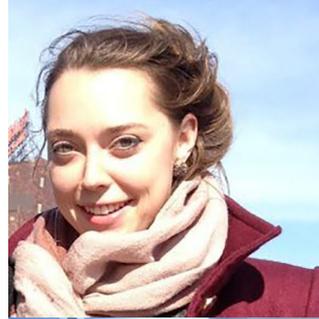# AI in healthcare

examining challenges related to justice and trust, and the contribution of procedural justice

Dr Anna Pacholczyk
Lecturer in Medical Ethics and Law
MBBS Medical Ethics and Law Lead
Institute for Medical and Biomedical Education
St. George's, University of London, UK

St George's
University of London

Background

- ❏ AI is a fast advancing field
- ❏ a wealth of current and possible healthcare applications including applications in the 'neuro' arena (Yu, Beam & Kohane 2018)
- ❏ AI in healthcare raises considerable ethical unease (Di Nuci 2019)
- ❏ The field requiring intense public and academic scrutiny to facilitate ethical implementation

Aims

(1) consider AI from the lense of a public good

(2) outline ethical challenges facing the use of AI in healthcare

(3) evaluate the possible contribution of procedural justice approaches to addressing these challenges

St George's
University of London

Aim (1) Is AI/ are AI infrastructure elements a public good?

Question
Could AI in healthcare/elements of AI technology/infrastructure be considered as a public good? (for the concept of public good see: Stiglitz 1999)

(1)   Local public goods

(2)   Global public goods
   (a)   Nonrivarlous consumption
   (b)   Nonexcludability

Conclusion
Elements of AI can be considered to be local public goods.

Knowledge related to AI (e.g. code, algorithms and datasets) fulfil criteria or nonrivarlous consumption and nonexcludability necessary for global public goods.

Why does it matter?
Weather we consider something to be a public good will change the way we approach regulation, funding, access, transparency & sharing. It might also have profound influence on behaviours of stakeholders, for example willingness of citizens and patients to contribute to AI datasets.

Example: 2021 'NHS Data Grab' (UK)

AI's effectiveness in healthcare crucially depends on availability of rich, good quality and representative datasets in usable formats.

The General Practice Data for Planning and Research scheme aimed to facilitate data sharing from GP practices and integration with existing datasets for medical research and planning purposes (Renault 2021)

The scheme was dubbed 'NHS data grab' in the media, after nearly **1.3 million people opted out in June 2021**

The public did not know about the planned data sharing, despite short deadlines for opting out before data transfer. Concerned GPs and activists run an online campaign.

Information was widely circulated on WhatApp/Facebook to unsuspecting public and subsequently covered by the media.

NHS Digital

Return to contents

Table 1: Number of national data opt-outs, England, August 2020 - July 2021

Date	ONS code	Code	Name	Opt-out	List size	Opt-out Rate	Deceased
01 July 2021	E92000001	Eng	England	3,032,917	60,970,002	4.97	89,331
01 June 2021	E92000001	Eng	England	1,757,764	60,860,759	2.89	88,127
01 May 2021	E92000001	Eng	England	1,650,335	60,798,847	2.71	87,064
01 April 2021	E92000001	Eng	England	1,646,698	60,744,002	2.71	85,905
01 March 2021	E92000001	Eng	England	1,616,294	60,650,740	2.66	84,671

Example: 2021 'NHS Data Grab'

Perception of data as a public good conflicted with objections to private companies access

Main objections included (Renault 2021)

(1) lack of consultation
(2) lack of transparency
(3) access of private companies to data

In the UK, the NHS (a public health service) is seen as a collective good, enjoying high levels of public support. Health data sharing relies on perception of sharing data as contribution to the public good, and was undermined (among other impacts) by a perception that date will be shared with private companies and potentially used for profit.

Scheme was paused indefinitely in August 2021.

PRIMA FACIE IMPORTANCE OF EXAMINING THE 'PUBLIC GOOD' ASPECTS OF AI INFRASTRUCTURE (such as HEALTH DATASETS) IN DELIBERATION

IMPORTANCE OF PROCEDURAL JUSTICE PRINCIPLES AND CONSULTATION

St George's
University of London

Aim (2) Outline ethical challenges facing AI applications in healthcare

Challenges related to:

- adaptations needed to enable successful adoption (e.g. creation of large scale usable data infrastructures)
- stakeholder attitudes and perceptions (e.g. user trust)
- limitations of the technology (e.g. bias and explainability)
- potential negative impacts of adoption (e.g. privacy and workforce impacts, perpetuating and/or increasing discrimination due to dataset characteristics etc.)

Academy of Medical Sciences, Medical Research Council and National Institute for Health Research (2019). *AI and health.*
RSA (2018). *Artificial intelligence: real public engagement.*

St George's
University of London

(2) Ethical challenges of AI applications in healthcare: example

Trust
IBM Watson Oncology

FOCUS ON: trust, explainability and bias

TRUST

IBM Watson Health (Strickland 2019; Hamilton et al 2019; Watson 2019; Schmidt 2017)

- Hype in Watson Oncology. Applications undermined by hype, technical challenges (concordance rate varied across different cancers) & lack of user trust.

