BMJ Health & Care Informatics

Patient perspective on predictive models in healthcare: translation into practice, ethical implications and limitations?

Sarah Markham 💿

To cite: Markham S. Patient perspective on predictive models in healthcare: translation into practice, ethical implications and limitations? *BMJ Health Care Inform* 2025;**32**:e101153. doi:10.1136/ bmjhci-2024-101153

Received 29 May 2024 Accepted 02 January 2025



© Author(s) (or their employer(s)) 2025. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ Group.

King's College London, London, UK

Correspondence to Dr Sarah Markham; sarah.markham@kcl.ac.uk

ABSTRACT

In this perspective article, we consider the use of predictive models in healthcare and associated challenges. We will argue that patients can play a valuable role in supporting the safe and practicable embedding of such tools and provide some examples.

INTRODUCTION

The National Health Service currently uses predictive tools to help decide how to treat patients.¹ Awareness of the concept and use of predictive models in healthcare have been increasing in the media during the last decade.² Such predictive tools hold the promise of potentially being able to predict our future health needs, including preventative needs, thereby enabling clinicians to target and deliver care in a more timely and equitable manner and hopefully leading to healthier and happier us.³ However, research on predictive models in healthcare has highlighted both their benefits and limitations. Examples of the benefits that predictive models can bring to healthcare include a machine learning model which has been successfully developed and deployed to predict sepsis in hospitalised patients before clinical signs appeared.⁴ Early sepsis detection is crucial because timely treatment can prevent organ failure and death. Through the analysis of a range of patient data, the model was able to predict sepsis up to 12 hours before traditional clinical detection methods, significantly improving the time to treatment.⁴ The model demonstrated that predictive analytics could lead to early interventions that reduce mortality rates and improve patient outcomes. However, predictive models are not always beneficent; there is also evidence of bias in predictive models for healthcare needs, for example, in 2019 a study examined an algorithm used to predict which patients would benefit from extra healthcare services. It was found that the algorithm used historical healthcare spending data, which

disproportionately favoured white patients, as they had historically received more care than black patients, leading to underestimating the health needs of black patients.⁵ Another limitation of predictive tools is their limited generalisability across populations. For example, in 2018, a study found that certain predictive tools for hospital readmissions performed less well for minoritised populations predominantly due to differences in healthcare access, and treatment patterns, in addition to the social determinants of health.⁶ We will argue that patients through public and patient involvement (PPI) can play a valuable role in mitigating these potential harms and supporting the safe and practicable embedding of predictive models.

BACKGROUND

As a long-term patient with several health issues and a mathematician, I am dually interested in the phenomenon of clinical predictive models and how their use may be integrated in healthcare and public health.⁴ I am also an advocate for the value of PPI in the development and use of such tools in addition to wider discourse and decision-making regarding the ethics, commissioning and regulation of these technologies. For the last 6 years, I have cochaired a Biostatistics PPI Group at King's College London. We meet regularly online to inform the design of quantitative clinical studies including the development of predictive models.

DEVELOPMENT OF PREDICTION TOOLS

Risk prediction tools are becoming increasingly en vogue.⁷ Examples of predictive tools used in healthcare include those used for predicting disease risk, readmission rates, patient outcomes and optimising treatment plans. For instance, Google Deepmind developed a model to identify whether someone is at risk of experiencing acute kidney injury.⁸



These tools are based on models which have been trained on healthcare data to predict various potential harms such as the possibility of a patient, as an individual or as someone possessing certain characteristics, developing a certain health condition, or experiencing certain treatment outcomes.⁹ Theoretical predictive models are created through a process of data collection, model building, exploratory data analysis, evaluation and deployment as shown in the diagram below. Data can be drawn from healthcare in significant amounts. Healthcare providers record and aggregate data from diverse sources: electronic health records (EHR), laboratory test results, medical imaging and routine administrative data. Before the data can be used to develop a predictive model, they require preprocessing: cleaning and management of missing data, outliers and data inconsistencies.¹⁰ In essence predictive model development involves the identification and selection of relevant features that are

relevant to optimising the predictive accuracy of the resulting model. Practices vary and there are multiple methodologies and techniques which may be deployed; the specific methods used being determined by the nature of the predictive problem at hand and the characteristics of the data available.⁷ These methods are drawn from data science, statistics and artificial intelligence (AI) to make forecasts about future outcomes based on historical data, including traditional statistical methods such as regression analysis and survival analysis; supervised machine learning techniques such as decision trees and random forests; time series analysis which is crucial for making healthcare predictions that involve sequences of data points over time (eg, vital signs, blood glucose levels, patient monitoring data) and natural language processing (NLP) which is used to analyse unstructured text data from sources like clinical notes, medical literature and EHR.⁶ Explainability and interpretability techniques are increasingly being called for in healthcare to understand how predictive models come to their outcomes. This is recognised as being crucial for trust and adoption within practice (figure 1).¹



