**We need timely access to mental health data – A brief response to the Goldacre Review**

Duncan E. Astle1,2, Anna Moore1, Louise Marryat3 , Essi Viding4, Karen L. Mansfield5, Mina Fazel5, Matthias Pierce6, Kathryn.M.Abel6,7, Jonathan Green6  Matthew R. Broome8, Helen Bould9, Helen Minnis11, Elizabeth Liddle16, Kapil Sayal 16,Chris Hollis12, Madeleine J Groom 12, 16 , Stephan Collishaw15, Vashti Berry13, Ruchika Gajwani11 Helen Dawes 18 Samuele Cortese 19 , Lucy Thompson21, Edward Watkins14 Aditya Narain Sharma 22, 23 Sarah-Jayne Blakemore24 Emily Simonoff 25 Matthias Schwannauer 26, Dasha Nicholls 28 , Johnny Downs 17 John R. Geddes29 [please insert your names, don’t forget to add affiliations]Rachel Upthegrove30, Matthew Hotopf17, Ann John7, Tamsin J. Ford1 , Atif Rahman20, Cherie Armour26 , Ellen Townsend30 , Rachel M Hiller 31, Jack Pollard32, Mara Violato32, James H MacCabe17

1. Department of Psychiatry, University of Cambridge, Cambridge, UK
2. MRC Cognition and Brain Sciences Unit, University of Cambridge, UK
3. School of Health Sciences, University of Dundee, UK.
4. Division of Psychology and Language Sciences, University College London, London, UK
5. Department of Psychiatry, University of Oxford, Oxford, UK
6. Centre for Women’s Mental Health, Division of Psychology and Mental Health, University of Manchester, UK
7. Greater Manchester Mental Health NHS Foundation Trust, Manchester, UK
8. Population Data Science, Swansea University Medical School. UK
9. Institute for Mental Health, University of Birmingham, UK and Birmingham Women’s and Children’s NHS Foundation Trust, UK.
10. Centre for Academic Mental Health, Population Health Sciences, University of Bristol, Oakfield House, Oakfield Grove, Bristol, UK & MRC Integrative Epidemiology Unit, University of Bristol, Oakfield House, Oakfield Grove, Bristol, UK
11. Mental Health and Wellbeing, School of Health and Wellbeing, University of Glasgow, UK
12. NIHR MindTech MIC and Nottingham Biomedical Research Centre, Mental Health and Clinical Neurosciences, School of Medicine, University of Nottingham, UK
13. Children & Young People’s Mental Health Research Collaboration (ChYMe), University of Exeter, Exeter, UK
14. School of Psychology, University of Exeter, Exeter, UK
15. Wolfson Centre for Young People’s Mental Health, Cardiff University, UK
16. Centre for ADHD & Neurodevelopmental Disorders Across the Lifespan (CANDAL), Institute of Mental Health, School of Medicine, University of Nottingham, UK
17. Institute of Psychiatry Psychology and Neuroscience, King’s College London.
18. Exeter BRC NIHR, Medical School, University of Exeter
19. Centre for Innovation in Mental Health, University of Southampton, UK & Solent NHS Trust
20. Institute of Population Health, University of Liverpool, UK
21. Institute of Applied Health Sciences, University of Aberdeen, UK
22. Translational and Clinical Research Institute, Newcastle University UK
23. Cumbria Northumberland Tyne and Wear NHS Foundation Trust
24. Department of Psychology, University of Cambridge, UK
25. Department of Child and Adolescent psychiatry, King’s College London, Institute of Psychiatry, Psychology and Neuroscience, London UK
26. Research Centre for Stress, Trauma and Related Conditions (STARC), School of Psychology, Queen's University Belfast (QUB), Belfast, Northern Ireland
27. Centre for Applied Developmental Psychology, The University of Edinburgh, School of Health in Social Science, Edinburgh, UK.
28. Division of Psychiatry, Imperial College London
29. NIHR Oxford Health Biomedical Research Centre and University of Oxford
30. School of Psychology, University of Nottingham, UK.
31. Division of Psychology & Language Sciences, University College London, London UK
32. Nuffield Department of Population Health, Health Economics Research Centre, University of Oxford, UK

The Goldacre Review, published in April 2022 (Goldacre & Morley, 2022), reflects a landmark evaluation of the utility, availability and safety of NHS data. It underscores the necessary role of data in driving healthcare 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, adolescent and youth mental health, we support many of the review’s recommendations. We agree that NHS analytics 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. Likewise, we agree that long-term investment in the training of analysts and stronger public and patient participation in research governance are both essential. However, based on our experience, we also see potential pitfalls of some recommendations. If not carefully managed these could seriously undermine the utility of much valuable data.

