Emerging opportunities and challenges for clinical prediction models in psychiatry

Dominic Oliver[1,2], Gabrielle Beaudry[3,4,5], Fiona Coutts[6,7], Seena Fazel[1,2]

1 Department of Psychiatry, University of Oxford, Oxford, UK;

2 NIHR Oxford Health Biomedical Research Centre, Oxford, UK;

3 Faculty of Medicine and Health Sciences, McGill University, Montréal, Canada;

4 Centre for Outcomes Research and Evaluation, Research Institute of the McGill University Health Centre, Montréal, Canada

5 Centre for Clinical Brain Sciences, University of Edinburgh, UK;

6 Centre for Addiction and Mental Health, Toronto, Ontario, Canada

7 Department of Psychosis Studies, Institute of Psychiatry, Psychology & Neuroscience, King's College London, London, UK

Word count: 1,023/1,000

References: 9/9

Corresponding author: Dr Dominic Oliver (dominic.oliver@psych.ox.ac.uk)

POWIC Building, University of Oxford, Warneford Lane, Headington, Oxford, OX4 1PD, United Kingdom

Clinical prediction models use individual-level patient data to inform clinical decisions across medicine. However, few are currently used in psychiatry. Here, we examine the opportunities and challenges presented by the new generation of clinical prediction models in psychiatry and consider what it would mean for a model to be “good enough” for clinical use.

Opportunity one: prognosis

Prognostic models estimate individual-level risk of adverse outcomes to inform clinical decision making. Several such models have been developed and validated in psychiatry, including those estimating risk of psychosis onset, relapse, metabolic syndrome and suicide (1). Reliable risk estimation for these outcomes could inform follow-up intensity, continuity of care and collaborative safety planning. However, despite the potential of prognostic models for suicide, current UK guidelines for suicide risk assessment are unusually definitive, strongly recommending against the use of structured prediction tools (2,3). This contrasts sharply with other areas of medicine, such as the QRISK for the prediction of cardiovascular events, where imperfect tools are continuously improved and refined rather than abandoned.

Opportunity two: treatment matching

As “one size fits all” treatment approaches in psychiatry are increasingly questioned, prediction models may help estimate differential treatment response and align care with patient preferences (4). Even modest improvements in matching individuals to more effective or acceptable interventions could yield substantial benefits at the population level. For example, the most recent Cochrane review supports the use of cognitive behavioural therapy (CBT) for self-harm (5), but the recommendation in the Feature that everyone who self-harms should receive the same management (6) does not account for real-world constraints on service capacity. Clinical prediction models could ease this burden by recommending CBT to those predicted to benefit most.

Challenge one: low base rates

Concerns are often raised about low base rates of psychiatric outcomes but advances in data availability and methodology have largely mitigated this issue. Large population-based datasets have provided the sample sizes needed to model

uncommon but clinically important outcomes, accompanied by improved handling of missing data, modern modelling techniques and adherence to reporting standards (e.g. Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis [TRIPOD+AI]). Consequently, research in psychiatry is increasingly characterised by robust models that can be transparently evaluated, externally validated and implemented.

Challenge two: generalisability

Concerns about generalisability are frequently raised, particularly when models are developed using data from a single country or healthcare system. As long as high-quality methods are used, a model demonstrating adequate performance in its development setting sets a strong foundation for validation elsewhere. For example, differences in national suicide rates do not undermine a model's ability to estimate relative risk (the central function of prognostic models). Absolute risks can be recalibrated for new populations when required, as is standard practice across medicine (7). Decades of evidence demonstrate the generalisability of well-constructed prediction models across settings (1,8).

Challenge three: mechanisms and causality

Psychiatric models are often criticised for predicting outcomes without advancing understanding of underlying mechanisms. Models closest to clinical use typically rely on readily available, clinically relevant predictors, rather than data directly reflecting aetiopathology. Biomarker-based models may better capture biological processes, but remain rarely validated in psychiatry (8), limiting their clinical applicability and highlighting the trade-off between mechanistic ambition and real-world feasibility.

Causal artificial intelligence has been proposed as a means of modelling individual patient trajectories under different hypothetical scenarios by generating "counterfactual" predictions. While conceptually promising, evaluating such models is challenging, as counterfactual outcomes cannot be observed.

Importantly, the absence of biological or causal interpretation does not preclude clinical usefulness. While simpler, more feasible models using routine clinical data might not fully capture the complexities of mental disorders, many widely used prediction models across medicine inform decision making without explaining

disease mechanisms. Conflating prediction with explanation risks holding clinical prediction models in psychiatry to a standard not applied elsewhere in medicine.

Challenge four: what is “good enough”?

Whether a prediction model is “good enough” is not a single challenge. Rather it comprises several distinct elements: adequate and fair predictive performance, stakeholder acceptability, system-level benefit and a clearly defined clinical decision.

First, a model must demonstrate adequate performance, namely discrimination, calibration, clinical utility and fairness (Panel 1). Performance should be evaluated using robust external validation in independent datasets representative of intended use, and should be compared against real-world practice rather than an idealised notion of clinical judgement.

Second, a model must be acceptable to relevant stakeholders. Even models with perfect accuracy have limited value if clinicians are unwilling to adopt them. Patients value models that are accurate, transparent, secure and empowering (9,10). Patient and clinician input can therefore be useful to inform effective model development and implementation. For example, small gains in performance may not justify substantial increases in clinical burden or assessment time.

Third, a model must show potential benefit at a system level and consider the downstream effects of implementation. For example, a model detecting risk for psychosis may increase referrals to early detection services. Whether such effects are beneficial depends on the available resources. Cost-effectiveness is therefore critical: a model that systematically recommends more expensive interventions with limited incremental benefit may be inappropriate for widespread adoption.

