The Evolution of Psychiatric Epidemiology: Where to Next?

Oleguer Plana-Ripoll, PhD[1], Mathias Lasgaard, PhD[2],
Zeina N. Mneimneh, PhD[3], and John J. McGrath, MD, PhD[1,4,5]

We have learnt a great deal about the epidemiology of mental disorders over the last 30 years. This has been based on the development and application of a suite of complementary methods—the essential "tool kit" of epidemiology. Surveys based on population-based samples and register-based studies (usually based on help-seeking individuals) have enriched our understanding on how mental disorders impact on society. The content and format of these tools have evolved in the face of shifting diagnostic boundaries and the steady march of improved technology. In this concise perspective, we reflect on the nature and pace of recent changes within the field of psychiatric epidemiology. Mindful of how bad we are at predicting the future, we will also take the risk of speculating on the shape of things to come.

While the definition of epidemiology has evolved over time,[1] we will use a simple, easy-to-understand definition. Coggon, Rose and Barker open their (freely available, online) book with this definition: "Epidemiology is the study of how often diseases occur in different groups of people and why. Epidemiological information is used to plan and evaluate strategies to prevent illness and as a guide to the management of patients in whom disease has already developed" (p. 1).[2] Epidemiology is a field of health science that assesses the frequency and rates of disorders and their risk factors (i.e., the "determinants" of health) over time and place (including the distribution of disorders and their risk factors). Within the taxonomy of frequency measures, we have measures of incidence (e.g., a rate that counts new cases per background population per unit of time) and prevalence (e.g., a count of current cases assessed over different periods of time such as 1 month, 1 year, or a lifetime). Around these measures, epidemiology explores the causes and consequences of mental disorders. Research related to identifying the causes of mental disorders operates within an unbounded search space—but progress has been made in the assessment of environmental exposures and genetic risk factors. With respect to the consequences of mental disorders,

measures related to morbidity and mortality will continue to figure prominently.

If the future follows the past, we will continue to have changes in the diagnostic criteria. While we can build algorithms for backward-looking crosswalks (e.g., *DSM*-5 to 4), we cannot future-proof studies against unforeseen changes in diagnostic criteria. Sometimes there are grounds to "freeze" items related to the measurement of symptoms and/or diagnoses—especially in panel studies over time (i.e., repeated measures). Here, fixing the items can provide insights into secular changes across time (i.e., cohort effects). Waves of individuals within a fixed age range can be compared across different decades. Recently, the Wellcome Trust and the National Institute of Mental Health have recommended a set of self-report measures for youth anxiety and depression.[3] Having uniform criteria (at the diagnostic or symptom level) can aid between-study comparisons and meta-analyses. This recommendation has attracted some criticism[4] and may hinder innovation.

The number of studies based on administrative registers has increased exponentially over the past 2 decades. Population-based registers are a gold mine for epidemiologists; nowadays, it is possible to use de-identified data to link

[1] National Centre for Register-based Research, Aarhus University, Denmark
[2] DEFACTUM—Public Health & Health Services Research, Central Denmark Region, Aarhus, Denmark
[3] Survey Research Center, Institute for Social Research, University of Michigan, Ann Arbor, MI, USA
[4] Queensland Brain Institute, University of Queensland, St Lucia, Queensland, Australia
[5] Queensland Centre for Mental Health Research, The Park Centre for Mental Health, Queensland, Australia

Corresponding Author:
John J. McGrath, MD, PhD, Queensland Centre for Mental Health Research, The Park Centre for Mental Health, Queensland, Australia.
Email: j.mcgrath@uq.edu.au

each individual to their family members, their vital status, their health conditions (e.g., hospitalizations, visits to the general practitioner, drug prescriptions), and their socioeconomic characteristics (e.g., annual income, highest education achieved), among other information. The main strengths of using registers for epidemiological studies are that data are already collected—and collected prospectively—and registers include the entire population followed over many years; the risk of selection bias is thus negligible. Additionally, sample sizes are often large, which makes it possible to study rare exposes or outcomes. Many important findings from psychiatric epidemiology have been based on register-based studies, for example, recent studies estimating the incidence and prevalence of a broad spectrum of mental disorders.[5] Canada also has a proud tradition of psychiatric epidemiology based on register (or administrative) data.[6-8]

