

Editorial

Towards personalised predictive psychiatry in clinical practice: an ethical perspective

Natalie Lane and Matthew Broome



Summary

Personalised prediction models promise to enhance the speed, accuracy and objectivity of clinical decision-making in psychiatry in the near future. This editorial elucidates key ethical issues at stake in the real-world implementation of prediction models and sets out practical recommendations to begin to address these.

Keywords

Ethics; philosophy; information technologies; patients; stigma and discrimination.

Copyright and usage

© The Author(s), 2022. Published by Cambridge University Press on behalf of the Royal College of Psychiatrists.

Natalie Lane (pictured) is a Year 3 Core Psychiatry Trainee with NHS Greater Glasgow & Clyde and an MA Bioethics and Society graduate from King's College London, UK.
Matthew Broome is Chair in Psychiatry and Youth Mental Health and Director of the Institute for Mental Health at the University of Birmingham, UK.

Uncertainty regarding patient diagnosis, prognosis and optimal management are ever-present challenges in clinical practice. In psychiatry in particular, clinical decision-making has been critiqued as overly subjective. Personalised prediction models represent a novel approach, whereby statistical and machine learning models detect patterns in big data repositories of sociodemographic, clinical, behavioural, cognitive and biological (e.g. neuroimaging and genetic) information to generate probabilistic, individualised risk estimates of a particular outcome occurring. In clinical psychiatric practice, this could facilitate targeting of the type and duration of interventions offered and thus improve patient outcomes. It is crucial that the myriad ethical implications of the use of prediction models are well understood in advance of their clinical implementation and that scientific and technical advances do not leave such considerations in their wake.

Background

The clinical use of personalised prediction models in psychiatry is becoming increasingly feasible. Much research has focused on their use in early intervention in psychosis (in particular, to predict the likelihood of transition to psychosis in individuals at increased risk or to predict outcomes in first-episode psychosis), but it is envisaged that they could be applied to a wide range of psychiatric illnesses occurring across the lifespan.¹ Published examples include models predicting the individual likelihood of diagnosis of autism spectrum disorder; of the onset of bipolar disorders in individuals at familial risk; and of treatment response to antidepressants in major depression.²

Practical considerations

Why do personalised prediction models warrant ethical scrutiny? After all, we routinely make informed estimates of patients' likely outcomes in clinical practice, based on the available evidence and

professional experience. However, personalised prediction models differ from these conventional means in two salient ways: first, their reported predictive accuracy may overstate that achievable in clinical practice; and second, their workings often lack transparency.

In brief, it has been widely observed that personalised prediction models tend to perform more poorly when applied to a different population from the one in which they are derived. However, most personalised prediction models under development for use in psychiatry have not yet undergone external validation using an independent sample or had their accuracy prospectively assessed in a real-world patient population. Therefore, the reported predictive accuracy of currently available models is likely to overestimate what they can achieve in clinical psychiatric practice.²

Moreover, current personalised prediction models often lack transparency in terms of not supplying the user with an explanation of how or why a particular risk estimate is made. In addition, the complexity of statistical and machine learning methods makes it extremely difficult (and often impossible) for the clinician, patient or relative to independently decipher the prediction model's workings.³

Ethical implications

Issues of accuracy and transparency are of ethical importance as they determine the degree of influence that personalised predictions are likely to exert on clinical decision-making. In effect, if prediction models claim a heightened level of predictive accuracy, they are likely to be ascribed greater weight in determining clinical recommendations. Moreover, if clinicians cannot access how or why a particular risk estimate has been made, they are unable to properly evaluate the relevance of it in guiding clinical management. This inability to effectively question risk estimates, combined with their high level of reported accuracy, may leave clinicians feeling pressurised into adopting an automaticity of decision-making, whereby they act in accordance with prediction model outcomes by default.³

Within psychiatry, this effect could be particularly problematic in high-risk scenarios (e.g. crisis management and in-patient discharge). Typically, such situations are approached through shared decision-making and safety planning with the patient and multidisciplinary colleagues, with preference given to mitigating risk in the least restrictive manner. However, if a prediction model confidently

estimates a high chance of a poor outcome, this could be taken to outweigh all other considerations, not least owing to the challenge of justifying acting contrary to this information and fear of medico-legal repercussions if one were to do so. Moreover, the lack of transparency of prediction models renders clinicians ill-equipped to discuss in detail with patients the rationale behind clinical recommendations, thus jeopardising their ability to provide valid informed consent. Therefore, instead of being a useful decision-making adjunct, prediction models could, in some instances, subvert the roles of the healthcare team and patient in shared decision-making, encourage defensive medicine at the expense of patient autonomy and undermine the principle of informed consent.³

Prediction models could further disempower patients by undermining their sense of agency (the belief that one can shape one's own life).⁴ This may occur if patients misconstrue personalised predictions to infer that significant outcomes (e.g. onset of mental illness, relapse or recovery) are entirely foreseeable and thus entirely predetermined and beyond their ability to influence. This could leave patients feeling hopeless in the face of negative predictions or disengaged in response to positive ones. In a clinical context, patients may be less motivated to play an active role in shared decision-making and treatment, to the detriment of clinical outcomes.⁴ More broadly, patients (and relatives) may impose limitations on their own aspirations for the future, thus curtailing their life experiences. Further, discrimination from external sources could occur if third parties gain access to risk estimates, for example reduced opportunities for employment in certain sectors (e.g. the military) and difficulty obtaining insurance or travel visas.⁵

Last, the clinical use of personalised prediction models risks perpetuating inequities in psychiatric care. Specifically, the real-world accuracy of predictions is likely to be further compromised in individuals from minority groups, owing to underrepresentation in the derivation sample on which the prediction model is trained and tested. For example, if a model is derived from a predominantly young adult, White male population (as is typical of research cohorts in the UK), its risk predictions may be less accurate for individuals from ethnic minorities, females or those at the extremes of age, thus disadvantaging these groups.³

Ethical recommendations

It is imperative that the above ethical considerations are taken into account in efforts to implement personalised prediction models in clinical psychiatry. In support of this, we suggest the following practical ethical recommendations and highlight avenues for future research.

