



## AI, Surgery and Inequalities

Artificial Intelligence (AI) technologies offer new opportunities to improve surgery but there is a growing recognition that the development of these technologies can unintentionally create or exacerbate inequalities. In the context of surgery, AI can be used in several ways, including detecting and characterising signs of disease, providing guidance to surgeons during procedures (for example through improving endoscopic navigation or providing augmented reality overlays) and in surgical robots. The future success of these emerging technologies in surgery will depend on ensuring they are designed with, and for the public and patients.

### About the report

This report is produced by the Wellcome EPSRC Centre for Interventional Surgical Sciences (WEISS) as part of “In Theatre”, a programme developed to help people discuss challenging topics and break down barriers to engagement.

The report is based on a series of discussions that took place during two online workshops in July and September of 2022, with participants from the UK Government, academia and the wider community. The main objectives of the first workshop were to map the different ways in

which AI and surgery might lead to inequalities and how those inequalities relate to current public understanding and concerns. As a follow-up, the second workshop focussed on exploring what steps might be taken to ensure that the development of new surgical technologies don't inadvertently contribute to health inequalities.

Through the planning and delivery of workshops, and to contextualize discussions, we took inspiration from a few existing frameworks to evaluate sources of inequalities related to AI<sup>1,2</sup>.

<sup>1</sup> Source: Leslie, D., Mazumder, A., Peppin, A., Wolters, M.K. and Hagerty, A., 2021. Does “AI” stand for augmenting inequality in the era of covid-19 healthcare?. *bmj*, 372.

<sup>2</sup> Uche-Anya, E., Anyane-Yeboah, A., Berzin, T. M., Ghassemi, M., & May, F. P. (2022). Artificial intelligence in gastroenterology and hepatology: how to advance clinical practice while ensuring health equity. *Gut*, 71(9), 1909–1915.

## Key points

- Artificial Intelligence (AI) technologies offer new opportunities to improve surgery but there is a growing recognition that the development of these technologies can unintentionally create or exacerbate inequalities.
- Not including patients and the general public in the design and choice of new research areas may mean that some conditions and groups of people are more likely to benefit from new technologies than others.
- Selecting and designing research projects that minimise the risk of unequal outcomes may require researchers to work in new ways and to build or access new skills.
- Unrepresentative and incomplete datasets combined with incompatibilities in the underlying NHS infrastructure make it very difficult to evaluate health outcomes.
- Transparency and clear messaging on how health data is going to be used is very important as it directly affects trust in institutions and how much patients and the public are willing to engage.
- **Patients and the public should be more involved in research and governance around data stewardship. This may facilitate collecting important information such as demographics, while maintaining high standards in privacy and security.**
- **Researchers and clinicians need more capacity to engage with complex governance and time to actively consider positive and adverse effects of technologies being developed.**
- **Policy makers and funders should promote and support data sharing initiatives on creating large scale national level databases. They should also promote greater diversity and inclusion in research funding in both the definition of new priority areas and the safe deployment of new technologies.**

## Potential issues in the design and development of medical devices (specifically focusing on Artificial Intelligence and Surgery)

There are many ways in which artificial intelligence technologies may lead to inequalities. In this series of workshops, we prompted participants to think about the use of AI in surgical technologies and whether some groups are unfairly benefiting or being disadvantaged, what the impacts on infrastructure and institutions might be and how will this affect access to and quality of healthcare. We also discussed barriers that would prevent action to address health inequalities in the development and implementation of AI-based surgical technologies.

In our discussions we identified two main sources of inequalities: 1) issues surrounding the research and innovation process preceding development and manufacturing; and 2) the datasets and infrastructure available to researchers when developing new technologies.

### 1) RESEARCH AND INNOVATION PRECEDING DEVELOPMENT AND MANUFACTURING;

#### Research design

The choice of which problems to address can result in inequalities: some technologies are developed to solve a particular problem that happens to be more prevalent among certain groups in society whilst conditions that affect other groups are under-researched. This means that some groups of people are more likely to benefit from new technologies than others.

