Cambridge Public Health

Causal AI and its uses in public health

WORKSHOP REPORT

Prepared by Becky Wolfe and Danielle Cannon,
Cambridge Public Health

June 2026

Introduction

Artificial Intelligence (AI) is increasingly used in health and medicine, including in medical imaging, risk prediction and the analysis of large datasets. However, in public health — where outcomes are shaped by a complex web of poverty, place and social conditions — standard AI has limitations. Traditional machine learning is good at spotting patterns, but it struggles to explain why outcomes occur or what might change if a new policy is introduced.

In March 2026, Cambridge Public Health convened an interdisciplinary workshop to explore the potential use of causal AI in public health, funded by The CHAI (Causality in Healthcare AI) Hub. Causal AI is an emerging subfield of AI that integrates machine learning with causal inference. It draws on theoretical frameworks such as Pearl's structural causal models and Rubin's potential outcomes framework, and uses methods such as targeted maximum likelihood estimation, double/debiased machine learning, causal forests, and counterfactual reasoning.

Causal AI moves beyond pattern recognition to understand cause and effect, allowing researchers to ask "what if" questions, such as what might happen if services are changed or people are exposed to new awareness campaigns. This report summarises the main discussion points from the workshop.

Key points

- Causal AI offers a distinct methodological orientation for public health by foregrounding cause-and-effect reasoning rather than pattern recognition, but its boundaries require sharper definition, particularly its relationship to causal inference more broadly and to predictive machine learning.
- Promising applications include the evaluation of public health interventions where randomised trials are infeasible, risk stratification in screening, and informing the design of targeted forms of clinical and policy action that adapt to need and context.
- To be used well, the application of causal AI in public health requires a shared language, stronger mutual understanding, and closer collaboration across disciplines.
- Ethical and social issues need to be considered from the start of the research process, both in relation to the use of AI and to how causality is understood and framed in public health.
- Causal AI may be a useful tool for supporting public health decision-making, but it should not replace human judgment. Its value lies in helping professionals think through options, not in removing the need for interpretation and decision-making.

What is causal AI?

The workshop began by exploring causal AI and how it differs from standard predictive machine learning.

> The CHAI hub defines causal AI as follows:
>
> *"An artificial intelligence designed to untangle complex causal relationships and to use them in robust problem-solving. This differs from traditional AI, which identifies patterns and connections in data. Causal AI digs deeper to figure out what drives outcomes. This means it doesn't just spot links — it understands why things happen."*

From predictive patterns to causal approaches

The session highlighted important differences between predictive AI and causal AI. Much current machine learning in healthcare relies on predictive systems that use signals such as medical imaging or patient characteristics to estimate the probability of an outcome. While these models can be effective for estimating risk, they are shaped by patterns in historical data and may be less reliable when those patterns reflect underlying processes that are not well understood or no longer hold in a changing environment.

One example illustrated this problem. A model suggested that asthma reduced the risk of death from pneumonia. In practice, however, this pattern reflected differences in care, with patients with asthma more likely to receive intensive monitoring and treatment. Without understanding these underlying processes, such findings could be misinterpreted.

Causal AI was presented as a way of untangling these spurious associations by accounting for how human decisions and interventions shape outcomes. In contrast to standard predictive approaches, it focuses on cause and effect rather than pattern recognition alone. It aims to account for confounders and system changes, so that estimates and analyses remain meaningful when an intervention, policy change, or treatment alters the conditions under which outcomes arise.

Clarifying terminology: AI vs machine learning

The discussion highlighted some uncertainty about how broadly the term 'causal AI' is used, reflecting differences in how it is defined and understood. One example discussed was research described as causal AI that used machine learning methods to estimate outcomes in a randomised trial setting — suggesting that the term may, at times, be applied too broadly.

The discussion also noted that, while machine learning generally refers to methods that learn patterns to make predictions, AI may also involve planning, reasoning and interaction. However, causal AI should not be seen as a way to "take the thinking out" of causal inference. Rather, it should be understood as a tool to support and enhance human judgement in complex decision-making.

Example: Lessons from applying double machine learning to mental health research

During the workshop, an example was presented of research using double machine learning (DML), a causal inference approach using birth cohort data from the Millennium Cohort Study and survey data from the UK Household Longitudinal Survey, to estimate average treatment effects on young people's mental health outcomes.

The example highlighted two lessons about the practical use of DML in real-world research. First, data collected at three-year intervals, as in the Millennium Cohort Study, did not appear to be frequent enough to estimate effects prospectively. This meant that contemporaneous models had to be used instead, although these still helped identify important risk and protective factors.