Figure 1 Building a predictive model. https://www. analyticsvidhya.com/blog/2022/09/the-6-steps-of-predictiveanalytics/

PPI IN THE DEVELOPMENT OF RISK PREDICTION TOOLS

PPI is recognised to increase the quality and outcomes of research in healthcare.¹¹ Integrating PPI in all stages of predictive model development is essential to ensuring that models are accurate, understandable, fair and aligned with patient needs and realities. First, PPI can be of value in prioritising which health risks and other uncertainties would merit the development of better prediction tools. Involving patients in determining which data should be collected and how can help reduce the likelihood of types of data essential to representing the patient ground truth being omitted. For instance, patients can be invited to participate in surveys or focus groups or surveys to help identify the most relevant features for predictive models. Patients can be involved alongside academics and clinicians within research studies in codesigning predictive models through recruitment to patient advisory boards. They can help refine the questions that predictive models are trying to answer and ensure that the model is designed with patient care needs in mind. PPI contributors can also contribute insight into which risk factors to focus on in development and how best to safely use anonymised patient data in the development of predictive tools. An example of such involvement is found in a research study which developed a machine learning model trained on observational data from electronic healthcare records to predict mortality in patients suffering from schizophrenia.¹² Patients were recruited to join a Research Advisory Group (RAG) and although the central hypothesis of the project was defined by the researchers, the need to develop an explainable AI predictive model emerged from discussions with the members of the RAG.¹³ The members of the RAG remained involved throughout the study contributing to decisions and interpretations of findings via regular online meetings and coproduced questions in surveys that were sent out to patients. In order to empower RAG members to make informed contributions, the researchers used various tools, for instance, the Teachable Machine, which enabled patients to interact with and develop an understanding of neural networks by using a web browser.¹⁴ Patient-reported outcomes, such as selfreported health status, symptoms and psychoemotional well-being, are often needed to inform the development of predictive models. Patients are often best placed to advise on which outcomes and outcome measures should be used. Evaluating how well the model aligns with their lived experience and whether its predictions resonate with their experiences can help identify where the model may be inaccurate or insufficient. Patients can also provide feedback on whether the model's outputs are understandable and actionable for them. In order to empower patients to make informed decisions, they need to be involved in communicating how a predictive tool works, what data it uses, what its outputs mean and how it can impact care.

BIAS IN PREDICTION TOOLS

Predictive models are statistically or computationally trained on the preprocessed data, with the overall aim of optimising the quality of the predictions. All parts of this process are vulnerable to bias and other forms of error, for instance due to missing data and how it is management and representativeness of the training data.¹⁵ Hence, the need for model validation, that is, checking how well the model works by testing it on different sets of data, and ongoing evaluation. Model validation involves the assessment of model performance using separate datasets not used during the training phase and techniques such as cross-validation.¹⁶ Appropriate metrics are then applied to the output of the predictive model to evaluate the sensitivity, specificity, precision and other indicators of accuracy.¹⁷ Although predictive models have the potential to help improve patient outcomes and optimise resource allocation, if not appropriately developed or implemented, they can unintentionally discriminate against certain patient groups, leading to inequitable care and exacerbating health inequalities. This often occurs when the data on which the predictive models are trained isn't representative of all the groups to which the tool may be applied. Examples of patient groups that can be discriminated against in predictive models in healthcare include racial and minority ethnic groups, gender and gender minority groups (especially transgender patients if gender identity is not appropriately recognised or accounted for, especially in terms of healthcare needs like hormone therapy), and patients with rare diseases as the training of predictive models often relies on large datasets, which may not be available for rare diseases. In the case of racial and ethnic minorities, predictive models may deploy historical data from populations primarily composed of certain racial or ethnic groups. If these models are not robustly adjusted for racial or ethnic differences, they may underestimate risk for minority populations, leading to inequalities in care. For instance, in 2019, a study found that an algorithm used to predict healthcare needs for patients discriminated against black patients by underestimating their risk of needing care, because the model was trained primarily on data from white patients.⁵

PPI IN PROTECTING AGAINST BIAS IN RISK ASSESSMENT TOOLS AND PROMOTING ETHICAL CONSIDERATIONS

PPI in predictive model development is key to ensuring that the model doesn't inadvertently discriminate against certain groups. Patients from diverse backgrounds can help identify potential biases in the model and provide feedback on how the model might affect different demographic groups and how to resolve this.¹⁸ A further requirement of predictive models is that they adhere to relevant healthcare regulations, which in turn may be revised or replaced over time.⁶ This leads to the need for further remodelling or replacement of the predictive tool. Thus, the development, translation and maintenance of predictive tools in healthcare are a complex and potentially resource heavy process.

