

Developing effective policy to support Artificial Intelligence in health and care

Jessica Morley^{1,2} and Indra Joshi²

1. Oxford Internet Institute, University of Oxford, 1 St. Giles OX1 3JS
2. NHSX, Skipton House, 80 London Road, London, SE1 6LH

Abstract: The increased availability of data has enabled the development of Artificially Intelligent Systems (AIS) for health, but implementing these systems and capitalising on the associated opportunities is not straightforward. To mitigate these risks, outdated governance mechanisms need to be updated and key questions answered. To achieve this, whilst still supporting innovation, a new joint organisation for digital, data and technology in the English NHS (NHSX) is developing a ‘principled proportionate governance’ model that involves focusing on proactively and objectively evaluating current AIS technology and regularly involving all those who rely on and serve the health and care system.

Keywords: Digital health, Artificially Intelligent Systems, Machine Learning, NHS, Health Policy

Introduction

Compared to other industries, it might appear that health care has been slow to respond to the ‘data revolution’¹. However, over the last few years this has changed significantly. The increased availability of data has enabled the development of Artificially Intelligent Systems (AIS) for diagnostics, drug discovery, public health and epidemiology, operational efficiency, and ‘P4’ (predictive, preventive, personalised and participatory) Medicine^{2 3 4}.

Implementing these systems and capitalising on the associated opportunities is, however, not straightforward. The fall-out from the care.data scheme¹ demonstrates the challenge we face in trying to find a balance between investing in the development of data-driven technologies that have the potential to improve the quality of health and care services, whilst respecting ethical values such as autonomy, transparency, confidentiality and privacy⁵. Poorly designed AIS for health could result in significant harm to individuals, groups, or society, ultimately resulting in a loss of public trust.

To mitigate these risks, outdated governance mechanisms (policies, standards and regulations) need to be updated and key questions answered in terms of: liability in cases of medical error; doctors’ and patients’ understanding and control over how AIS produce predictions or recommendations that are used in treatment plans; and access to and protection of patient data.⁶ In order to do this, whilst still supporting

¹ Care.data was a programme intended to create a national database of patients’ medical that could be used for the purposes of research and ensuring consistency of care across primary and secondary care. However, it failed to win clinician and patient trust and was shut down in 2016.

innovation that can deliver genuine system and patient benefit, NHSX (the central governing arm of the English NHS related to digital transformation) is developing a ‘principled proportionate governance’ model focused on: (1) guiding principles and best practice; and (2) safe, effective and proportionate regulation⁷. This involves focusing on proactively and objectively evaluating current AIS technology to ensure best practices are developed and implemented in an evidence-based manner⁸ and with regular involvement from all those who rely on and serve the health and care system.

We hope that by adopting this approach we can encourage the development of AIS in a way that secures patient and health care practitioner trust, serves the public interest, and strengthens shared social responsibility⁹. What follows are some examples of how NHSx is doing this. The intention in outlining these is not to imply that we have ‘solved’ the problem of how to foster the development and implementation of AIS in health and care in a safe, ethical and robust manner, but to provide examples of what can be done and to encourage an open and collaborative approach to policymaking.

Guiding principles and best practice

First, building on existing frameworks, such as the Department of Digital, Culture, Media and Sport *Data Ethics Framework*¹⁰, is the Code of Conduct for Data-Driven Health and Care Technology. The Code aims to promote the development of AIS for health and care in accordance with the Nuffield Council on Bioethics’ principles for data initiatives (i.e. respect for persons, respect for human rights, participation, accounting for decisions¹¹) by clearly setting out the behaviours that the central governing organisations of the NHS, expect from those developing, deploying and using AIS for health and care.

We developed the Code, which is a series of 10 principle behaviours, using a Delphi methodology and published the first draft on 5 September 2018 along with a questionnaire for members of the public to offer feedback. We combined the feedback gleaned from this survey with that gathered from an extensive period of face-to-face engagement with industry experts, academics, regulators and patient representative organisations over the last quarter of 2018, to produce a revised version of the Code in February 2019, which has since been updated (see Figure 1).

