Exploring biases in predictive modelling across diverse populations

In their article, Willeke G van der Plas-Krijgsman and colleagues[1] underscore the limitations of existing predictive tools for older patients with breast cancer through the development of PORTRET, a tool for patients aged 65 and older to predict recurrence and mortality. In this Correspondence, we explore the underlying reason behind PORTRET's development—namely, the inadequacy of current predictive models to assess risk across diverse populations, and we offer potential solutions.

Biases in predictive tools and algorithms have been documented in medicine. In one algorithm used to assess health needs, Black populations were found to be significantly less healthy than White populations with the same risk score.[2] When these disparities in scoring were addressed, the proportion of Black patients receiving additional help increased nearly three-fold.[2] Vyas and colleagues[3] similarly noted that current algorithm approaches in clinical practice bias decisions and direct more resources towards White populations.

Racial disparities are likely to be mirrored across ethnicities, languages, LGBTQ+ identities, and disabilities. Inherent in these findings is the question of whether algorithms or tools should be intended for use in all populations. Futoma and colleagues summarise this sentiment well, "If hospitals want to have useful machine learning systems at the bedside, the broader research community needs to stop focusing solely on generalisability and consider the ultimate goal: will this system be useful in this specific case?"[4] We suggest steps to improve equity in predictive risk modelling development and implementation.

First, the populations whose data were used to train predictive algorithms should be acknowleged. Clinicians should be informed about the populations in which predictive algorithms have been developed and should assess the value of the tool's findings for the patient in question. As many patients do not perfectly fit trial or algorithm inclusion criteria, we call on the global community for high-quality collection and sharing of real-world data, as exemplified by the European project OPTIMA.[5]

Second, the use of algorithms more appropriate for subpopulations should be increased. Various risk algorithms have been created for subpopulations, such as the Black Women's Health Study Breast Cancer Risk Calculator, which was developed with data solely from American-born Black women. Medical professionals should consider tools validated in minority populations that might better assess risk than the so-called gold standard.

Third, equity-focused precision medicine should be supported. With an increased focus on genomics and patients' individualised risk, precision medicine might soon be the primary method of predicting risk in patient populations. It should be ensured that these tools are developed and validated with diverse populations in mind.

We declare no competing interests

*Nishwant Swami, Chiara Corti, Giuseppe Curigliano, Leo A Celi, *Edward Christopher Dee*
deee1@mskcc.org

University of Massachusetts Medical School, Worcester, MA, USA (NS); Department of Epidemiology (NS), and Department of Biostatistics (LAC), Harvard TH Chan School of Public Health, Boston, MA, USA; Division of New Drugs and Early Drug Development, European Institute of Oncology IRCCS, Milan, Italy (CC, GC); Department of Oncology and Hemato-Oncology, University of Milan, Milan, Italy (CC, GC); Laboratory for Computational Physiology, Massachusetts Institute of Technology, Cambridge, MA, USA (LAC); Division of Pulmonary, Critical Care and Sleep Medicine, Beth Israel Deaconess Medical Center, Boston, MA, USA (LAC); Department of Radiation Oncology, Memorial Sloan Kettering Cancer Center, New York, NY 10065, USA (ECD)

1 van der Plas-Krijgsman WG, Giardiello D, Putter H, et al. Development and validation of the PORTRET tool to predict recurrence, overall survival, and other-cause mortality in older patients with breast cancer in the Netherlands: a population-based study. *Lancet Healthy Longev* 2021; **2:** e704–11.
2 Obermeyer Z, Powers B, Vogeli C, Mullainathan S. Dissecting racial bias in an algorithm used to manage the health of populations. *Science* 2019; **366:** 447–53.
3 Vyas DA, Eisenstein LG, Jones DS. Hidden in plain sight—reconsidering the use of race correction in clinical algorithms. *N Engl J Med* 2020; **383:** 874–82.
4 Futoma J, Simons M, Panch T, Doshi-Velez F, Celi LA. The myth of generalisability in clinical research and machine learning in health care. *Lancet Digit Health* 2020; **2:** e489–92.
5 Innovative Medicines Initiative. OPTIMA. Optimal treatment for patients with solid tumours in Europe through artificial intelligence. http://www.imi.europa.eu/projects-results/project-factsheets/optima (accessed Nov 28, 2021).

For more on the **Black Women's Health Study Breast Cancer Risk Calculator** see https://www.bu.edu/slone/bwhs-brcarisk-calculator/