***Different analytics have different requirements***

The review focuses on analytics to drive NHS operational decision making and service planning, particularly using GP data. Many of the recommendations flow from this focus. For example, a recurring theme is an enhanced role for NHS analysts in accessing and analysing data. Whilst this will make highly sensitive data more secure, *will the same approach be applied to all datasets and all applications?* Access for analysts outside of the NHS is mentioned, but the potential implication is that external access to all NHS data will face additional constraints and delays.

Many uses of data extend well beyond the planning of service provision. Identifying new antecedents to clinical symptoms as potential therapeutic targets, or new constellations of symptoms that enhance existing clinical classification frameworks, are just two examples. This kind of innovation is particularly vital within child and adolescent mental health, where shifting to a prospective and 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 is tailored to the types of analyse s typically run by NHS analysts.

Maximising the potential of data requires collaborative innovation, flexible and time sensitive data processing, teams of scientists trying alternative approaches, and early career researchers utilising tools and techniques from other fields. All these activities necessitate direct access to data, which may be at risk under some interpretations of the review. Beyond the incorporation of some of these elements within an NHS analytic service, there is little mention of how this access would be enabled and afforded to other, appropriate, stakeholders. Relatedly, the recommended 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.

***Not all datasets need the same treatment***

Identifiable, routinely collected patient data indeed 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 consented for research purposes (such as the national mental health surveys, longer-term follow-ups of clinical trials and cohort studies )? *The review suggests that explicitly consented data can be treated differently, but our experience is that the most restrictive practice quickly becomes the default for all health-related data.* This was demonstrated by the dramatic reduction in downloads and uses of the Adult Psychiatry Morbidity Study 2014 following restrictions, relative to its 2007 counterpart, despite it being explicitly consented research data (Ford et al. 2021): f our years following the 2007 release, data had been accessed 165 times, producing 30 published papers. By the equivalent time point following the 2014 release there were *just 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 (Jones et al, 2022). Special consideration is required for engagement with vulnerable groups, especially those who may experience barriers to conventional consent procedures, not least because they may be more readily identifiable. We regularly work with datasets collected for the explicit purpose of tracking mental health. Recent work 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 (MindKind Final Report, 2022). We can increase this trust by involving them in decision making and priority setting with regard to their data. This further highlights both the time and longitudinal sensitivity of any data concerning children and adolescents ; a good enough contextual understanding of the societal and individual developmental context in which the data were created as well as the developmental plasticity of any child and adolescent clinical data is absolutely essential in making this approach work for this population.

***Trusted Research Environments (TREs)***

A key recommendation of the report is the centralisation of TREs so that a small number would curate all NHS data. Our current experience is of substantial delays in accessing NHS controlled data, with TREs imposing additional bottlenecks 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 only available in late 2020, while the pandemic follow-ups collected in 2020 and 2021 are *still* not available at the time of writing. Almost three years into the pandemic and associated restrictions, we still do not know its full impact on children and young people’s mental health. We fear that the proposed centralisation 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 TREs then who will bear this cost and how will access be prioritised? TREs will need sufficient capacity (personnel and processing) to undertake different analyses, using a range of different coding languages. What will happen to existing TREs? What will the waiting time be before the TRE and application process are ready? What happens in the meantime to datasets currently within NHS digital that are barely accessed?

In summary, we welcome many of the Goldacre Review’s recommendations, but are concerned about potential interpretations of this landmark review. Child and adolescent mental health is particularly dynamic, complex and context dependent, as has been amply demonstrated by the Covid-19 pandemic. Timely access to *current* data is essential if we are to answer the most urgent questions and provide responsive, evidence based care for children and young people.

**References**

Goldacre, B & Morley, J. (2022). 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.

Ford, T., Mansfield, K. L., Markham, S., McManus, S., John, A., O'Reilly, D., ... & Shenow, S. (2021). The challenges and opportunities of mental health data sharing in the UK. *The Lancet Digital Health*, *3*(6), e333-e336.

Jones, L. A., Nelder, J. R., Fryer, J. M., Alsop, P. H., Geary, M. R., Prince, M., & Cardinal, R. N. (2022). 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, 12(4), e057579.

MindKind Final Report (2022), Sage Bionetworks, <https://cms.wellcome.org/sites/default/files/2022-09/mindkind-final-report-2022.pdf>