A further concern is the ethics of such tools and their use. Ethical concerns include issues relating to patient

privacy, consent for their data to be input into such tools and ensuring that the model does not introduce biases that may disproportionately impact certain patient cohorts. Sadly, the risk remains that such predictive tools could cause unintentional harm. In the context of such tools being used to guide clinical decision-making, this may affect the quality of care delivered to patients and potentially widen health inequalities. Here, it is crucial that PPI plays a substantive role in delineating better understanding of good practice and how best to safeguard patients against harm.

Other ethical concerns include the privacy and security of patient data. The data used to develop predictive models may include sensitive and confidential data, hence the need for effective data anonymisation and pseudomisation to prevent inappropriate access or misuse. Furthermore, there is a real need for patients to be comprehensively informed about how their data may be used to develop predictive models and that their informed consent is given before their data are input into predictive tools and potentially used to inform their care and treatment. Furthermore, patients often experience unique aspects of the healthcare process and can help to enhance situational awareness of whether a given predictive tool would be of practicable value and fit into specific diagnostic, care and treatment contexts.

DEPLOYMENT OF PREDICTIVE TOOLS IN HEALTHCARE

The embedding of predictive modelling into healthcare systems isn't necessarily straightforward.¹ Very often a user interface which clinicians and other professionals find accessible, and which fits into clinical workflows needs to be developed. Within this, the outputs of such predictive models should be meaningfully interpretable and clinically deployable by healthcare professionals.¹⁹ Even if a predictive model passes all these tests, regular monitoring is required to measure the model's performance in realworld settings.⁶ Predictive models and their value are intrinsically tied to the quality and relevance of the data used to train them and the state of individuals and the wider world such data pertain to.¹⁹ Complex modelling techniques applied to noisy data may manifest overfitting meaning that they perform less well on different data sets.¹⁹ Furthermore, the characteristics of individuals and the environments which influence their behaviours and health, in addition to healthcare policies and practices, are not static and evolve over time.²⁰ Inevitably predictive models will manifest the consequences of such concept drift and will need to be retrained on more recent data and restructured to maintain their predictive accuracy.²⁰

Another key issue is the transparency and understandability of these models.²¹ Some predictive models, typically those derived using machine learning, can be metaphorical 'black boxes' and it can be difficult if not impossible to determine how given the data to which they are applied, how they derive their outputs.²¹ This is of special concern in treating complex or high-risk cases where it is vital that clinicians make the right clinical treatment decisions. Another complex issue is how to determine who is legally responsible for any harm caused to patients due to clinical decision-making informed by the outputs of predictive tools.²² It is important to establish accountability mechanisms and clarify the responsibilities of healthcare professionals when using these models. As far as I am aware, no clear solutions have been proposed and this issue may significantly impact the trust both patients and clinicians are able to place in such tools and their future use in healthcare. The degree to which predictive tools can be incorporated into existing healthcare systems and workflows presents a further challenge. These multiple and potentially interacting dimensions of complexity and indeed risk necessitate the meaningful and informed consultation and collaboration between patients, clinicians, methodologists and other stakeholders to create standards and practices regarding the development and deployment of predictive tools in healthcare which optimise the health and well-being of patients. At the end of the day, it may be the disadvantages more than the benefits that predictive tools may bring that will make or break their future in healthcare.

FUTURE DIRECTIONS FOR PREDICTIVE MODELS

The future of predictive models in healthcare will be driven by advancements in data science, statistical modelling and AI including machine learning, and hopefully supported and strengthened by embedded and robust PPI. Personalised medicine promises more individually tailored care and treatment and to deliver on this, predictive models need to move beyond general populationbased designs to provide more person-specific treatment recommendations and other information, based on an individual's genotype, phenotype, lifestyle and aspirations; medical includes predicting how an individual patient will respond to specific medications or therapies.

Prevention is universally recognised to be better than cure and there is significant scope for predictive models to play a greater role in early disease identification and the prediction of disease onset. Predictive models need to become more adept at identifying early signs of diseases, preferably before symptoms manifest. This could be actualised by using data from EHR, imaging and biomarkers, to predict the likelihood of conditions arising years in advance so that preventative measures can be taken. Wearable devices could be deployed at scale to collect individual-level data, to be analysed by predictive models to identify anomalies or early symptoms, leading to more timely intervention.

Predictive models will be integrated into clinical decision support systems to help doctors make more accurate diagnoses and provide better advice to patients when engaging in shared decision-making regarding future treatment and care. Predictive models will be developed to optimise service provision and operations by forecasting patient demand and resource utilisation. This will help healthcare systems improve efficiency, reduce wait times and manage resources more effectively. Similarly, predictive models will be used to forecast the demand for medications, equipment and staffing needs, especially in up-tempo clinical contexts such as emergency departments and pandemics.

