UNIVERSITY OF WOLLONGONG AUSTRALIA

Submission to the Select Committee on Adopting Artificial Intelligence (AI)

Date: 10 May 2024

Prepared by: Dr. Yves Saint James Aquino, Emma Frost, and Prof Stacy Carter

Australian Centre for Health Engagement, Evidence and Values, University of Wollongong; https://www.uow.edu.au/the-arts-social-sciences-humanities/research/acheev/

Contact details: yaquino@uow.edu.au; +612 4221 3294; Level 3, Building 29, University of Wollongong, Northfields Avenue, Wollongong, 2522, NSW Australia

---

We welcome the opportunity to respond to the Senate Committee Inquiry on Adopting Artificial Intelligence (AI), which aims to inquire into and report on the opportunities and impacts for Australia arising out of the uptake of AI technologies in Australia.

OUR WORK ON HEALTHCARE AI

The Australian Centre for Health Engagement, Evidence and Values (ACHEEV) includes a team of researchers focused on *Data, AI and other health technologies*.[1] Our work focuses on the ethical, legal and social implications of the use of machine learning in healthcare.[2,3,4,5,6,7,8]

We particularly draw the Senate's attention to the first national citizens' jury held worldwide on artificial intelligence in healthcare,[8] which ACHEEV ran in 2023. This robust process used best-practice methods to randomly select diverse Australians to learn about AI from experts and then work together over 18 days, including three days face to face, to make recommendations on the following question: "Under what circumstances, if any, should artificial intelligence be used in Australian health systems to detect or diagnose disease?" The jurors generated 15 recommendations, which are already informing policymaking processes at state and national levels. The report from this jury, published in the Medical Journal of Australia, is attached for your consideration.

Our response addressing considerations from the inquiry's terms of reference is based on our findings across our studies. We focus on health contexts as this is our area of specialisation. While the Terms of Reference for the Senate Inquiry notes the potential benefits of healthcare AI, we note also that health contexts demonstrate the urgency of questions regarding governance. Healthcare and public health, while potentially offering pathways to benefit, are also high-risk and high-stakes areas for any application of AI and

---

[1] Our healthcare AI research has been funded by NHMRC Ideas Grant 1181960, NHMRC CRE 2006545, and grants from the University of Wollongong.

[2] "Facilitating public involvement in research about healthcare AI: A scoping review of empirical methods." *International Journal of Medical Informatics* (2024): https://doi.org/10.1016/j.ijmedinf.2024.105417.

[3] "Measures of Socioeconomic Advantage Are Not Independent Predictors of Support for Healthcare Ai: Subgroup Analysis of a National Australian Survey." *BMJ Health & Care Informatics* (2023). https://informatics.bmj.com/content/30/1/e100714

[4] "The Adoption of Artificial Intelligence in Health Care and Social Services in Australia." *Journal of Medical Internet Research* (2022). https://doi.org/10.2196/37611.

[5] "Practical, Epistemic and Normative Implications of Algorithmic Bias in Healthcare Artificial Intelligence: A Qualitative Study of Multidisciplinary Expert Perspectives." *Journal of Medical Ethics* (2023). https://jme.bmj.com/content/early/2023/02/22/jme-2022-108850

[6] "Utopia Versus Dystopia: Professional Perspectives on the Impact of Healthcare Artificial Intelligence on Clinical Roles and Skills." *International Journal of Medical Informatics* (2023). https://doi.org/10.1016/j.ijmedinf.2022.104903.

[7] Australian women's judgements about using artificial intelligence to read mammograms in breast cancer screening. *Digital Health* (2023), https://doi.org/10.1177/20552076231191

[8] "How should artificial intelligence be used in Australian health care? Recommendations from a citizens' jury." *Medical Journal of Australia* (2024), 220. https://doi.org/https://doi.org/10.5694/mja2.52283

automated decision-making (ADM). In addition, Australia is behind relative to other countries in developing both technology and strategies for safe and responsible implementation of healthcare AI and ADM.[9]

KEY COMMENTS ADDRESSING THE TERMS OF REFERENCE

a. Recent trends and opportunities in the development and adoption of AI technologies in Australia and overseas, in particular regarding generative AI

Consideration of the adoption of AI in healthcare raises questions around the governance required for such adoption to be legitimate and justified.