The choice of research topic is largely driven by the availability of funding and funders tend to set their own individual priorities without a central body for overview or coordination. The result is that some conditions are overlooked because they are not prioritised by any funders.

In addition, the clinical research question is often developed without input from the communities it is meant to serve. There is a lack of focus on mapping unintended consequences which may affect them negatively.

Finally, fairness and diversity can sometimes be seen as a tick-box exercise at the beginning of projects. The lack of frameworks to evaluate the ethical dimensions of the development of AI technologies makes it hard for researchers to assess whether their intended research could lead to inequalities.

### **Patient and public involvement**

Workshop participants held a general view that greater patient and public involvement could be one way of choosing the ‘right’ research topics and avoiding poor research design. However, there are a number of barriers to achieving this.

- Insufficient efforts to engage with traditionally underrepresented or marginalised groups. This can be anything from finding volunteers for trials, creating patient groups or developing survey/focus group work. Limiting advertising or ignoring the unique challenges of engaging with seldom heard groups will typically lead to the same homogenous groups being involved (audiences are often self-selecting).
- The necessary time to establish the relationships and trust that allow meaningful and continued engagement with these groups is not compatible with research project lifecycles, which are generally too short to allow meaningful relationship-building. There is also a lack of patient groups that focus on underrepresented and marginalised people, which might make it harder to reach them.

### **Insufficient resources and capacity (to deal with data, patients/clinicians and regulations)**

Selecting and designing research projects that minimise the risk of unequal outcomes may require researchers to work in new ways and to build or access new skills. For example:

- delivering effective and representative patient and public involvement processes;
- undertaking relevant analyses (and ongoing monitoring) to determine potential sources of bias and negative outcomes that might arise from their projects; and
- understanding complex guidance on ethical development of AI technologies, from complying with standards on the data that is used, to accurately reporting clinical evaluations and assessing performance.

## **2) DATASETS;**

While it is widely acknowledged that using unrepresentative datasets to train AI is problematic and may lead to unequal outcomes, there is not yet a consensus on what a balanced/representative dataset would look like in practice. Governance mechanisms to deal with data and the underlying infrastructure also pose a challenge to how good datasets can be.

### **Incomplete and underrepresented data**

Data is often unrepresentative of all the communities that AI technologies have the potential to serve, but there is a question as to what it should be representative of: the patient population? The local community? The country? Global? Which characteristics should be considered? UK equalities legislation defines 9 protected characteristics, but these are not necessarily the same in other jurisdictions.

Furthermore, existing (big retrospectively collected) datasets are anonymised and do not have sufficient information, such as demographics, which makes it impossible to evaluate health outcomes on this basis.

To gather this data, datasets need to be recreated from scratch, but this problem may reoccur as “new” variables emerge in the future. It is important to consider how to manage these issues and create platforms that would directly consider underrepresented groups.

### **Complex governance and difficulties accessing data**

Information governance and requirements for data protection become very complex when personal and/or identifiable data is collected. Furthermore, there is no standardisation of local information governance criteria and processes across the NHS.

As a result, researchers seeking to access and securely store data can encounter large administrative barriers.

### **Infrastructure challenges**

There are major disparities in different health settings regarding IT structures and capacity from the NHS to support requests from researchers.

Access to AI technologies varies across the country and it is not clear whether everyone would be able to benefit from these tools regardless of where they live. Participants questioned whether every health setting possesses the necessary expertise to operate these

technologies. They also asked if there are effective strategies in place to deal with specific local health contexts and ensure support for those who need to travel to benefit from these technologies.

The pressure to develop complex infrastructures and inter-linkages (which need shared skills as well as shared standards) has sometimes been addressed by turning to the tools big tech firms have available, which may not be sustainable both from a reputational and practical point of view.

Transparency and clear messaging on how the data is going to be used is important. Often the value of data and who will benefit from it is not clear. Participants noted that patients may have concerns that NHS data will be exploited for monetary gain, where private firms are involved in handling patient data.

## Potential solutions in the design and development of medical devices (specifically focusing on Artificial Intelligence and Surgery)

Based on the issues highlighted above, participants suggested actions that might help to improve research practices and deliver more equitable health outcomes. Below we present ideas for further investigation based on the insights and experience of all participants involved, rather than fully worked up recommendations.