There are several limitations related to register-based epidemiology. The accuracy of clinical diagnoses recorded in administrative data can be of variable quality—clinical validation studies can help determine the extent of this measurement error. While some health registers are based on services that cover the entire population, in order to appear in these registers, an individual has to first seek help. Thus, registers are more likely to reflect severe cases, while people who do not seek help are not included in these registers. Those included in modern health registers can be "left truncated"—researchers do not know what happened before the registers were implemented, which make it difficult to differentiate prevalent from incident cases of a particular disease. Additionally, the long follow-up and the availability of all data at different points in time might require specific analytical tools to deal with potential biases,[9] which are less prominent in studies with nonhistorical data collection. Registers can be a good source of age of first presentation/diagnosis; however, this date may be some time after the actual age of onset of the presenting condition. Details on recovery (i.e., the "age of offset") are rarely available. Finally, data are usually precollected for nonresearch purposes, and some key information might be poorly registered, misclassified, or simply unavailable. In order to enrich information from register-based studies, it is necessary to link registers with other data sources. In Denmark, a successful example in psychiatric epidemiology is the iPSYCH case–cohort study,[10] which cross-links register and genetic data from a representative sample randomly selected from the entire population.[11]

As mentioned above, psychiatric epidemiology based solely on health registers may be biased because of an over-representation of clinically severe cases.[12] This misclassification can be addressed by linking health registers to representative surveys including questions about mental health symptoms, which would also identify the milder cases of mental disorders. Canadian researchers have confirmed this by linking survey and administrative data on anxiety and mood disorders in Ontario.[13] Additionally, surveys usually provide information on health behavior, social relationships, and other variables such as health literacy that are rarely available through administrative registers. The impact of unmeasured confounders or potential misclassification in register-based epidemiology could therefore be assessed in a subsample for which information from surveys is available.

Moreover, mental health data from general surveys may help monitor population-based trends over time. In order to carry out effective mental health policy planning, accurate general population data on current mental health prevalence are vital. These can be measured with specific structured interviews (e.g., the Composite International Diagnostic Instrument). Mental health can be assessed in routine general health surveys; however, with some notable exceptions[14], it is difficult to find brief measures of mental disorders with high sensitivity and specificity that are validated for population-based studies. Future research would benefit from further validation of psychometrically robust measures that enable the monitoring of specific mental health disorders in the general population and the identification of (sub)-clinical cases.

The Danish National Health Surveys are examples of population-based survey (conducted every 4 years) that examine health-related factors in the adult population not covered in administrative registers.[15] The sample is identified through the Civil Registration System using unique personal identification numbers (i.e., survey data can be linked at the individual level to register data), and the data collection is based on a self-administered questionnaire distributed to approx. 300,000 citizens including all municipalities (response rate 54% to 60%).[15]

However, response rates to community surveys have been declining in recent years,[16,17] which may comprise the representativeness and generalizability of population-based data. Thus, nonrespondent analysis and reweighting techniques are more crucial than ever but still depend on information available about the nonrespondent population. An interesting approach to this is the linkage of survey and register data (e.g., demographic and health-related data) to calibrate weights that account for nonresponse.

While surveys will continue to be important for mental health research, over the past decade, innovative data collection methods have emerged. Such methods include the use of experience sampling and the use of social media data—these methods can be blended with surveys or health registers.

Experience sampling methods (ESM) is a self-report diary technique derived from an ecological psychology concept that assess "mood, symptoms, context and appraisals thereof as they occur in daily life" (p. 123).[18] ESM uses a signaling-mobile, wearable, or alternative personal device to collect real-time individual data within the moment, on repeated occasions, similar to a diary method.[19] Although ESM was initially conceptualized decades ago, the widespread use of mobile phones has made the method more accessible.[20] ESM has several applications within the mental health research field including improving the understanding of specific symptoms, capturing emotional variability,

identifying internal and situational determinants, uncovering person–environment interactions, examining real-world social interactions, and evaluating treatments and interventions.

Many methodological questions need to be answered for evaluating the quality of mental health data from ESM studies. Currently, ESM studies employ vastly different design and implementation features including devices, frequency of collection, and time interval between data collection points. This makes it difficult to derive solid methodological recommendations for its use, especially in general population samples.[21] Moreover, compliance, retention, and missing data rates seem to vary by respondent characteristics including gender and mental health status, leading to sample imbalances[21,22] and loss of precision. For example, the meta-analysis by Vachon et al.[21] found that a range of simple measures were associated with better compliance and retention rates (e.g., fixed sampling scheme, higher incentives, fewer evaluations per day). As technology improves (e.g., access to physiological measures related to movement and pulse rate), it will be feasible to link time-stamped self-report and physiological measures. We encourage the next generation of psychiatric epidemiologists to explore these developments which can provide fine-grained data related to health and well-being.