Figure 1: The 10 Principles in the Code of Conduct for Data-Driven Health and Care Technologies, May 2019

1	Understand users, their needs and the context
2	Define the outcome and how the technology will contribute to it
3	Use data that is in line with appropriate guidelines for the purpose for which it is being used
4	Be fair, transparent and accountable about what data is being used

5	Make use of open standards
6	Be transparent about the limitations of the data used
7	Show what type of algorithm is being developed, or deployed, the ethical examination of how the data is used, how its performance will be validated, and how it will be integrated into health and care provision
8	Generate evidence of effectiveness for the intended use and value for money
9	Make security integral to the design
10	Define the commercial strategy

Source: ¹²

For the most part, these principles reflect behaviours that are already required by regulation, such as the Data Protection Act 2018, or existing NHS guidance, such as the NHS Digital Design Manual. However, principles 7 and 8 (and 10 although we do not discuss this here), are entirely new and require further supporting policy work.

For principle 8, we worked with the National Institute for Health and Care Excellence (NICE), Public Health England, and Med City (the life sciences sector cluster organisation for the Greater South East of England), to create the *Evidence Standards Framework for Digital Health Technologies*.¹³ The framework establishes the evidence of effectiveness and economic impact required before digital health interventions can be deemed appropriate for adoption by the health and care system. In keeping with our principled proportionate approach, the framework is based on a hierarchical classification determined by the functionality (and associated risk) of the tool, which indicates the level of evidence required so that a more complex tool (such as one providing diagnosis) requires considerably more evidence than one simply communicating information¹⁴.

Following from this, for principle 7, NHSx are currently working with a number of think tanks, academic, industry and patient groups, to create a 'how-to' guide. The guide takes the form of a set of processes that NHSx will encourage developers to undertake. The processes are divided into two: i) recommendations for general processes that apply across all aspects of principle 7; and ii) recommendations for specific processes that apply to certain subsections (Figure 2).

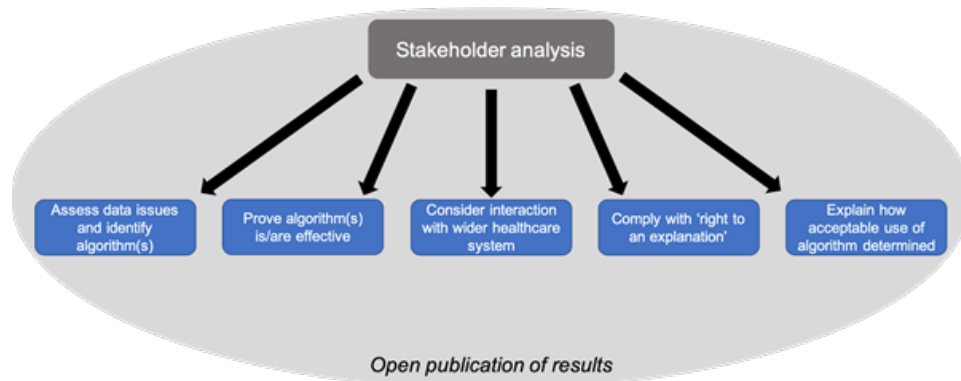


Figure 2: A schematic outlining the different components of the principle 7 ‘how-to’ guide
Produced by: Matt Fenech, Olly Bostrum and Nika Strukelj of Future Advocacy

In both cases, our intention is to make it very clear to developers not only *what* we expect them to do in order to develop AIS for use in health and care, but *how* they might go about doing it. This is because, as has been pointed out several times recently^{15 16}, ethical and behavioural principles are necessary but not sufficient to ensure the design and practical implementation of responsible AI. Indeed, this is why NHSx have turned the entire Code into a self-assessment workbook, which much like the assessment list included in the European Commission’s Framework for Trustworthy AI¹⁷, aims to operationalise responsible AIS for health and care.