### 1) RESEARCH AND INNOVATION PRECEDING DEVELOPMENT AND MANUFACTURING;

In general, there was agreement that the research and innovation process should be more open and inclusive. There is a need to develop more meaningful public engagement with patients and the public, and to identify strategies that have had successful and continued engagement from traditionally underrepresented or marginalised groups.

Promoting greater involvement of healthcare professionals (such as clinicians, nurses, care workers or administrators) might help to identify practical solutions to overcoming administrative barriers.

Finally, developing collaborations between academia, clinicians, NHS organisations and industry will enable scaling of new and good ideas.

## Specific ideas for further exploration:

- Embed incentives in research awards for greater consideration of Diversity and Inclusion in the design of research projects;
- Encourage researchers to spend time actively considering positive and adverse effects of technologies being developed and provide appropriate risk assessments for their use (Along similar lines to the “Equality & Health Inequality Impact Assessment” being developed by NHS England);
- Share methodology and code openly so that others can adopt/reproduce good approaches others have built to monitor inequalities, we should all be learning from each other;
- Design clear guidance / training for Patient/Public Involvement and Engagement in AI research;
- Make evaluation a priority in funding programmes, so that independent evaluations are commissioned and include top level representation of patient populations and their respective conditions.

### 2) DATASETS;

There is no single solution to create more robust and representative datasets.

First of all, in order to increase representation, participants highlighted the importance of policy makers and researchers promoting an open discussion with the public about sharing and governance of health data. The involvement of any third parties and their interests about the data being collected must be clearly articulated as it will greatly influence how much people trust institutions and are willing to engage.

<sup>3</sup> STANDING Together initiative: [https:// www.datadiversity.org/delphi](https://www.datadiversity.org/delphi)

It is also important to develop consensus on what a good dataset looks like. This will make the argument for data sharing easier<sup>3</sup>. Furthermore, the platforms where this data is shared and accessed can also be improved with data stewardship or cooperative initiatives, which may make it easier for researchers to comply with regulations. Once the above are in place, we will have the key ingredients to develop a secure system where data sharing between institutions is the default and patients are able to make informed choices about their data. This will also allow a public record of decision making to make healthcare more democratic and allow scrutiny from the wider public.

## Specific ideas for further exploration:

- Promote and support data sharing initiatives to create large scale national level databases;
- Empower patients to be part of governance around data stewardship. Trusted research environments (TREs) have been suggested as a route for increasing trust around security and privacy of data. Participants suggested that involving patients and the public in the design and governance of TREs might help to increase trust further;
- Develop more flexible and agile research infrastructure and proportionality in regulations – further investing in diversifying the actors in charge of developing TREs<sup>4</sup>;
- Develop training for researchers on early and targeted interventions to improve communication with underrepresented and marginalised audiences. For example, making sure that information is accessible (providing non-digital options if necessary), and making it easier for patients that are willing to be involved with data research to do so;
- Devise systems to disseminate and implement health dataset standards, including an active voice of patients, clinicians, researchers and policy makers.

<sup>4</sup> Data and Analytics Research Environments UK (<https://dareuk.org.uk/>)

## Public and patient concerns and trust regarding digital technologies

Participants reported that although there is not enough public involvement in medical research in general, artificial intelligence is a highly complex, technical field, and it can be hard to bring everyone 'up to speed'. This may make it even harder to deliver effective public engagement on Artificial Intelligence topics.

The population in the UK is very diverse and contains multiple cultures, nationalities, religions and ethnicities. There may be different appetites for using AI technologies between groups, leading some to miss out on potential benefits if they 'opt out'. For example, if information for patients is not accessible or if there is not enough support for people with disabilities, some people may not feel comfortable with AI technologies being used.

Even for some healthcare professionals, there may be a fear that engagement with new technologies will negatively impact on service delivery. This in turn also affects how patients react to AI.

We also set out to discuss current and future concerns that patients and the public have towards these technologies.