Future predictive models may also be designed to better incorporate data related to social determinants of health to better analyse, understand and resolve health inequalities. As predictive models become integral to healthcare decision-making, there will be an increased emphasis on developing explainable systems that enable patients and clinicians to better understand how predictions are made, thereby promoting proportionate trust and accountability. In the context of patient empowerment and autonomy, hopefully future predictive models will aid patients in predicting their own health patterns and needs, potentially providing information to enable individuals to make more informed decisions about their lifestyle behaviours, treatment and care. PPI will no doubt play a key part in optimising the design, delivery and impact of future innovations in predictive modelling in healthcare.

X Sarah Markham @DrSMarkham

Funding The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests None declared.

Patient consent for publication Not applicable.

Ethics approval Not applicable.

Provenance and peer review Commissioned; externally peer reviewed.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/.

ORCID iD

Sarah Markham http://orcid.org/0000-0002-8755-5935

REFERENCES

- 1 NHS. Case study: AI tool improving outcomes for patients by forecasting A&E admissions. 2023. Available: https://www.england. nhs.uk/long-read/case-study-ai-tool-improving-outcomes-forpatients-by-forecasting-ae-admissions/
- 2 Jung K, Kashyap S, Avati A, et al. A framework for making predictive models useful in practice. J Am Med Inform Assoc 2021;28:1149–58.
- 3 Takura T, Hirano Goto K, Honda A. Development of a predictive model for integrated medical and long-term care resource consumption based on health behaviour: application of healthcare big data of patients with circulatory diseases. *BMC Med* 2021;19:15.
- 4 Henry KE, Hager DN, Pronovost PJ, *et al.* Early detection of sepsis using predictive analytics. *N Med* 2016.
- 5 Obermeyer Z, Powers B, Vogeli C, et al. Dissecting racial bias in an algorithm used to manage the health of populations. Science 2019;366:447–53.
- 6 Efthimiou O, Seo M, Chalkou K, et al. Developing clinical prediction models: a step-by-step guide. BMJ 2024;386:e078276.
- 7 NHS. Artificial intelligence (AI) and machine learning version1.1. 2023. Available: https://www.england.nhs.uk/long-read/artificialintelligence-ai-and-machine-learning/
- 8 Reynolds R. DeepMind's new AI predicts kidney injury two days before it happens. Science; 2019. Available: https://www.wired.com/ story/deepmind-streams-ai-algorithm-kidney-injury/

Open access

- 9 Powles J, Hodson H. DeepMind and healthcare in an age of algorithms. *Health Technol (Berl)* 2017;7:351–67.
- 10 Bharati S, Mondal MRH, Podder P. A Review on Explainable Artificial Intelligence for Healthcare: Why, How, and When? *IEEE Trans Artif Intell* 2024;5:1429–42.
- 11 Arumugam A, Phillips LR, Moore A, et al. Patient and public involvement in research: a review of practical resources for young investigators. BMC Rheumatol 2023;7:2.
- 12 Banerjee S, Lio P, Jones PB, *et al*. A class-contrastive humaninterpretable machine learning approach to predict mortality in severe mental illness. *NPJ Schizophr* 2021;7:60.
- 13 Banerjee S, Alsop P, Jones L, et al. Patient and public involvement to build trust in artificial intelligence: A framework, tools, and case studies. Patt N Y 2022;3:100506.
- 14 Teachable Machine. Available: https://teachablemachine.withgoogle.com/
- 15 Estiri H, Strasser ZH, Rashidian S, et al. An objective framework for evaluating unrecognized bias in medical AI models predicting COVID-19 outcomes. J Am Med Inform Assoc 2022;29:1334–41.

- 16 Habli I, Lawton T, Porter Z. Artificial intelligence in health care: accountability and safety. *Bull World Health Organ* 2020;98:251–6.
- 17 Ivanescu AE, Li P, George B, *et al*. The importance of prediction model validation and assessment in obesity and nutrition research. *Int J Obes (Lond)* 2016;40:887–94.
- 18 Norori N, Hu Q, Aellen FM, et al. Addressing bias in big data and Al for health care: A call for open science. Patt N Y 2021;2:100347.
- 19 Shipe ME, Deppen SA, Farjah F, et al. Developing prediction models for clinical use using logistic regression: an overview. J Thorac Dis 2019;11:S574–84.
- 20 Xiang Q, Zi L, Cong X, et al. Concept Drift Adaptation Methods under the Deep Learning Framework: A Literature Review. Appl Sci (Basel) 2023;13:6515.
- 21 Yang CC. Explainable Artificial Intelligence for Predictive Modeling in Healthcare. *J Healthc Inform Res* 2022;6:228–39.
- 22 Parikh RB, Obermeyer Z, Navathe AS. Regulation of predictive analytics in medicine. *Science* 2019;363:810–2.