https://eprints.gla.ac.uk/295557/

Deposited on 7 June 2023

We need timely access to mental health data: implications of the Goldacre review

The Goldacre review, published in April, 2022[1], is a landmark evaluation of the use, availability, and safety of National Health Service (NHS) data across all four nations of the UK. The review underscores the necessary role of data in driving health-care improvement and innovation, and the potential risks inherent in using data routinely contributed by health service users. The review recommends a radical overhaul in NHS data curation, access, and analysis, and, crucially, argues that substantial new resources must be marshalled to make this aspiration a reality.

As experts in child and adolescent mental health, we support many of the review's recommendations. With regard to mental health data, many international systems struggle with inconsistent case definitions, fragmented record keeping, breaches in confidentiality, and the frequency and format of dissemination[2]. For example, in the USA the monitoring of child and adolescent mental health is spread across multiple different agencies operating at both state and federal levels, with entirely different reporting systems for different sets of symptoms and conditions[2,3]. In contrast, in some European nations, universal health-care systems combined with routine registration of health and life events and enhanced individual-level linkage provide a powerful and relatively user-friendly resource for informing public policy, basic science, and health-care provision[4]. Despite the UK's universal health-care system, as the Goldacre review outlines, there are substantial problems with the management, curation, processing, and overall use of NHS data. We agree that use of NHS data needs to be brought in line with best practice in contemporary data science, including online open source code, reproducible pipelines, accessible data dictionaries, standardised formatting, and strong information governance. Similarly, we agree that long-term investment in the training of data analysts and stronger public and patient participation in research governance are essential. However, based on our experience, we also see pitfalls of some potential interpretations of the review's recommendations. If not carefully managed, these interpretations could seriously undermine the use of many valuable data.

The Goldacre review focuses on analytics to drive NHS operational decision making and service planning, particularly using data from general practitioners. These data contain patient-level information, including demographics, any conditions, medications, and test results. Many of the recommendations flow from this focus on general practitioner data. For example, a recurring theme is an enhanced role for data analysts within the NHS in accessing and analysing data, reducing the need for these sensitive data to be accessed from outside the NHS. While this change will make highly sensitive data more secure, will the same approach be applied to all NHS-controlled data and all applications? Access for analysts outside of the NHS is mentioned in the review, but the potential implication is that external access to all NHS data will face additional constraints and delays. Many uses of health-related data extend well beyond the planning of service provision. For example, such data can be used to identify unknown antecedents to clinical symptoms as potential therapeutic targets, or new constellations of symptoms that enhance existing clinical classification frameworks. This kind of innovation is particularly crucial within child and adolescent mental health, where shifting to an active model of prevention is an urgent priority. Timely discovery science would continue to be impeded if most analyses were run by NHS analysts, or if data sharing were tailored to the types of analyses typically run by NHS analysts.

Maximising the potential of health-care data requires collaborative innovation, flexible and time-sensitive data processing, teams of scientists trying alternative approaches, and early career researchers using tools and techniques from other fields. All these activities necessitate direct access to data, which might be at risk with some interpretations of the Goldacre review. Beyond the

incorporation of some of these elements within a new NHS analyst service, there is little mention in the review of how this access would be enabled and afforded to other, appropriate, stakeholders, such as those working in university settings. Relatedly, the review's recommendation of enhanced training and professional recognition for NHS analysts is welcome, but this is no replacement for experienced multidisciplinary teams working together on datasets, all of whom would derive benefit.

Identifiable, routinely collected patient data do require tight control, but our fear is that the same model will be applied to all data. What constraints will extend to data that have been anonymised, and to data specifically designed and obtained with informed consent for research (such as the national mental health surveys, longer-term follow-ups of clinical trials, and cohort studies)? The Goldacre review suggests that data with explicit consent can be treated differently, but our experience is that the most restrictive practice quickly becomes the default for all health-related data. This restriction in practice was shown by the substantial reduction in downloads and uses of the Adult Psychiatry Morbidity Study 2014, relative to its 2007 counterpart, following the imposition of restrictions despite the data being obtained with explicit consent for research purposes:[5] 4 years after the 2007 release, data had been accessed 165 times, producing 30 published papers. By the equivalent time point following the 2014 release there were only two publications.

Different populations hold different attitudes about the use of their data, but a UK survey of nearly 30,000 people suggests net support for researchers' access to de-identified data[6]. Special consideration is required for engagement with vulnerable groups, especially those who might experience barriers to conventional consent procedures, not least because they might be more readily identifiable. We regularly work with datasets collected for the explicit purpose of tracking mental health. A recent preprint across different global contexts has shown that when young people trust the environment in which their data are used, they actively want their mental health data to be used for public good[7]. We can increase this trust by involving them in decision making and priority setting with regard to their data. These context-dependent attitudes to data sharing highlight both the temporal and longitudinal sensitivity of any data concerning children and adolescents; a good understanding of the societal and developmental context in which the data were collected, as well as the developmental plasticity of any child and adolescent clinical data, is essential in making this approach work for this population.