Second, the method appeared less suitable when the available data lacked sufficient relevant information to capture the relationship between treatment and the outcome, especially when the treatment was relatively uncommon. In the UK Household Longitudinal Survey example, fewer than 10% of participants were exposed to the treatment of interest, which made propensity weighting more appropriate.

Ethical and social considerations

Ethical and social considerations were a prominent part of the opening session and continued to feature in later discussions. While ethical issues associated with AI in medicine are clearly defined, those associated with AI in public health are less so, particularly in relation to causal AI. Issues relating to AI and causality need to be considered from the outset of the research process, rather than being treated as secondary concerns at the end. Specific considerations include those outlined below.

Area	Issue	Considerations
AI	**Data**	Privacy; access to the right types of data; data equity and bias; governance
	AI models	Transparency, explainability, accountability, uncertainty and updating, deployment, usability and access
	Sustainability	Material resources and costs; access to compute
Causality	**Scales of causation**	Individual, population, causes of causes
	Epistemology and politics	Conflicting narratives of causation; plausibility of causes; individual/societal attribution; authority of AI
	Consequences of causation	Balance of individual/system-level interventions; responsibility and blame

In relation to data, one issue concerns the limitations of some commonly used datasets. Models based on electronic health records or epidemiological data often provide only a partial picture, potentially overlooking the social determinants of health. Over-reliance on these sources could overlook the 'causes of the causes,' making it difficult to understand the underlying drivers of population health.

The session also raised questions about how causality is framed in public health discourse and about the implications this may have for public policy. There is a risk that, if causal AI identifies 'lifestyle choices' as a primary cause of disease, it could inadvertently reinforce a culture of individual blame.

Public health challenges from research and practice

The second session set out a series of public health challenges drawn from research and practice to ground the interdisciplinary group discussions.

Health inequalities
Inequalities are evident in access to healthcare, the care people receive, and health outcomes. They affect a wide range of groups and circumstances, including people experiencing socioeconomic disadvantage, marginalised groups, and those with disease-specific needs. A key challenge is understanding how different drivers of inequality vary across communities and how multiple forms of disadvantage interact. Another is determining how interventions and services should be adapted for different groups, rather than assuming that a single approach will work for everyone.

Complex health and care systems
Managing patient flow across the entire care journey, including within hospitals and into community care, is challenging because these systems are complex, involve many actors across organisational boundaries, and do not follow a single standard model. Difficulties may arise not because of a fault in any one part of the system, but because different parts are working to different pressures, timescales and objectives.

Cancer screening and early detection
A key challenge in screening and early detection is deciding who is most likely to benefit from an intervention, and when. Decisions about screening often rely on evidence of average benefit, even though the likely value of screening may differ substantially between individuals and groups. This can be seen in breast cancer screening, where different countries have reached different decisions based on the same trial evidence. Another challenge is the time required to generate robust evidence through trials.

Evaluating public health interventions
Evaluation of public health interventions is often challenging because randomisation is either unethical or impractical, and causal relationships must then be inferred from observational data. In some cases, such as smoking or substance abuse, it would not be ethical to assign people to potential harm; in others, such as social media, controlled comparison is difficult because access cannot realistically be managed. Other challenges include establishing who has been exposed to an intervention (for example, a public health message at a bus stop), dealing with selection biases, and measuring outcomes that may take years or decades to emerge.

Local authority priorities and challenges
Local authorities face several practical challenges, particularly in relation to the social determinants of health. One is identifying where action should be targeted at the population level, given the volume of data and the difficulty of determining the main drivers of worsening health outcomes. Another is responding to rising demand and budget pressures, including in adult social care, where there is interest in identifying which residents might benefit most from earlier support. A further challenge lies in understanding where intervention is most needed across children's pathways, including early years development, special educational needs and disabilities (SEND), and mental health services.

Areas where causal AI may be useful

Drawing on these challenges, participants discussed where causal AI might be useful and identified the following areas.

Risk stratification. Causal models may support risk stratification in cancer screening, particularly where the introduction of screening changes the data available over time and makes conventional prediction more difficult.

Precision public health and precision medicine. Causal models may also help explore how social determinants of health interact to drive inequalities, supporting more tailored recommendations or referrals in general practice. This could include more personalised advice or targeted forms of referral, such as social prescribing, informed by a fuller understanding of patient context.

Research, trials and evaluation. Approaches such as conditional average treatment effects may help identify which groups are more likely to benefit in cancer screening trials and inform later trial design. They may also help support earlier adaptation of trials where strong evidence emerges. Synthetic controls may also be useful where randomised trials are not feasible and real-world controls are weak.

