



City Research Online

City, University of London Institutional Repository

Citation: Pierce, M., McManus, S., Jessop, C., John, A., Hotopf, M., Ford, T., Hatch, S. L., Wessely, S. & Abel, K. M. (2020). Says who? The significance of sampling in mental health surveys during COVID-19. *Lancet Psychiatry*, doi: 10.1016/S2215-0366(20)30237-6

This is the accepted version of the paper.

This version of the publication may differ from the final published version.

Permanent repository link: <https://openaccess.city.ac.uk/id/eprint/24320/>

Link to published version: [https://doi.org/10.1016/S2215-0366\(20\)30237-6](https://doi.org/10.1016/S2215-0366(20)30237-6)

Copyright: City Research Online aims to make research outputs of City, University of London available to a wider audience. Copyright and Moral Rights remain with the author(s) and/or copyright holders. URLs from City Research Online may be freely distributed and linked to.

Reuse: Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge. Provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.

City Research Online:

<http://openaccess.city.ac.uk/>

publications@city.ac.uk

Says Who? The significance of sampling in mental health surveys during COVID-19

The COVID-19 pandemic is expected to have profound and enduring effects on mental health but, until we have data, we will not know its form, extent, duration, or distribution. An appropriate public health response to mitigate and manage mental health sequelae is likely to require significant diversion of resources. Such decisions must be underpinned by reliable information: policymakers, commissioners and services need to know both the scale of need and who is most vulnerable. A recent position paper in this journal¹ highlights “an immediate priority is collecting high-quality data on the mental health effects of the COVID-19 pandemic across the whole population and vulnerable groups”. This should be a clarion-call for governments to fund, and researchers to gather, timely, high-quality population mental health data which represents the true need arising from the pandemic.

Instead, our desire for quick information has driven rapid propagation of online surveys using non-probability and convenience samples, some of which claim to be representative.

Understandably, many are receiving widespread media attention. These early insights might be valuable, but we caution against relying on them to drive policy and resource because they are prone to substantial bias: acting on misleading information could be worse than having no information at all.

Survey sampling and design choices must be led by their purpose. If the survey is to generate quick ideas, consult on perspectives, or foster community engagement, rapid, low-cost convenience sampling are appropriate. However, to understand prevalence in a population, *how* survey respondents are recruited is critically important. Non-probability samples are usually recruited by approaching membership lists; through service providers; existing large convenience panels; or from snowball recruitment using word-of-mouth, often via social media. Such samples attract volunteers who are already well-engaged, interested in the topic, and who do/can access the internet. Bias can affect any survey, but can be particularly problematic for social/mental health surveys where those excluded are often most in need. Individuals with existing or severe mental illness are less likely to participate online², while half of over 75's and many with mental illness, (key COVID-19 at-risk groups), are not regular internet users; access to digital devices is also limited amongst the most vulnerable and deprived children. Most surveys will weight their sample to match their target population by certain characteristics; however, these adjustments miss crucial elements of bias and

cannot account for groups not included at all, particularly if the response rate is unknowable. A common misconception is that larger samples solve these biases. One Chinese study of mental health responses to the pandemic gathered an impressive 52,730 respondents³; however, 65% were female, indicating a highly skewed sample in a population with significantly fewer women than men⁴. We recommend all surveys detail their sampling strategy and to publish comparative statistics with the population they are sampled from, so informed judgements can be made about representativeness.

The value of a survey depends on its data utility. Non-probability sampling lacks a sound theoretical basis for statistical inference⁵, which means basic descriptive analyses and explorations of potential associations are appropriate but measures of uncertainty i.e. confidence intervals around estimates of prevalence are generally not valid. Moreover, our ability to compare the population's mental health pre- and post-COVID is compromised if surveys do not use standardised measures which are reliable and stable over time and if pre-pandemic baseline data from the same population are not available.

The current crisis has compromised several established data sources: health registries that previously quantified mental illness prevalence⁶ are reporting a drop in patient contacts - in the UK during March 2020, significantly fewer people presented to A&E and primary care with mental illness (personal communications). National registries of mental illness and suicide will catch-up, but are a poor tool in the short-term. Many official surveys have suspended data-collection in response to social distancing guidelines or transferred to remote interviews affecting comparability with prior waves⁷ and creating new challenges, such as how to gather sensitive data on self-harm/suicidality or intimate partner violence⁸.

We believe it is both possible and cost-effective to generate high-quality evidence of mental health need in the current crisis. We recommend using random sampling to reduce risk of bias, allow quantification of non-response, and permit valid statistical analysis. A major investigation into online survey panels⁹ concluded that “Researchers should avoid nonprobability online panels when...[the] objective is to accurately estimate population values.” Determining prevalence of COVID-19 mental health effects should use rigorous methods that sample from the whole population to reduce erroneous conclusions and potentially damaging actions. This approach may be more expensive, but is essential to gain reliable insights into how we mitigate psychological risks during this and future pandemics.

