Correspondence on population health surveys and screening tools for depressive disorders: aims and uses by Arias de la Torre et al

Felix Fischer, Dario Zocholl, Geraldine Rauch, Brooke Levis, Andrea Benedetti, Brett Thombs, Matthias Rose

Arias de la Torre et al argue in ‘Population health surveys and screening tools for depressive disorders: aims and uses’ that population-based studies like the European Health Interview Survey (EHIS) are primarily intended to identify and compare vulnerable groups and factors associated with depression on a population level, not to estimate prevalence of major depressive disorders (MDD). We fully agree that the use of brief depression questionnaires in such studies provides meaningful data on the perceived burden by depressive symptoms.

While questionnaire sum scores can be easily analysed, reported and compared, the common practice of reporting the share of participants that score above a given threshold as prevalence is highly problematic. For their widely cited prevalence estimates can only be compared meaningfully on a population level, not to estimate prevalence of major depressive disorders (MDD). We fully agree that the use of brief depression questionnaires in such studies provides meaningful data on the perceived burden by depressive symptoms.

When using different evaluation tools, prevalence estimates can only be compared meaningfully when the imperfect accuracy of each tool in the context of an accepted illness definition is taken into account. Analysis of depression screening tool data in a Bayesian framework enables us to derive prevalence estimates that a cut-off may be false positive screens.

When using different evaluation tools, prevalence estimates can only be compared meaningfully when the imperfect accuracy of each tool in the context of an accepted illness definition is taken into account. Analysis of depression screening tool data in a Bayesian framework enables us to derive prevalence estimates that a cut-off may be false positive screens.

Arias de la Torre et al argue that they employed a broad definition of ‘current depressive disorders’, including MDD, dysthymia and mere occurrence of depressive symptoms or ‘probable cases of depression’. However, there is no evidence that a cut-off of 10 for the eight-item Patient Health Questionnaire depression scale (PHQ-8) accurately distinguishes persons fulfilling these vague definitions from those who do not. On the contrary, at a true prevalence of 5%, over 75% of persons with a score above the cut-off may be false positive screens.

Arias de la Torre et al argue that they employed a broad definition of ‘current depressive disorders’, including MDD, dysthymia and mere occurrence of depressive symptoms or ‘probable cases of depression’. However, there is no evidence that a cut-off of 10 for the eight-item Patient Health Questionnaire depression scale (PHQ-8) accurately distinguishes persons fulfilling these vague definitions from those who do not. On the contrary, at a true prevalence of 5%, over 75% of persons with a score above the cut-off may be false positive screens.

When using different evaluation tools, prevalence estimates can only be compared meaningfully when the imperfect accuracy of each tool in the context of an accepted illness definition is taken into account. Analysis of depression screening tool data in a Bayesian framework enables us to derive prevalence estimates that a cut-off may be false positive screens.

The imprecision of our estimates and lack of power to establish MDD prevalence differences across countries should be understood as a consequence of the PHQ-8’s imperfect diagnostic accuracy, not as evidence for absence of such differences. Revealing this uncertainty is not a weakness but rather a strength of the Bayesian approach.

We are concerned that exaggerations of depression prevalence and the precision of estimates due to ignoring imperfect diagnostic accuracy are misleading and harmful. They potentially identify non-existent and magnify negligible prevalence differences and could therefore lead to mis-targeted research and public health initiatives. The proportion of persons scoring above a cut-off on a screening test should not be reported as prevalence without appropriate adjustment for diagnostic accuracy.

Correspondence to
Dr Felix Fischer, Department of Psychosomatic Medicine, Center for Internal Medicine and Dermatology, Charité – Universitätsmedizin Berlin, corporate member of Freie Universität Berlin, Humboldt-Universität zu Berlin, and Berlin Institute of Health, Berlin, Germany, felix.fischer@charite.de

Received 13 July 2023
Accepted 16 July 2023
Published Online First 8 August 2023

For numbered affiliations see end of article.

Author affiliations
1Department of Psychosomatic Medicine, Center for Internal Medicine and Dermatology, Charité – Universitätsmedizin Berlin, corporate member of Freie Universität Berlin, Humboldt-Universität zu Berlin, and Berlin Institute of Health, Berlin, Germany
2Institute of Biometry and Clinical Epidemiology, Charité – Universitätsmedizin Berlin, corporate member of Freie Universität Berlin, Humboldt-Universität zu Berlin, and Berlin Institute of Health, Berlin, Germany
3Lady Davis Institute for Medical Research, Jewish General Hospital, Montreal, Quebec, Canada
4Department of Epidemiology, Biostatistics and Occupational Health, McGill University, Montréal, Québec, Canada
5Respiratory Epidemiology and Clinical Research Unit, McGill University Health Centre, Montréal, Québec, Canada
6Department of Medicine, McGill University, Montréal, Québec, Canada
7Department of Psychiatry, McGill University, Montréal, Québec, Canada
8Biomedical Ethics Unit, McGill University, Montréal, Québec, Canada
9Biomedical Ethics Unit, McGill University, Montréal, Québec, Canada
10Department of Quantitative Health Sciences, University of Massachusetts Medical School, Worcester, Massachusetts, USA
11Faculty of Public Health, University of Thessaly, Karditsa, Greece

Funding The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests None declared.

Patient consent for publication Not applicable.

Ethics approval Not applicable.

Provenance and peer review Not commissioned; internally peer reviewed.

Data availability statement No data are available.

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
Felix Fischer http://orcid.org/0000-0002-9693-6676
REFERENCES


