UNIVERSITY
OF WOLLONGONG
AUSTRALIA

Response to public consultation

Date: 13/10/2024

Subject: **Response to "Safe and Responsible Artificial Intelligence in Health Care – Legislation and Regulation Review"**

To: Department of Health and Aged Care

From: Dr Yves Saint James Aquino, Australian Centre for Health Engagement, Evidence and Values, University of Wollongong

Prof Stacy Carter, Australian Centre for Health Engagement, Evidence and Values, University of Wollongong

Prof Farah Magrabi, Australian Institute of Health Innovation, Macquarie University

CONTEXT/BACKGROUND

We are honoured to respond to the Department of Health and Aged Care's public consultation on *Safe and Responsible Artificial Intelligence (AI) in Health Care – Legislation and Regulation Review*[1], which seeks submissions to consider the benefits and risks of AI and potential regulatory changes across Australia's healthcare system.

Our responses are based on an ongoing program of research that examines ethical, social, quality and safety issues in healthcare applications of AI. Specifically, we draw on three documents we have provided as attachments. First, we draw on the published recommendations from a national citizens' jury on healthcare AI.[2] Second is a 2024 report *AI Implementation in Hospitals: Legislation, Policy, Guidelines and Principles, and Evidence about Quality and Safety*, which provides the findings of a literature review and environmental scan we conducted for the Australian Commission for Safety and Quality in Health Care (ACSQHC).[3] Finally, we draw from the Australian Alliance for AI in Healthcare's (AAAiH) *National Policy Roadmap for AI in Healthcare.*[4]

We've constructed our responses to the 19 questions under four themes that our research actively engages with: a) what makes healthcare AI different; b) risks; c) accountability and liability; and d) consent, privacy and transparency.

OUR RESPONSE

a. What makes AI in healthcare different?

Questions 6-9 and 15 speak to the notion of AI being exceptional in some ways. This can be understood as AI being different than other medical technologies, or healthcare AI being different than AI applied to other industries.

We draw attention to three of the ways healthcare is exceptional.

---

[1] https://consultations.health.gov.au/medicare-benefits-and-digital-health-division/safe-and-responsible-artificial-intelligence-in-he/

[2] Carter, S. M., Aquino, Y. S. J., Carolan, L., Frost, E., Degeling, C., Rogers, W. A., ... & Magrabi, F. (2024). How should artificial intelligence be used in Australian health care? Recommendations from a citizens' jury. Medical Journal of Australia, 220(8), 409-416. https://doi.org/10.5694/mja2.52283

[3] https://www.safetyandquality.gov.au/publications-and-resources/resource-library/ai-implementation-hospitals-legislation-policy-guidelines-and-principles-and-evidence-about-quality-and-safety

[4] https://aihealthalliance.org/wp-content/uploads/2023/11/AAAiH_NationalPolicyRoadmap_FINAL.pdf

First, healthcare involves language, norms and expertise that are specific to healthcare practice. There is real danger that AI developers are not paying attention to these particularities when repurposing general purpose AI technologies to healthcare. For example, healthcare is rightly held to highly formalised and mandatory ethical, safety and effectiveness standards. This means that, unlike in some other industries, it is not acceptable to deploy untested AI—especially general-purpose AI—and mitigate risks after deployment.

Second, healthcare arguably has a more complex governance ecosystem than other industries such as finance or manufacturing. Healthcare governance involves e.g. regulation of health practitioners and clinicians, accreditation of institutions, medical device and pharmaceutical regulation, and health laws. All of these components of healthcare regulation must respond to the challenges of AI.

Third, healthcare is a 'high stakes domain', as acknowledged by the Department of Industry, Science and Resources' *Safe and Responsible AI in Australia* consultation report[5] and in the *International Scientific Interim Report* delivered from the Bletchley AI Safety Summit.[6] Decades of reasoning in bioethics suggests that health has special moral importance. This is both because we value our health (that is, health is a good in itself), and because being healthy enables us to access other opportunities.[7] This special moral importance of health means changes to healthcare provision, such as the introduction of transformative and potentially disruptive technologies, requires particular attention.

We endorse the prescient summary of the exceptional characteristics of AI technologies in the *Introducing mandatory guardrails for AI in high-risk settings: proposals paper*.[8] These characteristics are particularly salient in a healthcare setting, where the autonomy and thinking and explanation skills of human clinicians are critically important. The adaptability and opacity of machine learning models makes them potentially less amenable to the tenets of evidence-based practice that have previously assured quality and safety in healthcare. The speed and scale of AI technology means any harms caused by AI systems could be greatly amplified relative to the harm that might currently be done by variations in practice or negligence by individual clinicians. These factors together risk undermining public trust in healthcare systems, and make maintenance of high standards of evidence and effectiveness even more important as we move to widespread AI implementation. This need for high-quality evidence and ongoing monitoring and evaluation were key concerns of the Citizens' Jury.[9]

We support an integrated approach to healthcare AI governance rather than an approach that relies on creating new, AI-specific laws. Our environmental scan[10] shows that healthcare is already a highly regulated space. An integrated approach builds on existing governance processes within the healthcare system and entails explicit clarification of how these apply to AI.