Cutting corners to provide quick, cheap answers **will** result in poorer quality evidence, poorer policy, and wasted resources in the longer term. We can and must do better.

798 words

References

- 1 Holmes EA, Connor RCO, *et al.* Position Paper Multidisciplinary research priorities for the COVID-19 pandemic : a call for action for mental health science. 2020; **0366**: 1–14. [https://www.thelancet.com/journals/lanpsy/article/PIIS2215-0366\(20\)30168-1/fulltext](https://www.thelancet.com/journals/lanpsy/article/PIIS2215-0366(20)30168-1/fulltext)
- 2 Office for National Statistics. Internet users, UK:2018 Office for National Statistics. 2018; : 8.
- 3 Qiu J, Shen B, Zhao M, Wang Z, Xie B, Xu Y. A nationwide survey of psychological distress among Chinese people in the COVID-19 epidemic: implications and policy recommendations. *Gen Psychiatry* 2020; **33**: e100213.
- 4 Chao F, Gerland P, Cook AR, Alkema L. Systematic assessment of the sex ratio at birth for all countries and estimation of national imbalances and regional reference levels. *Proc Natl Acad Sci U S A* 2019. DOI:10.1073/pnas.1812593116.
- 5 Neyman J. On the Two Different Aspects of the Representative Method: The Method of Stratified Sampling and the Method of Purposive Selection. *J R Stat Soc* 1934. DOI:10.2307/2342192.
- 6 Abel KM, Hope H, Swift E, *et al.* Prevalence of maternal mental illness among children and adolescents in the UK between 2005 and 2017: a national retrospective cohort analysis. *Lancet Public Heal* 2019; **4**: e291–300.
- 7 Office for National Statistics Statistical. COVID-19 and the production of statistics. Statements Lett. 2020. <https://www.ons.gov.uk/news/statementsandletters/covid19andtheproductionofstatistics> (accessed April 16, 2020).
- 8 Townsend E, Nielsen E, Allister R, Cassidy SA. Key ethical questions for research during the COVID-19 pandemic. *Lancet Psychiatry*.
- 9 American Association of Public Opinion Research report <https://www.aapor.org/Communications/Press-Releases/AAPOR-Releases-Report-on-Online-Survey-Panels.aspx>

Author list and affiliations:

Dr Matthias Pierce, PhD Centre for Women's Mental Health, Faculty of Biology, Medicine and Health Sciences, University of Manchester, Manchester, UK

Sally McManus, City, University of London; National Centre for Social Research, London

Curtis Jessop, National Centre for Social Research, London

Prof Ann John, MD PhD Population Data Science, Farr Institute of Health Informatics Research, Swansea University Medical School, Swansea University

Prof Matthew Hotopf, MD PhD Department of Psychological Medicine, Institute of Psychiatry, Psychology & Neuroscience, King's College London, London, UK

Prof Tamsin Ford, MD PhD Professor of Child and Adolescent Psychiatry, Department of Psychiatry, University of Cambridge, UK

Prof Stephani Hatch, PhD Department of Psychological Medicine, Institute of Psychiatry, Psychology, and Neuroscience, King's College London, London, UK

Prof Sir Simon Wessely, MD PhD Department of Psychological Medicine, King's College London, London, UK

Prof Kathryn M Abel, MD PhD Centre for Women's Mental Health, Faculty of Biology, Medicine and Health Sciences, University of Manchester, Manchester, UK

Corresponding author:

Dr Matthias Pierce, PhD | Centre for Women's Mental Health, Division of Psychology and Mental Health | Faculty of Biology, Medicine and Health | 3.307 Jean McFarlane Building | 0161 275 5225 | University of Manchester | Oxford Road | Manchester M13 9PL

Email: matthias.pierce@manchester.ac.uk

Conflict of interest statement

Prof John is collaborating with the Mental Health Foundation in their survey of Mental Health during the Covid-19 pandemic. **Dr. Hotopf** reports grants from Innovative Medicines Initiative/European Federation of Pharmaceutical Industries and Associations, outside the submitted work. **Prof. Hatch** reports grants from NIHR, grants from Wellcome Trust, grants from ESRC, from Guy's and St. Thomas' Charity, outside the submitted work; and I am a member of the following advisory groups: NHS England and NHS Improvement - Patient and Carers Race Equalities Framework [PCREF] Steering Group NHS England and NHS Improvement - Advancing Mental Health Equalities Taskforce Health Education England - Mental Health Workforce Equalities Subgroup Maudsley Learning - Maudsley Learning

Advisory Board South London and Maudsley NHS Foundation Trust (SLaM) - Independent
Advisory Groups, the SLaM Partnership Group Lambeth Public Health - Serious Youth
Violence Public Health Task and Finish Group NHS England - Workforce Race Equality
Standard Advisory Group Thrive London - Thrive London Advisory Board Black Thrive -
Black Thrive Advisory Board. All other authors have no conflicts of interest to disclose