Another source of data that has started to emerge in mental health research is digital trace data from social media applications such as Twitter. Sinnenberg and colleagues[23] systematically reviewed 137 published articles that have used Twitter for health-related research between 2010 and 2015 and found that there was a 2-fold increase in such publications each year. The most commonly studied fields were public health (22%), infectious disease (20%), behavioral medicine (18%), and psychiatry (11%). Gamon and colleagues[24] provide an example on how Twitter data have been used to build a classifier that provides an estimate of the risk of depression at an individual level. However, fundamental methodological questions (e.g., sample representation, measurement accuracy) remain unanswered since the data generation process of social media is still largely unknown.[25] Linking multiple data sources such as survey data and Twitter data could shed light on the data properties of social media[26] providing mental health researchers a reliable framework to use when such data are used solely or in combination with other data sources.

The contribution of mental disorders to the overall burden of disease has been highlighted by the Global Burden of Disease studies.[27] In 2017, these disorders accounted for 18.0% of the entire global burden of disease.[28] However, there are concerns about how these studies underestimate the contribution of premature mortality in those with mental disorders.[29] Recently introduced demographic method have provided innovative ways to capture the association between particular mental disorders and premature mortalitys.[30]

Finally, it should be noted that major advances in psychiatric genetics have been made in recent years.[31] While these advances are generally not yet ready for clinical translation, this will certainly change in the decades ahead.[32] From a research perspective, methods designed to combine genetic and environmental data are advancing rapidly. This augurs well for innovation in our understanding the risk architecture of mental disorders. Epidemiology can play an important role in these developments.

If we are to meet the future challenges of psychiatric epidemiology (e.g., the impact of COVID-19 on mental health), we need to continue to improve our measurement instruments, and we need to strength our capacity to leverage multiple data sources through different types of linkages. Linking different sources of data will require us to increase our collaboration across disciplines, institutions, and industry. The discoveries from epidemiology will be vital to guide many rapidly evolving fields of health research (e.g., precision medicine, decision support tools in clinical practice). The era of relying solely on one source of data might be coming to an end, as we open the door to a new era of interdisciplinary multi-data source collaborations. To be successful in this new era, we need to grow the work force, remain creative, reach out for new collaborations, blend our methods, and set traps for discovery.

Declaration of Conflicting Interests

Funding

ORCID iD

John J. McGrath, MD, PhD ⓘ https://orcid.org/0000-0002-4792-6068

References

1. Frerot M, Lefebvre A, Aho S, Callier P, Astruc K, Aho Glele LS. What is epidemiology? changing definitions of epidemiology 1978-2017. PLoS One. 2018;13(12):e0208442.

2. Coggon D, Barker D, Rose G. Epidemiology for the uninitiated. 5th ed. London, UK: BMJ Publishing Group; 2003.

3. Wolpert M. Funders agree first common metrics for mental health science. London, UK: Welcome Trust; 2020.

4. Patalay P, Fried EI. Prescribing measures: unintended negative consequences of mandating standardized mental health measurement. PsyArXiv. 2020.

5. Pedersen CB, Mors O, Bertelsen A, et al. A comprehensive nationwide study of the incidence rate and lifetime risk for

treated mental disorders. JAMA Psychiatry. 2014;71(5): 573-581.

6. Butler AL, Smith M, Jones W, et al. Multi-province epidemiological research using linked administrative data: a case study from Canada. Int J Popul Data Sci. 2018;3(3):443.

7. Lesage A, Rochette L, Emond V, et al. A surveillance system to monitor excess mortality of people with mental illness in Canada. Can J Psychiatry. 2015;60(12):571-579.

8. Kisely S, Sadek J, MacKenzie A, Lawrence D, Campbell LA. Excess cancer mortality in psychiatric patients. Can J Psychiatry. 2008;53(11):753-761.

9. Hansen SN, Overgaard M, Andersen PK, Parner ET. Estimating a population cumulative incidence under calendar time trends. BMC Med Res Methodol. 2017;17(1):7.

10. Pedersen CB, Bybjerg-Grauholm J, Pedersen MG, et al. The iPSYCH2012 case-cohort sample: new directions for unravelling genetic and environmental architectures of severe mental disorders. Mol Psychiatry. 2018;23(1):6-14.

11. Albinana CC, Vilhjalmsson BJ, McGrath JJ. Families, health registers, and biobanks: making the unmeasurable measurable. Biol Psychiatry. 2020;88(6):440-441.