In addition, an integrated approach requires clarification of how cross-sectoral governance mechanisms apply to healthcare AI. For example, appeals to ethics frameworks are common across sectors when developing guidance for AI implementation, in both Australian and international jurisdictions. In Australia, key ethics frameworks include the *Australia's AI Ethics Principles*[11] and the NSW Government's AI Ethics Policy[12] (embedded in the *NSW AI Assurance/Assessment Framework*[13]). These frameworks provide common principles, such as fairness and privacy, that decision makers are asked to pay attention to across different stages of the AI lifecycle. These principles tend to be abstract, and work is needed to guide stakeholders in how to apply them in specific health contexts.

Although we strongly advocate for an integrated approach where AI considerations are embedded into existing governance mechanisms, we also note that this risks fragmentation and inconsistency. To work against these risks, an Australian body specifically dedicated to overseeing AI in healthcare would ensure an integrated approach to healthcare AI governance. The AAAiH Roadmap recommends a National AI in Healthcare Council to "better coordinate and harmonise the responsibilities and activities of those entities

---

[5] Australian Government Department of Industry Science and Resources (2024). Safe and Responsible AI in Australia: Proposals paper for introducing mandatory guardrails for AI in high-risk settings https://consult.industry.gov.au/ai-mandatory-guardrails.
[6] Department for Science, Innovation and Technology and AI Safety Institute, International Scientific Report on Safety of Advanced AI: Interim Report, UK Government, 2024, page 12.
[7] Daniels, N. (2008). Just Health: Meeting Health Needs Fairly. Cambridge, Cambridge University Press.
[8] Reference 5, p. 11.
[9] Reference 2, p. 6-7., Recommendations 3, 8, 9, 10, 13, 14.
[10] Reference 3, Section 3.3.1
[11] https://www.industry.gov.au/publications/australias-artificial-intelligence-ethics-framework/australias-ai-ethics-principles
[12] https://www.digital.nsw.gov.au/policy/artificial-intelligence/artificial-intelligence-ethics-policy
[13] https://www.digital.nsw.gov.au/policy/artificial-intelligence/nsw-artificial-intelligence-assessment-framework

responsible for oversight of AI safety, effectiveness, and ethical and security risks".[14] Similarly, the Australian citizens' jury recommended a charter for AI in healthcare, and an independent decision-making body tasked to manage the charter.[15]

b. Risks to patients/consumers, healthcare professionals, and marginalised groups

This theme is relevant to questions 2, 3 and 4.

The risk of algorithmic bias is well established. Algorithmic bias refers to the problem where AI systems perform poorly for some individuals or groups—typically those who are marginalised or underserved. One of the main causes of algorithmic bias is imbalanced datasets that exclude data from marginalised and underserved communities. The citizens' jury recommended that AI training datasets should be "adequately representative and inclusive to capture Australia's multiculturalism and diversity".[16] The National Roadmap particularly emphasised the need for culturally safe and trusted health-related data collation, in line with principles of Indigenous Data Sovereignty.[17]

AI can pose serious risks to patient safety. Our report for ACSQHC outlines emerging evidence of safety concerns associated with AI systems and their impact on patient care.[18] Studies describing AI implementation reported lack of quality assurance measures, including post-deployment monitoring, auditing, or performance reviews. AI implementation in healthcare should be supported by a strong evidence base, strong ethics review, and suitable research governance to ensure the evidence base for AI use in healthcare is built via ethical, responsible, high-quality and safe research. Our research program includes ongoing work on how to assure adequate governance for AI research in health.

Another risk to patient safety associated with AI is the potential for misuse due to lack of skills and training. The ACQSHC Report identified the need for tailored training modules that cover AI functionalities, limitations, potential biases, and ethical considerations.[19] Similarly the National Australian Citizens' Jury repeatedly emphasised the need for adequate training and support for health workers to use AI responsibly, including guidance from professional bodies such as Colleges and other bodies that regulate healthcare workers (e.g. the Australian Health Professional Regulation Agency).[20] The National Policy Roadmap calls for a shared code of conduct and the development of profession-specific codes of practice for the responsible use of AI.[21]

c. Accountability and liability

This theme is relevant to questions 11, 12 and 13.