NHSx are currently piloting the workbook with a series of public and private companies to assess how it can be tailored in a proportionate manner, where it needs further development and where, like with principles 7 and 8, NHSx might need to provide more detailed guidance on translating between the *what* and the *how*. Once it has been tested, NHSx will be able to embed it into existing assessment processes, such as the Government Digital Service spend controls process, and the Digital Assessment Questionnaire for the NHS Apps Library, to provide us with a mechanism for conducting algorithmic assurance.

Safe, effective and proportionate regulation

NHSx believe that as the Code of Conduct, and its associated guidance, is strengthened over time will significantly improve the governability of AIS developed, deployed and used within the health and care sector in the UK. Nevertheless, NHSx also believe that this will not provide sufficient protection for higher risk systems. For these systems, harder governance mechanisms will be needed. To develop these, NHSx are conducting two parallel processes: (a) assessing the supply of and demand for AIS and (b) strengthening the existing regulatory framework as part of an ongoing process involving all of the health and care regulators (Information Commissioner’s Office, Care Quality Commission, Medicines and

Healthcare products Regulatory Agency, Health Research Authority) and the Better Regulation Executive, to assess the gaps in the existing framework and transition it to one of ‘regulation as a service.’²

First, to assess the supply of AIS, NHSx are conducting a nationwide survey (available to all in the UK), in collaboration with the Academic Health Science Network AI Initiative, to answer key questions such as: (i) what outcomes are developers of data-driven technologies expecting to achieve for their identified user(s)?; (ii) where in the system do developers anticipate their data-driven technology to be deployed and how far away from that being a reality do they feel?; (iii) where and how are data-driven technology developers accessing data for training, testing, validation and evaluation?; and (iv) have the resultant models been assessed for possible issues of bias, optimised (in terms of architecture, procedures and outcomes) for fairness and designed with explainability in mind?

Second, to assess demand, NHSx are working with the Academy of Medical Royal Colleges, think tanks and social research groups, to conduct qualitative research with staff in hospital trusts to understand: (i) the primary issues they face on a daily basis that may prevent them from delivering the high standard of care NHSx expect; and (ii) the barriers that might be preventing them from using an available AIS to overcome relevant issues. NHSx hope that the results from these two projects will give us the evidence NHSx require to determine where to invest next in the development of AIS and which areas of regulation NHSx should focus on strengthening first.

Conclusion

The above provides a brief and non-exhaustive snapshot of the work that the central governing organisations of the NHS are leading on to develop effective policy to support Artificial Intelligence in health and care. The constantly developing nature of AIS means that these will not be one-off exercises, but part of an ongoing programme of work. Nor is it work that NHSx have been, or are planning to, conduct alone. The implications of the implementation of AIS in health and care are so significant, that close collaboration with our regulators, with innovators, with patients (who must be seen as part of the solution, not a problem to be overcome¹⁸, with commissioners, with policymakers and with those on the frontline will continue to be essential if NHSx are to successfully embed the values that matter to all voices in the NHS into AIS from the very beginning. NHSx will continue to regularly assess whether NHSx are striking the right balance between supporting innovation and protecting patient safety whilst creating a trusted environment that is in alignment with the NHS Constitution.

² This particular programme of work is led by Adrian Price in the Department of Health and Social Care.