### CURRENT CONCERNS

- **Lack of transparency:** there are concerns about the lack of transparency in institutions regarding the use of data, its anonymisation, ownership and distribution;
- **Insufficient regulations:** there are also worries that there are insufficient checks and evaluations of how data is collected and how private information is kept secure, as well as common standards to regulate and support the safe application of new technologies.

## FUTURE CONCERNS

- **Significant security and privacy standards:** As the trend towards digitalisation increases, participants speculated that further concerns towards data privacy and security may arise. This has implications in terms of designing new infrastructures and safety protocols;
- **Ethical principles:** There is the need to both inform the population regarding their attitudes towards data whilst at the same time shifting the responsibility towards institutions to be more transparent about their intentions. Patients should not suffer consequences deriving from institutional failures in upholding ethical principles and agreed uses of their patient's data.

## Specific ideas for further exploration:

- Expanding the public debate on the benefits/harms of AI, including understanding of ethics, social impacts of tech, and algorithmic harm;<sup>5</sup>
- Expanding on current initiatives to track public attitudes towards AI by further examining the influence of socio-demographic and health-related variables in the acceptance of AI;<sup>6</sup>
- Investigate whether one approach will universally improve people's trust in AI and/or whether a different approach would be needed for different groups.<sup>7</sup>

<sup>5</sup> <https://www.gov.uk/government/publications/public-attitudes-to-data-and-ai-tracker-survey>

<sup>6</sup> Antes, A. L., Burrous, S., Sisk, B. A., Schuelke, M. J., Keune, J. D., & DuBois, J. M. (2021). Exploring perceptions of healthcare technologies enabled by artificial intelligence: an online, scenario-based survey. *BMC Medical Informatics and Decision Making*, 21(1), 221. <https://doi.org/10.1186/s12911-021-01586-8>

<sup>7</sup> Lee, M. K., & Rich, K. (2021). Who Is Included in Human Perceptions of AI?: Trust and Perceived Fairness around Healthcare AI and Cultural Mistrust. *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, 1–14. <https://doi.org/10.1145/3411764.3445570>



## Annex 1: Summary of resources and initiatives

### Research and statistics

- European Health Data Space  
[https://ec.europa.eu/health/publications/communication-commission-european-health-data-space-harnessing-power-health-data-people-patients-and\\_en](https://ec.europa.eu/health/publications/communication-commission-european-health-data-space-harnessing-power-health-data-people-patients-and_en)
- Open Data Institute projects on AI  
<https://theodi.org/topic/emerging-tech/>
- Standing Together  
<https://www.health.org.uk/funding-and-partnerships/programmes/artificial-intelligence-and-racial-and-ethnic-inequalities/>  
<https://www.datadiversity.org/>

### Policy Papers and Consultations

- Data storage and processing infrastructure security and resilience – call for views  
<https://www.gov.uk/government/publications/data-storage-and-processing-infrastructure-security-and-resilience-call-for-views/data-storage-and-processing-infrastructure-security-and-resilience-call-for-views>
- Inclusive Britain  
<https://www.gov.uk/government/publications/inclusive-britain-action-plan-government-response-to-the-commission-on-race-and-ethnic-disparities/inclusive-britain-government-response-to-the-commission-on-race-and-ethnic-disparities>
- <https://www.bennett.ox.ac.uk/blog/2022/07/bennett-insights-an-overview-of-uk-data-policy-developments/>
- Common Regulatory Capacity for AI  
<https://www.turing.ac.uk/research/publications/common-regulatory-capacity-ai>
- Secure data environment for NHS health and social care data  
<https://www.gov.uk/government/publications/secure-data-environment-policy-guidelines/secure-data-environment-for-nhs-health-and-social-care-data-policy-guidelines>

- Government response to consultation on the future regulation of medical devices in the United Kingdom