Findings of the environmental scan in the ACQSHC Report indicated concern among government agencies in Australia and overseas about the lack of clarity in establishing legal liability—understood as apportioning blame when errors occur. There is a general agreement that in areas such as health, liability must lie with a natural or legal person if the AI system is used as an aide.[22] Where an AI system acts autonomously, legal and policy reform may be needed to determine, in respect of liability, the legally defined actors in the AI system lifecycle (from development to implementation).

Several organisations have proposed a collective responsibility model, in which all actors involved in the development and deployment of AI are held responsible in order to promote integrity. Examples include the WHO's *Ethics and governance of artificial intelligence for health: Guidance on large multi-modal models,*[23] and the US Government Accountability Office's *Artificial Intelligence: An Accountability Framework for Federal Agencies and Other Entities.*[24]

---

[14] Reference 4, p. 4. Recommendation 1.
[15] Reference 2. p. 6. Recommendations 1-2.
[16] Reference 2. p. 11. Recommendation 11.
[17] Reference 4. p. 4. Recommendation 9.
[18] Reference 3. p. 13 and 118. Key finding 5.
[19] Reference 3. p. 113-114. Worker training and support.
[20] Reference 2. p. 6, Recommendations 6-7.
[21] Reference 4. p. 4, Recommendations 6-7
[22] Reference 3. p. 112-113.
[23] https://www.who.int/publications/i/item/9789240084759
[24] https://www.gao.gov/products/gao-21-519sp

d. Consent, privacy and transparency

This theme is relevant to questions 16-18. In the ACQSHC Report's environmental scan, we found different conceptions of transparency, namely: a) transparency to support meaningful patient/consumer consent; b) transparency of using AI in patient care; c) transparency in data use to ensure compliance with privacy and data protection obligations; and d) transparency with respect to governance, including process to monitor and audit AI systems.[25]

For consent, findings from the report suggest two separate consent obligations: 1) patient consent to use AI in their care and 2) patient consent for their health data being used to develop AI by healthcare organisations and third-party developers, especially for commercial purposes. Key challenges in complying consent requirements include stakeholders not being able to anticipate all ways their data will be used in the future, upholding privacy obligations while ensuring representative datasets, and lack of guidance how to implement consent policies in clinical practice.

Further work is required to establish a national consensus that would provide guidance on consent, privacy and transparency. One of the recommendations of Australian citizens' jury on AI is that there should be a comprehensive and fully funded community education program to promote digital health literacy to ensure guidance on AI issues is consistent with community needs and values.[26]

CONCLUSION

AI implementation in healthcare requires an integrated approach to clarify how existing governance mechanisms and regulatory processes apply to healthcare applications of AI technologies. Legal and policy analyses and reform, supported by robust community and consumer engagement, are needed to address concerns about accountability and liability, consent, privacy and transparency. We are grateful for the opportunity to provide a submission to this consultation, and would be pleased to expand on any of the above if useful to the Department's deliberations.

Sincerely,

Dr Yves Saint James Aquino

Prof Stacy M. Carter

Prof Farah Magrabi

[25] Reference 3 p. 34.
[26] Reference 2, p. 7. Recommendation 15

LIST OF ATTACHMENTS

1. Report of the National Australian Citizens' Jury on AI in Healthcare

Citation: Carter SM, Aquino YSJ, Carolan L, Frost E, Degeling C, Rogers WA … Magrabi F. How should artificial intelligence be used in Australian health care? Recommendations from a citizens' jury. Med J Aust. 2024;220(8):409-16.

Available from: https://onlinelibrary.wiley.com/doi/full/10.5694/mja2.52283

2. Literature Review and Environmental Scan Report conducted for the Australian Commission on Safety and Quality in Healthcare

Citation: Magrabi F, Bates L, Brooke-Cowden K, Jayawardena T, Wang A, Coiera E, Aquino YSJ, Frost E, Carter S, Adams C, Almohanna A. AI Implementation in Hospitals: Legislation, Policy, Guidelines and Principles, and Evidence about Quality and Safety. Australian Commission on Safety and Quality in Healthcare, 2024.

Available from: https://www.safetyandquality.gov.au/sites/default/files/2024-08/artificial_intelligence_-_literature_review_and_environmental_scan.pdf

3. A National Policy Roadmap for Artificial Intelligence in Healthcare

Citation: Australian Alliance for Artificial Intelligence in Healthcare. A National Policy Roadmap for Artificial Intelligence in Healthcare. AAAiH, 2023.

Available from: https://aihealthalliance.org/wp-content/uploads/2023/11/AAAiH_NationalPolicyRoadmap_FINAL.pdf