Fourth, a model must relate to a clear clinical decision point. Ideally, this would correspond to a specific treatment decision, analogous to QRISK and statin prescription. However, psychiatry lacks analogues for many outcomes. Beyond treatment selection, models may inform referral pathways, support collaborative risk management, improve communication and consistency within and between services, and raise the baseline level of expertise in settings with variable clinical experience, including primary care and junior clinical staff (4). In many cases, the key decision is simply whether to intervene at all, where inaction is often the default.

Conclusion

Advances in data availability and methodology have created real opportunities for clinical prediction models to improve prognosis and treatment selection in psychiatry. Assessing risk in psychiatry depends on balancing clinical judgement with data: this will require promoting new research and openness to change current approaches to risk assessment. As in other fields of medicine, the key question should not be whether psychiatric prediction models are perfect, but whether they are good enough to improve upon current practice.

Acknowledgements

DO and SF are supported by the NIHR Oxford Health Biomedical Research Centre. The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health and Social Care.

Author Contributions

All authors drafted the manuscript and revised the content.

Conflicts of Interest

No authors report any conflicts of interest.

Panel 1 Definitions of key considerations for clinical prediction model performance: discrimination, calibration, clinical utility and fairness. Adapted from (1).

Discrimination	Calibration	Clinical utility	Fairness
• Ability to separate individuals with and without the outcome. • Typically measured by Harrell's C: the proportion of randomly selected cases who receive a higher risk score than randomly selected non-cases. • There is no strict cut-off for discrimination that is good enough, as this depends on the context of the clinical use case and on the available alternatives.	• Clinical prediction models should not only discriminate well but also provide accurate risk estimates, which are assessed by calibration. • Calibration assesses the relationship between predicted probabilities and observed risk proportions. • Miscalibrated models result in over- or under-estimation of risk. i.e. a model may discriminate well between those with and without the outcome, but probabilistic estimates of absolute risk may be systematically off target, which is important if this is being communicated to the patient or used for a clinical decision. • A miscalibrated model can lead to patients being misinformed about their true risk and may also have treatment implications: a patient may be recommended an unnecessary intervention (over-estimation) or not receive care that is needed (under-estimation). • Therefore, calibration is essential to prevent potential harm caused by the overestimation or underestimation of risk.	• Measured by net benefit. • Allows us to weigh the benefits and costs associated with using the model. • Net benefit is compared to reference strategies (e.g. treat all, treat none or the current gold standard approach) across different assumptions of the number-needed-to-test (e.g. a number-needed-to-test/treat of 10 equates to an odds of 1:9, indicating that missing the outcome of interest once is nine times worse than an unnecessary intervention). • It therefore considers a range of preferences for whether you are more worried about missing the outcome or giving an unnecessary intervention to evaluate potential clinical benefit.	• Clinical prediction models should not exacerbate healthcare inequalities. • Fairness can be tested through subgroup analyses, assessing model performance in potentially under-represented sub-populations (sex, ethnicity etc.). • Can also be assessed through metrics like demographic parity (proportion of individuals assigned a positive prediction is similar across protected groups, regardless of true outcome status) or equalised odds (sensitivity and specificity are similar across protected groups), though these are based on classification performance so are dependent on the definition of cut-offs.

References

1. Oliver D, Arribas M, Perry BI, Whiting D, Blackman G, Krakowski K, et al. Using electronic health records to facilitate precision psychiatry. Biological Psychiatry. 2024 Feb;S0006322324011077.

2. NICE. Self-harm: assessment, management and preventing recurrence [Internet]. 2022. Available from: https://www.nice.org.uk/guidance/ng225/chapter/Recommendations#risk-assessment-tools-and-scales

3. Seyedsalehi A, Fazel S. Clinical guidelines on self-harm and suicide prevention: taking uncertainty into account in the evidence base. BMJ Ment Health. 2026 Feb;29(1):e302420.

4. Seyedsalehi A, Scola G, Fazel S. Precision psychiatry: thinking beyond simple prediction models – enhancing causal predictions: commentary, Seyedsalehi et al. Br J Psychiatry. 2025 Nov 13;1–2.

5. Witt KG, Hetrick SE, Rajaram G, Hazell P, Taylor Salisbury TL, Townsend E, et al. Psychosocial interventions for self-harm in adults. Cochrane Common Mental Disorders Group, editor. Cochrane Database of Systematic Reviews [Internet]. 2021 Apr 22 [cited 2026 Jan 5];2021(4). Available from: http://doi.wiley.com/10.1002/14651858.CD013668.pub2

6. Cockburn A, Large M. A clinician's guide to probabilistic suicide risk prediction tools: cautions and pitfalls. Br J Psychiatry. 2025 Nov 5;1–5.

7. Steyerberg EW, Vergouwe Y. Towards better clinical prediction models: seven steps for development and an ABCD for validation. European Heart Journal. 2014 Aug;35(29):1925–31.

8. Salazar de Pablo G, Studerus E, Vaquerizo-Serrano J, Irving J, Catalan A, Oliver D, et al. Implementing Precision Psychiatry: A Systematic Review of Individualized Prediction Models for Clinical Practice. Schizophrenia Bulletin. 2021 Mar 16;47(2):284–97.

9. Kéri P, White LA, Oliver D, Fusar-Poli P. Empowering Experts by Experience to Guide Ethical Precision Psychiatry. Biological Psychiatry. 2023 Dec;S0006322323017444.

10. Whiting D, Glogowska M, Mallett S, Maughan D, Lennox B, Fazel S. Use of a violence risk prediction tool (Oxford Mental Illness and Violence) in early intervention in psychosis services: mixed methods study of acceptability, feasibility and clinical role. Br J Psychiatry. 2025 Mar 20;1–10.