Supporting question formulation. AI tools could serve as useful support to help practitioners, including those in local authorities, formulate questions in ways that can be taken forward more easily by modellers.

Bridging AI and public health communities

In addition to exploring potential uses, the discussion considered how public health researchers and researchers working with AI and data science might be brought into closer conversation. A recurring theme was the importance of clearer shared language and a more common understanding across disciplines. This included both the need to define and explain what is meant by AI, particularly given that the term may be used differently by different groups, and the need for more opportunities for interaction across disciplines and organisations, so that researchers can better understand one another's assumptions, constraints and ways of working.

Data opportunities and challenges

Causal AI approaches may create opportunities to use a wider range of data, including free-text clinical conversations or medical information. A major opportunity identified was the use of large language models to extract useful information from clinical narratives. Potential uses discussed included extracting determinants of health from GP consultation notes, supporting digital scribes in clinical settings, identifying relevant contextual information for intervention design, and translating practical questions into forms that modellers can use.

At the same time, practical difficulties remain in linking datasets held in different places, working across different types of data, and ensuring that sufficiently rich and relevant variables are available for causal analysis. This is a particular issue for research on the wider determinants of health, where relevant data — for example on housing, income or benefits — may not be routinely collected, consistently recorded or readily accessible across organisations. Some datasets may be stronger than others in this respect; for example, children's data may offer particular opportunities to track development and outcomes over time through education and related services.

Trust, privacy and data governance

The use of AI and health data raised important questions of privacy, governance and public trust. The need for transparency in how models are designed and trained was highlighted to assess potential bias.

Privacy and disclosure risk were also discussed, including whether models trained in secure data environments might still leak sensitive information, and how data access committees can judge that risk. Technical approaches such as federated learning and differential privacy were mentioned, but so too were the limits of purely technical fixes, the need for mitigation and accountability, and concern about placing long-term trust in proprietary AI companies whose policies may change.

More broadly, the discussion highlighted the importance of public trust and included a clear caution that AI should not replace the thinking involved in causal inference. Careful human judgement is essential in deciding what questions matter, what data are relevant, and how findings should be interpreted.

What future success might look like

In the closing discussion, participants reflected on what future success in this area might look like in practice.

Starting with the problem. The starting point should be the public health challenge that needs to be addressed, rather than the capabilities of a particular technology.

Becoming more concrete about its use. Progress is likely to depend on moving towards a narrower and more nuanced understanding of where causal AI is most useful, with a smaller number of clearer and better-defined use cases.

Having the right data in place. This work depends on linking datasets and having sufficiently rich, theory-informed variables relevant to the exposure and outcome being studied. Without the right data, these models will not support robust causal inference.

Supporting study design and analysis. Causal AI may help strengthen study design and analysis, including through the development of synthetic controls, and may help address complex causal relationships that are difficult to capture using current methods alone.

Supporting more targeted action. Used well, causal AI may support both precision policymaking and precision medicine, particularly in situations shaped by multiple interacting factors rather than a single clear cause.

Conclusion

Causal AI may have useful applications in public health, particularly in areas where outcomes are shaped by complex, interacting causes and decisions need to be made under uncertainty. Its value lies not in replacing human judgement, but in helping to support more informed analysis, discussion and action across research, policy and practice.

At the same time, the workshop made clear that this is not simply a technical challenge. Progress will depend on clearer shared language, closer collaboration across disciplines, and careful attention to ethics, trust and governance. The challenge now is to identify where causal AI can make a meaningful contribution to real public health problems, and under what conditions it can be used responsibly.

Acknowledgements

The workshop was generously funded by The CHAI Hub. Many thanks to the following for their help in planning the workshop and contributing to this report:

Dr Tom Callendar, Department of Public Health and Primary Care
Dr Danielle Cannon, Cambridge Public Health
Professor John Clarkson, Department of Engineering
Dr Richard Milne, Department of Education
Dr Rakesh Modi, Department of Public Health and Primary Care
Dr Sharon Neufeld, Department of Psychiatry
Professor Niels Peek, Department of Public Health and Primary Care
Professor Nora Pashayan, Department of Public Health and Primary Care
Dr Sarah Dougan, Cambridgeshire County Council
Professor Nick Wareham, IMS Epidemiology
Professor Qingyuan Zhao, Department of Pure Mathematics and Mathematical Statistics

For more information please contact Becky Wolfe (rfh34@cam.ac.uk)

UNIVERSITY OF CAMBRIDGE

Cambridge Public Health

CHAI