### Guidance and Regulation

- HDRUK 5 safes  
<https://www.hdruk.ac.uk/news/building-a-robust-and-trustworthy-approach-to-health-data-research-and-innovation-across-the-uk/>
- 2022 artificial intelligence (AI) update publication  
<https://www.nice.org.uk/about/what-we-do/our-programmes/evidence-standards-framework-for-digital-health-technologies#ai-update>
- National AI strategy  
<https://www.gov.uk/government/publications/national-ai-strategy>
- Data protection and digital information bill  
<https://bills.parliament.uk/bills/3322>
- BSI White Paper – Overview of standardization landscape in artificial intelligence
- EU AI Act  
<https://artificialintelligenceact.eu/the-act/>
- AI international standards collaborations
- Software and AI as a Medical Device Change Programme – Roadmap  
<https://www.gov.uk/government/publications/software-and-ai-as-a-medical-device-change-programme/software-and-ai-as-a-medical-device-change-programme-roadmap>

## Annex 2: Summary institutions and participants engaged in the project

1. **Dan Stoyanov**, WEISS Director
3. **Laurence Lovat**, WEISS Clinical Director
4. **Evangelos Mazomenos**, WEISS Deputy Director for Engagement
5. **Matt Clarkson**, WEISS Deputy Director for Education
4. **Hani Marcus**, WEISS Deputy Director for Translation
6. **Rachael Rodell**, WEISS Quality Assurance Manager
7. **Simon Watt**, WEISS Public Engagement Manager
8. **Phil Wiles**, WEISS Public Engagement Coordinator
9. **Michelle McGrath**, WEISS Communication Manager
10. **Rawen Kader**, Clinical Research Fellow, WEISS
11. **Hannah Spiro**, Head of Public Attitudes, Centre for Data Ethics and Innovation
12. **Aleksandra Herbec**, Scientific Secretary and Head of Secretariat for Independent Review on Equity in Medical Devices, Department of Health and Social Care
13. **Nathalie Carter**, Patient and Public Involvement Manager, NHS England and NHS Improvement
14. **Julian Louis**, Project Support Officer, NHS England and NHS Improvement
15. **Sandra Nwokeoha**, Manager, AI Award, NHS England and NHS Improvement
16. **Georgios Onisiforou**, Research Manager, NHS AI Lab
17. **Nisha Shah**, Researcher in Social Science, Centre for Health, Law and Emerging Technologies (HeLEX) at University of Oxford
18. **Xiaoxuan Liu**, Clinical Researcher, Institute of Inflammation and Ageing
19. **Joe Alderman**, AI and Digital Health Clinical Research Fellow, Institute of Inflammation and Ageing
20. **Ellen Coughlan**, Programme Manager, Health Foundation
21. **Lara Grooves**, Researcher, Ada Lovelace Institute
22. **Pritesh Mistry**, Fellow, Digital Technologies, The King's Fund
23. **Lesley Booth**, Director of Research and Patient & Public Involvement/Engagement, Bowel Research UK
24. **Urte Macikene**, Policy and External Affairs Manager, Healthwatch England
25. **Rebecca Moore**, Senior Policy Analyst, Healthwatch England
26. **Dee Davison**, Public Engagement Manager, University of Edinburgh
27. **Robin Carpenter**, Senior Research Data Governance Manager, London Medical Imaging & Artificial Intelligence Centre for Value Based Healthcare
28. **Marzena Nieroda**, Assistant Professor, UCL's Global Business School for Health
29. **Melanie Smallman**, Associate Professor, UCL's Dept of Science & Technology Studies
30. **Cian O'Donovan**, Researcher UKRI Ethics Accelerator grant, UCL
31. **James Baggaley**, Head of Communications and Engagement, UCL Policy Lab
32. **Michelle Miller**, Research Manager, Newcastle University
33. **Ben Allcock**, APPG Inquiry Lead on Medical Research, Association of Medical Research Charities
34. **Anna Grey**, Impact strategy consultant, Cornish+Grey
35. **Charlotte Cornish**, Impact strategy consultant, Cornish+Grey

This briefing was developed with Dr Luis Lacerda at UCL Engineering's Policy Impact Unit.

Find out more at [www.ucl.ac.uk/steapp/PIU](http://www.ucl.ac.uk/steapp/PIU) or email us on: [PolicyImpactUnit@ucl.ac.uk](mailto:PolicyImpactUnit@ucl.ac.uk)

This work was funded by Wellcome grant number 203145/Z/16/A.