Bibliography

- ¹ Oborn E, Barrett SK. Digital health and citizen engagement: Changing the face of health service delivery. *Health Services Management Research* 2016;29(1–2):16–20. <https://doi.org/10.1177/0951484816637749>
- ² De Fauw J, Ledsam JR, Romera-Paredes B, et al. Clinically applicable deep learning for diagnosis and referral in retinal disease. *Nature Medicine*; 24(9):1342–50. <https://doi.org/10.1038/s41591-018-0107-6>
- ³ Fleming N. How artificial intelligence is changing drug discovery. *Nature* 2018;557(7707):S55–7. <https://doi.org/10.1038/d41586-018-05267-x>
- ⁴ Velasco E. Disease detection, epidemiology and outbreak response: the digital future of public health practice. *Life Sciences, Society and Policy* 2018;14(1). <https://doi.org/10.1186/s40504-018-0071-4>
- ⁵ Sterckx S, Rakic V, Cockbain J, Borry P. “You hoped we would sleep walk into accepting the collection of our data”: controversies surrounding the UK care.data scheme and their wider relevance for biomedical research. *Medicine, Health Care and Philosophy* 2016;19(2):177–90. <https://doi.org/10.1007/s11019-015-9661-6>
- ⁶ Ngiam KY, Khor IW. Big data and machine learning algorithms for health-care delivery. *The Lancet Oncology* 2019;20(5):e262–73. [https://doi.org/10.1016/S1470-2045\(19\)30149-4](https://doi.org/10.1016/S1470-2045(19)30149-4)
- ⁷ Sethi N, Laurie GT. Delivering proportionate governance in the era of eHealth: Making linkage and privacy work together. *Medical Law International* 2013;13(2–3):168–204. <https://doi.org/10.1177/0968533213508974>
- ⁸ Mittelstadt, B. (2019). *AI Ethics - Too Principled to Fail?* Retrieved from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3391293
- ⁹ Floridi L, Cowls J, Beltrametti M, et al. AI4People—An Ethical Framework for a Good AI Society: Opportunities, Risks, Principles, and Recommendations. *Minds and Machines* 2018;28(4):689–707. <https://doi.org/10.1007/s11023-018-9482-5>
- ¹⁰ Department for Digital, Culture, Media & Sport. *Data Ethics Framework*, 2018. Available at: <https://www.gov.uk/government/publications/data-ethics-framework/data-ethics-framework>
- ¹¹ Nuffield Council on Bioethics. The collection, linking and use of data in biomedical research and health care: ethical issues, 2015. Available at: <http://nuffieldbioethics.org/wp-content/uploads/Biodata-a-guide-to-the-report-PDF.pdf>
- ¹² Department of Health and Social Care. *Guidance Code of conduct for data-driven health and care technology*, 2019. Available at: <https://www.gov.uk/government/publications/code-of-conduct-for-data-driven-health-and-care-technology/initial-code-of-conduct-for-data-driven-health-and-care-technology>
- ¹³ NICE. *Evidence standards framework for digital health technologies*, 2019. Available at: <https://www.nice.org.uk/about/what-we-do/our-programmes/evidence-standards-framework-for-digital-health-technologies>
- ¹⁴ Greaves F, Joshi I, Campbell M, Roberts S, Patel N, Powell J. What is an appropriate level of evidence for a digital health intervention? *The Lancet* 2018;392(10165):266–7. [https://doi.org/10.1016/S0140-6736\(18\)33129-5](https://doi.org/10.1016/S0140-6736(18)33129-5)
- ¹⁵ Mittelstadt, B. (2019). The Ethics of Biomedical ‘Big Data’ Analytics. *Philosophy & Technology*, 32(1), 17–21. <https://doi.org/10.1007/s13347-019-00344-z>
- ¹⁶ Morley J, Floridi L, Kinsey L, Elhalal A. From What to How. An Overview of AI Ethics Tools, Methods and Research to Translate Principles into Practices, 2019. *ArXiv:1905.06876 [Cs]*. Retrieved from <http://arxiv.org/abs/1905.06876>
- ¹⁷ European Commission. *Ethics Guidelines for Trustworthy AI*, 2019. Available at: <https://ec.europa.eu/futurium/en/ai-alliance-consultation>
- ¹⁸ Aitken M, Tully MP, Porteous C, et al. Consensus Statement on Public Involvement and Engagement with Data-Intensive Health Research. *International Journal of Population Data Science* 2019;4(1). <https://doi.org/10.23889/ijpds.v4i1.586>