSSH+ Global PhD Fellowship Programme in Public Health Sciences of the Swiss School of Public Health, and the Swiss National Science Foundation (project number 176233). NL received funding for the COVID-19 Open Access Project from the Swiss National Science Foundation (grant number 176233) and the European Union’s Horizon 2020 research and innovation programme—project Epise (Grant agreement number 10103688) and acknowledges the contributions of Dr. Leonie Heron and Ms. Hira Imeri. This work was funded by the National Research Programme 78 COVID-19 of the Swiss National Science Foundation (grant number 198418). AC is supported by the National Institute for Health Research (NIHR) Oxford Cognitive Health Clinical Research Facility, by an NIHR Research Professorship (grant RP-2017-08-ST2-006), by the NIHR Oxford and Thames Valley Applied Research Collaboration and by the NIHR Oxford Health Biomedical Research Centre (grant BRC-1215-20005). DBG is a recipient of the Swiss government excellence scholarship (grant number 2019.0774), the SSSH+ Global PhD Fellowship Programme in Public Health Sciences of the Swiss School of Public Health, and the Swiss National Science Foundation (project number 176233). NL received funding for the COVID-19 Open Access Project from the Swiss National Science Foundation (grant number 176233) and the European Union’s Horizon 2020 research and innovation programme—project Epise (Grant agreement number 10103688) and acknowledges the contributions of Dr. Leonie Heron and Ms. Hira Imeri. This work was funded by the National Research Programme 78 COVID-19 of the Swiss National Science Foundation (grant number 198418). AC is supported by the National Institute for Health Research (NIHR) Oxford Cognitive Health Clinical Research Facility, by an NIHR Research Professorship (grant RP-2017-08-ST2-006), by the NIHR Oxford and Thames Valley Applied Research Collaboration and by the NIHR Oxford Health Biomedical Research Centre (grant BRC-1215-20005). DBG is a recipient of the Swiss government excellence scholarship (grant number 2019.0774), the SSSH+ Global PhD Fellowship Programme in Public Health Sciences of the Swiss School of Public Health, and the Swiss National Science Foundation (project number 176233). NL received funding for the COVID-19 Open Access Project from the Swiss National Science Foundation (grant number 176233) and the European Union’s Horizon 2020 research and innovation programme—project Epise (Grant agreement number 10103688) and acknowledges the contributions of Dr. Leonie Heron and Ms. Hira Imeri. This work
reflects only the authors’ view. The European Commission is not responsible for any use that may be made of the information it contains. TL is supported by grant UG1 EY020522 from the National Eye Institute, National Institutes of Health.

Disclaimer The views expressed are those of the authors and not necessarily those of the Swiss National Science Foundation. The views expressed are those of the authors and not necessarily those of the UK National Health Service, the NIHR, or the UK Department of Health.

Competing interests TAF reports personal fees from Boehringer-Ingelheim, DT Axis, Kyoto University Original, Shionogi and SONY, and a grant from Shionogori, outside the submitted work; In addition, TAF has patents 2020-5485887 and 2022-082495 pending, and intellectual properties for Kokoro-app licensed to Mitsubishi-Tanabe. AC has received research, educational and consultancy fees from INCIPI (Italian Network for Paediatric Trials), CARRIPO Foundation, Lundbeck and Angelini Pharma. He is the CI/PI of randomised trial about seltorexant in depression, sponsored by Janssen. SL reports personal fees and honoraria from Alkerneis, angelini, Lundbeck, Lundbeck Foundation, Otsuka, Angelini, Eisai, Gedeon, Medicchem, Merck, Mitsubishi, Otsuka, Recordati, Sanofi-Aventis Recordati, Rovi, Teva.

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 4.0 Unported (CC BY 4.0) license, which permits others to copy, redistribute, remix, and transform and build upon this work for any purpose, provided the original work is properly cited, a link to the licence is given, and indication of whether changes were made. See: https://creativecommons.org/licenses/by/4.0/.

ORCID iDs
Thomy Tonia http://orcid.org/0000-0001-7896-6188
Diana Buitrago-García http://orcid.org/0000-0001-9761-206X
Natallie Luise Peter http://orcid.org/0000-0002-5058-3270
Cristina Mesa-Vieira http://orcid.org/0000-0001-6738-4203
Toshi A Furukawa http://orcid.org/0000-0003-2195-3776
Andrea Cipriani http://orcid.org/0000-0001-5179-8321
Stefan Leucht http://orcid.org/0000-0002-4394-4352
Nicola Low http://orcid.org/0000-0003-4817-8986
Georgia Salanti http://orcid.org/0000-0002-3830-8508

REFERENCES