Uptake of technology will be driven by clinicians (Sheetz et al 2021). Hype and overpromising can be detrimental and undermine trust. Stakeholder perceptions and trust will influence adoption (Morrison 2021).

Increasing accuracy (concordance) and designing systems with view of explainability proposed to overcome challenges in trust of professionals.

(2) Ethical challenges of AI applications in healthcare: example

Explainability

EXPLAINABILITY

Diagnostic systems based in correlational data with AI trained on data examples, have been criticised for lack of explainability and justification..

- The value of justification, explanation and expertise in healthcare diagnosis and doctor-patient relationship
- Problem: when AI system accuracy (concordance with experts) is sufficiently high, but explainability is low
- Explainability emerged as a serious objection to use of some AI systems and sparked explainability-enhancing designing

ETHICAL ISSUE: How much weight should be given to explainability and justification vs accuracy in AI in diagnostic medicine?

(2) Outline ethical challenges facing AI applications in healthcare: example

Bias

BIAS

Artificial intelligence is also likely to incorrectly estimate risks for patient groups with missing data or less data.

Example 1: Black women with breast cancer had a lower probability of being tested for high-risk germline mutations compared with White women (despite similar risk) (McCarthy et al 2016). An AI algorithm that depends on genetic test results is more likely to make an error in risk assessment of breast cancer for Black patients than White patients.

Example 2: AI trained to detect melanoma with research participants drawn from majority White European samples with high accuracy, would run a risk of mistakes when applied to diversity of skin colours. (Noor 2020; Adamson & Smith 2018).

ETHICAL ISSUES:: how best to assess, explore and minimise the impacts of AI technology increasing and/or perpetuating bias within inequitable/biased societies

Aim (3) Procedural justice as contributing to addressing ethical challenges

Procedural justice in the context of AI

1) Procedural fairness in decisions made with an aid, or following, an algorithm or other automated decision making procedure

2) **Procedural fairness in decisions and recommendations evaluating healthcare applications of AI**

Legitimacy of recommendations issued by agencies undertaking appraisals of technologies and treatments in medicine at the national/local level. Legitimacy depends, in part, on procedural justice (Daniels 1996; Daniels & Sabin 1997).

Procedural justice: perceptions of decision outcome fairness depend not only on the outcomes, but also on the process used to determine the outcome.

(3) Procedural justice & deliberation in HTA

Accountability for reasonableness (Daniels & Sabin, 2008)

Fair deliberative process principles

(i) *Publicity*: specifically, transparency about reasons
(ii) *Relevance*:rationales that rely on reasons that all can accept as relevant to meeting health needs fairly
(iii) *Revisability*
(iv) *Enforcement*

Such deliberations can be transformative, in the sense of seeking novel practical and normative solutions and reconceptualisations (Grin, & Van de Graaf 1996)

Deliberation might be appropriately managed and structured.

Ethical issues should be included in health technology assessments (HTA) (Daniels & van der Wilt, 2016 for argument in favour of extending HTA to ethical issues). Citizen juries constitute one of the methods for evidence-informed deliberative processes that can be used in HTA (Jansen et al 2017).

Let us consider ONE example...

St George's
University of London

(3) Procedural justice & ethical challenges facing AI applications in healthcare:

Example of explainability

EXPLAINABILITY

Diagnostic systems based in correlational data with AI trained on data examples, have been criticised for lack of explainability and failing to provide justification…

ETHICAL ISSUE: How much weight should be given to explainability/justification vs accuracy in AI in diagnostic medicine in HTA?

Method: CITIZENS' JURIES

Citizens' juries are a form of deliberative democracy eliciting informed judgment from a representative sample of the general public around policy questions (Van der Veet et al 2021)

(3) Example: public consultation in AI in healthcare (UK) and <mark>explainability</mark> in medical AI

Trading off accuracy and explainability in AI decision-making: findings from 2 citizens' juries

Sabine N van der Veer [1], Lisa Riste [2] [3], Sudeh Cheraghi-Sohi [2] [4], Denham L Phipps [3], Mary P Tully [3], Kyle Bozentko [5], Sarah Atwood [5], Alex Hubbard [6], Carl Wiper [6], Malcolm Oswald [7] [8], Niels Peek [1] [2]

FINDINGS: <u>In healthcare scenarios, jurors favored accuracy over explainability,</u> whereas in non-healthcare contexts they either valued explainability equally to, or more than, accuracy.

How much weight we should give to the problem of 'blackboxing' and explainability depends on context in views of users. .

The issue with explainability have been emphasised by professionals, but they may matter less from the perspective of users.

Consultation is an important contributor to developing recommendations in HTA of AI applications..

St George's
University of London

Conclusions & further directions

→ **Elements of AI technology and infrastructure can be seen as a public good**

→ **Ethical challenges in healthcare AI applications include issues of trust, explainability and bias**

→ **Procedural justice approaches can contribute to and motivate deliberative processes about ethical issues in AI in healthcare, with citizens' juries as one example.**

→ **This presentation aimed to provide a short overview. Further analysis include critical evaluation of the role, contribution and limitations of procedural justice and deliberative approaches in HTA with a focus on ethical assessment of AI in medicine.**