A key recommendation of the Goldacre review is the centralisation of **trusted research environments** such that a small number of these centres would curate all NHS data. Our experience over the past decade is of substantial delays in accessing NHS-controlled data, with trusted research environments imposing additional restrictions on data access, even when explicit consent for data sharing has been given. For example, the Mental Health Survey of Children and Young People 2017 data were available only in late 2020, while the COVID-19 pandemic follow-up data collected in 2020 and 2021, are still not available in January 2023. Almost 3 years into the COVID-19 pandemic and associated social restrictions, we still do not know their full effect on children and young people's mental health. We fear that the proposed centralisation of data curation and analysis will further constrain data access and delay policy and practice-relevant discovery science.

Centralisation also raises practical challenges. If all analysis must be done within trusted research environments, who will bear this cost and how will access be prioritised? The research environments will need sufficient capacity (personnel and processing) to undertake different analyses, using a range of different coding languages. What will happen to existing trusted research environments? What will the waiting time be before the trusted research environments and application process are ready? What happens in the meantime to NHS-controlled datasets that are barely accessed?

In summary, we welcome many of the Goldacre review's recommendations, but are concerned about potential interpretations. Child and adolescent mental health is particularly dynamic, complex, and context dependent, as has been amply shown during the COVID-19 pandemic[8-10]. Timely access to current data is essential, if we are to answer the most urgent questions about mental health and provide responsive, evidence-based care for children and adolescents.

Duncan E. Astle[1,2], Anna Moore[1], Louise Marryat[3], Essi Viding[4], Karen L. Mansfield[5], Mina Fazel[5], Matthias Pierce[6,] Kathryn. M. Abel[6,7], Jonathan Green[6], Ann John[8], Matthew R. Broome[9,10], Rachel Upthegrove[9,10], Helen Bould[11,12], Helen Minnis[13], Ruchika Gajwani[13], Madeleine J. Groom [14, 15], Chris Hollis[16], Elizabeth Liddle[15], Kapil Sayal [15,17], Vashti Berry[18], Stephan Collishaw[19], Helen Dawes[20], Samuele Cortese[21,22,], Mara Violato[23], Jack Pollard[23], James H. MacCabe[24], Sarah-Jayne Blakemore[25], Emily Simonoff[24], Edward Watkins[26], Rachel M. Hiller[4,27], Ellen Townsend[28], Cherie Armour[29], John R. Geddes[30], Lucy Thompson[31], Matthias Schwannauer[32], Dasha Nicholls[33], Matthew Hotopf[24], Johnny Downs[24], Atif Rahman[34], Aditya Narain Sharma[35,36], and Tamsin J. Ford[1]

1. Department of Psychiatry, University of Cambridge, England, UK
2. MRC Cognition and Brain Sciences Unit, University of Cambridge England, UK
3. School of Health Sciences, University of Dundee, Scotland, UK
4. Division of Psychology and Language Sciences, University College London, England, UK
5. Department of Psychiatry, University of Oxford, England, UK
6. Division of Psychology and Mental Health, University of Manchester, England, UK
7. Greater Manchester Mental Health NHS Foundation Trust, England, UK
8. Population Data Science, Swansea University, Wales, UK
9. Institute for Mental Health, University of Birmingham, England, UK
10. Birmingham Women's and Children's NHS Foundation Trust, England, UK
11. Centre for Academic Mental Health, Population Health Science, University of Bristol, England, UK
12. MRC Integrative Epidemiology Unit, University of Bristol, England, UK
13. School of Health and Wellbeing, University of Glasgow, Scotland, UK
14. Academic Unit of Mental Health & Clinical Neurosciences, School of Medicine, University of Nottingham, England, UK
15. Centre for ADHD & Neurodevelopmental Disorders Across the Lifespan, University of Nottingham, England UK
16. NIHR MindTech MIC & NIHR Nottingham Biomedical Research Centre, Mental Health & Clinical Neurosciences, School of Medicine, University of Nottingham, England, UK
17. Academic Unit of Mental Health & Clinical Neurosciences, School of Medicine, University of Nottingham, England, UK
18. Children & Young People's Mental Health Research Collaboration, University of Exeter, England, UK
19. Wolfson Centre for Young People's Mental Health, Cardiff University, Wales, UK
20. NIHR Exeter Biomedical Research Centre, University of Exeter, England, UK
21. Centre for Innovation in Mental Health, University of Southampton, England, UK
22. Solent NHS Trust, Southampton, England, UK
23. Nuffield Department of Population Health, University of Oxford, England, UK
24. Institute of Psychiatry, Psychology and Neuroscience, King's College London, England, UK
25. Department of Psychology, University of Cambridge, England, UK
26. School of Psychology, University of Exeter, England, UK