First, a realistic report of the predictive accuracy achievable by a prediction model in clinical practice should be made available to users, to allow them to make an informed judgement on the influence that risk estimates should have on decision-making. This necessitates that models undergo external validation at the research stage, and prospective assessment of real-world accuracy in a clinical context as part of implementation research, prior to widespread roll-out. Care must be taken to ensure that sample populations at each stage are diverse and inclusive, to avoid creating inequities in predictive accuracy that perpetuate disadvantage for underrepresented groups.

Second, the interpretability of prediction models should be maximised before clinical implementation. Developers should prioritise simpler models where doing so will not compromise clinical utility, for example those using basic clinical data rather than complex biomarkers, and standard statistical approaches rather

than sophisticated machine learning algorithms. Initial indications from psychiatric research suggest that simpler models display comparable predictive accuracy to more complex alternatives.² In addition, materials should be designed to educate clinicians on how the prediction model operates. In line with this, future research should investigate what level of detail healthcare professionals deem useful. This is likely to include, at a minimum, information on the underlying data-sets and relevant variables used, broadly how the results are derived, how outcomes are defined and, crucially, the reliability and limitations of predictions.

Third, patients must be supported to meaningfully understand their personalised predictions. This is contingent on the previous point of clinicians being adequately informed to educate their patients (and by doing so, facilitate informed consent for interventions based on risk estimates). However, it necessitates not just that clinicians know what information to impart, but also how to do so in a manner that instils hope and maximises agency.⁵ It should be made explicit that predictions are probabilistic estimates, as opposed to binary indicators that an outcome will or will not occur. In addition, it should be explained that an outcome as defined by the model (e.g. occurrence of a condition based on diagnostic criteria) does not equate to a uniform lived experience. Last, it must be emphasised that the individual can often influence outcomes, for example by addressing modifiable risk factors and engaging with treatment and support. Future empirical research involving patients is warranted to investigate how best to disclose predictive information in an empowering fashion.

Conclusions

Personalised prediction models in psychiatry hold the potential to enhance clinical decision-making and significantly benefit patients in the near future by expediting early intervention and optimal management. However, it is crucial that the ethical implications of prediction model use are well understood ahead of clinical implementation, to ensure that this occurs in an ethically justified manner that maximises these benefits and minimises inadvertent harms. This editorial opens this discussion by considering the impact of prediction model use on shared decision-making, informed consent, patient autonomy and agency, and pre-existing inequities in psychiatric care. It is important that clinicians have an accurate understanding of prediction models and their limitations and can explain results to patients in a manner that promotes informed consent, instils hope and maximises agency. Empirical study of the views of key stakeholders, including patients, relatives and healthcare professionals, will help achieve these aims. Ultimately, proactive engagement with the ethical ramifications of scientific and technical innovation in this field is essential to ensure that the great promise of personalised prediction models in psychiatry is fully realised in clinical practice.

Natalie Lane , Department of Psychiatry, Gartnavel Royal Hospital, NHS Greater Glasgow & Clyde, Glasgow, UK; **Matthew Broome** , Institute for Mental Health, University of Birmingham, UK

Correspondence: Natalie Lane. Email: natalie.lane3@ggc.scot.nhs.uk

First received 13 Dec 2021, final revision 13 Feb 2022, accepted 16 Feb 2022

Data availability

Data availability is not applicable to this article as no new data were created or analysed in its preparation.

Author contributions

N.L. conducted the ethical analysis, planned and prepared the initial draft of the manuscript and was primarily responsible for the subsequent revision of the manuscript. M.B. reviewed and edited the initial manuscript and subsequent revised versions, and provided guidance throughout.

Funding

This research received no specific grant from any funding agency, commercial or not-for-profit sectors.

Declaration of interest

M.B. is Deputy Editor of the *BJPsych* and did not take part in the review or decision-making process of this paper.

References

- 1 Leighton SP, Upthegrove R, Krishnadas R, Benros ME, Broome MR, Gkoutos GV, et al. Development and validation of multivariable prediction models of remission, recovery, and quality of life outcomes in people with first episode psychosis: a machine learning approach. *Lancet Digit Health* 2019; **1**(6): e261–70.
- 2 de Pablo GS, Studerus E, Vaquerizo-Serrano J, Irving J, Catalan A, et al. Implementing precision psychiatry: a systematic review of individualized prediction models for clinical practice. *Schizophr Bull* 2021; **47**: 284–97.
- 3 Grote T, Berens P. On the ethics of algorithmic decision-making in healthcare. *J Med Ethics* 2020; **46**: 205–11.
- 4 Houlders JW, Bortolotti L, Broome MR. Threats to epistemic agency in young people with unusual experiences and beliefs. *Synthese* 2021; **199**: 7689–704.
- 5 Lane NM, Hunter SA, Lawrie SM. The benefit of foresight? An ethical evaluation of predictive testing for psychosis in clinical practice. *Neuroimage Clin* 2020; **26**: 102228.