12. Musliner KL, Liu X, Gasse C, Christensen KS, Wimberley T, Munk-Olsen T. Incidence of medically treated depression in Denmark among individuals 15-44 years old: a comprehensive overview based on population registers. Acta Psychiatr Scand. 2019;139(6):548-557.

13. Edwards J, Thind A, Stranges S, Chiu M, Anderson KK. Concordance between health administrative data and survey-derived diagnoses for mood and anxiety disorders. Acta Psychiatr Scand. 2020;141(4):385-395.

14. Kessler RC, Andrews G, Colpe LJ, et al. Short screening scales to monitor population prevalences and trends in non-specific psychological distress. Psychol Med. 2002;32:959-976.

15. Christensen AI, Ekholm O, Glumer C, et al. The Danish national health survey 2010. study design and respondent characteristics. Scand J Public Health. 2012;40(4):391-397.

16. Mindell JS, Giampaoli S, Goesswald A, et al. Sample selection, recruitment and participation rates in health examination surveys in Europe—experience from seven national surveys. BMC Med Res Methodol. 2015;15:78.

17. Stedman RC, Connelly NA, Heberlein TA, Decker DJ, Allred SB. The end of the (research) world as we know it? understanding and coping with declining response rates to mail surveys. Soci Nat Resour. 2019;32(10):1139-1154.

18. Myin-Germeys I, Kasanova Z, Vaessen T, et al. Experience sampling methodology in mental health research: new insights and technical developments. World Psychiatry. 2018;17(2): 123-132.

19. Csikszentmihalyi M. Flow and the foundations of positive psychology: the collected works of Mihaly Csikszentmihalyi. 1st ed. Dordrecht, the Netherlands: Springer; 2014:298. 1 online resource (XXIV, 298 pages 234 illustrations).

20. Cherubini M, Oliver N. A refined experience sampling method to capture mobile user experience. In: Nakhimovsky D. Eckles, J. Riegelsberger, editors. Presented at the International Workshop of Mobile User Experience Research part of CHI. Boston, MA, USA: 2009.

21. Vachon H, Viechtbauer W, Rintala A, Myin-Germeys I. Compliance and retention with the experience sampling method over the continuum of severe mental disorders: meta-analysis and recommendations. J Med Internet Res. 2019;21(12): e14475.

22. Messiah A, Grondin O, Encrenaz G. Factors associated with missing data in an experience sampling investigation of substance use determinants. Drug Alcohol Depend. 2011;114(2-3):153-158.

23. Sinnenberg L, Buttenheim AM, Padrez K, Mancheno C, Ungar L, Merchant RM. Twitter as a tool for health research: a systematic review. Am J Public Health. 2017; 107(1):E1-E8.

24. Gamon M, Choudhury M, Counts S, Horvitz E. Predicting depression via social media. In: Proceedings of the Seventh International AAAI Conference on Weblogs and Social Media. California, USA: 2013.

25. Singh L, Bode L, Davis-Kean P, et al. Study designs for quantitative social science research using social media. PsyArXiv. 2020.

26. Mneimneh ZN, McClain C, Bruffaerts R, Altwaijri YA. Evaluating survey consent to social media linkage in three international health surveys. Res Social Adm Pharm. 2020.

27. Whiteford HA, Degenhardt L, Rehm J, et al. Global burden of disease attributable to mental and substance use disorders: findings from the global burden of disease study 2010. Lancet. 2013;382(9904):1575-1586.

28. GBD 2017 Disease and Injury Incidence and Prevalence Collaborators. Disease injury incidence prevalence collaborators. global, regional, and national incidence, prevalence, and years lived with disability for 354 diseases and injuries for 195 countries and territories, 1990-2017: a systematic analysis for the global burden of disease study 2017. Lancet. 2018;392(10159): 1789-1858.

29. Vigo D, Thornicroft G, Atun R. Estimating the true global burden of mental illness. Lancet Psychiatry. 2016;3(2): 171-178.

30. Plana-Ripoll O, Pedersen CB, Agerbo E, et al. A comprehensive analysis of mortality-related health metrics associated with mental disorders: a nationwide, register-based cohort study. Lancet. 2019;394(10211):1827-1835.

31. Hyman SE. New evidence for shared risk architecture of mental disorders. JAMA Psychiatry. 2019;76(3):235-236.

32. Murray GK, Lin T, Austin J, McGrath JJ, Hickie IB, Wray NR. Could polygenic risk scores be useful in psychiatry? a review. JAMA Psychiatry. 2020.