27. Anna Freud National Centre for Children and Families, University College London, England, UK
28. School of Psychology, University of Nottingham, England, UK
29. School of Psychology, Queen's University Belfast, Northern Ireland, UK
30. NIHR Oxford Health Biomedical Research Centre, University of Oxford, England, UK
31. Institute of Applied Health Sciences, University of Aberdeen, Scotland, UK
32. Centre for Applied Developmental Psychology, University of Edinburgh, Scotland, UK
33. Division of Psychiatry, Imperial College London, England, UK
34. Institute of Population Health, University of Liverpool, England, UK
35. Translational and Clinical Research Institute, University of Newcastle, England, UK
36. Cumbria Northumberland Tyne and Wear NHS Foundation Trust, Newcastle Upon Tyne, England, UK

Corresponding author:

Professor Duncan E. Astle
Department of Psychiatry, University of Cambridge, Cambridge, UK
duncan.astle@mrc-cbu.cam.ac.uk

Competing interests declaration:

MH leads the RADAR-CNS consortium, a private–public pre-competitive collaboration on mobile health, through which his university receives in-kind and cash contributions from Janssen, Biogen, UCB, Merck, and Lundbeck. All other authors declare no competing interests.

References

1. Goldacre B, Morley J. Better, Broader, Safer: Using health data for research and analysis. A review commissioned by the Secretary of State for Health and Social Care. Department of Health and Social Care. 2022. Available from: https://www.gov.uk/government/publications/better-broader-safer-using-health-data-for-research-and-analysis (accessed: January 26th, 2023)

2. Lyerla R, Stroup DF. Toward a public health surveillance system for behavioral health. Public Health Reports. 2018 Jul;133(4):360-5.

3. Bitsko RH, Claussen AH, Lichstein J, Black LI, Jones SE, Danielson ML, Hoenig JM, Jack SP, Brody DJ, Gyawali S, Maenner MJ. Mental health surveillance among children—United States, 2013–2019. MMWR supplements. 2022 Feb 25;71(2):1.4. Schmidt M, Schmidt SA, Adelborg K, Sundbøll J, Laugesen K, Ehrenstein V, Sørensen HT. The Danish health care system and epidemiological research: from health care contacts to database records. Clinical epidemiology. 2019;11:563.

5. Ford T, Mansfield KL, Markham S, McManus S, John A, O'reilly D, Newlove-Delgado T, Iveson MH, Fazel M, Munshi JD, Dutta R. The challenges and opportunities of mental health data sharing in the UK. The Lancet Digital Health. 2021 Jun 1;3(6):e333-6.

6. Jones LA, Nelder JR, Fryer JM, Alsop PH, Geary MR, Prince M, Cardinal RN. Public opinion on sharing data from health services for clinical and research purposes without explicit consent: an anonymous online survey in the UK. BMJ open. 2022 Apr 1;12(4):e057579. 7. Sieberts S, Marten C, Bampton E, Björling EA, Burn AM, Carey E, Carlson S, Fernandes B, Kalha J, Lindani S, Masomera H. Young people's data governance preferences for their mental health data: MindKind Study findings from India, South Africa, and the United Kingdom. medRxiv. 2022 Jan 1.

8. Bignardi G, Dalmaijer ES, Anwyl-Irvine AL, Smith TA, Siugzdaite R, Uh S, Astle DE. Longitudinal increases in childhood depression symptoms during the COVID-19 lockdown. Archives of disease in childhood. 2021 Aug 1;106(8):791-7.

9. Pierce M, Hope H, Ford T, Hatch S, Hotopf M, John A, Kontopantelis E, Webb R, Wessely S, McManus S, Abel KM. Mental health before and during the COVID-19 pandemic: a longitudinal probability sample survey of the UK population. The Lancet Psychiatry. 2020 Oct 1;7(10):883-92.

10. Pierce M, McManus S, Hope H, Hotopf M, Ford T, Hatch SL, John A, Kontopantelis E, Webb RT, Wessely S, Abel KM. Mental health responses to the COVID-19 pandemic: a latent class trajectory analysis using longitudinal UK data. The Lancet Psychiatry. 2021 Jul 1;8(7):610-9.