Australia has lagged the world in both development and implementation of healthcare AI. This is rapidly shifting, with a wide range of stakeholders recognising that the Australian health sector needs to urgently institute risk-based, ethically-informed and coordinated approaches to governance and implementation. In particular, the health sector has recognised the need for consensus-building and participatory governance that includes consumers and citizens. The Australian Alliance on Artificial Intelligence in Health (AAAiH) recently launched a policy roadmap[10] developed through collaboration with a wide range of stakeholders, establishing priority areas and key actions to drive this governance and implementation agenda. The Roadmap process recommended establishment of a National AI in Healthcare Council, led by Government given the legislative responsibilities of Government Agencies with respect to AI, and with broad membership. This is an essential piece of governance architecture for health, but is also a model that could be applicable in other sectors.

b. Risks and harms arising from the adoption of AI technologies, including bias, discrimination and error

Literature in AI ethics have raised numerous risks and harms. In this response, we are focussing on three sets of issues, namely bias and discrimination, safety and evidence base, and deskilling and dehumanisation.

Bias and discrimination

Risk of bias and discrimination are dominant concerns in healthcare AI. Without robust intervention, AI systems will disadvantage already under-represented and marginalised groups that could systematically worsen existing inequity in healthcare system.

Despite the growing evidence of bias in AI, the moral and political seriousness of the problem remains contested within data science and AI communities. Our interview study[11] with healthcare AI experts helps explain this. We found that:

1. a small number of experts did not consider bias a real problem in healthcare (which we argue is incorrect based on evidence);
2. experts disagreed about who is responsible in mitigating bias, with some developers and data scientists considering bias to be a result of complex injustices that were beyond their responsibility; and
3. experts disagreed about whether complex social identities (race, ethnicity, gender, sexuality) subject to social injustice and health inequities should be included in datasets used to train AI.

This suggests the importance of acknowledging the reality of AI bias, and putting governance structures in place that clearly allocate responsibility for addressing it.

In Australia, a number of healthcare organisations have made recommendations for how bias should be addressed in Australian healthcare, mostly via evaluation and monitoring.[12,13] Evaluation includes:

---

[9] "We Need to Chat About Artificial Intelligence." *Medical Journal of Australia* (2023). https://doi.org/10.5694/mja2.51992.
[10] https://aihealthalliance.org/wp-content/uploads/2023/11/AAAiH_NationalPolicyRoadmap_FINAL.pdf
[11] "Practical, Epistemic and Normative Implications of Algorithmic Bias in Healthcare Artificial Intelligence: A Qualitative Study of Multidisciplinary Expert Perspectives." *Journal of Medical Ethics* (2023). https://jme.bmj.com/content/early/2023/02/22/jme-2022-108850
[12] "Ethical Principles for AI in Medicine" (2023). https://www.ranzcr.com/college/document-library/ethical-principles-for-ai-in-medicine
[13] "A National Policy Roadmap for Artificial Intelligence in Healthcare" (2023). https://www.mq.edu.au/__data/assets/pdf_file/0005/1281758/AAAiH_NationalAgendaRoadmap_20231122.pdf

- Systematically scrutinising the training data used to develop AI tools to ensure it is representative of the populations for which the tools will be used;
- Testing, and potentially calibrating, AI tools imported from overseas locations on Australian data before deployment; and
- Engaging National Aboriginal Community Controlled Health Organisations (NACCHOs) to ensure that the use of AI benefits the health of Aboriginal and Torres Strait Islander communities.

Bias and fairness was a core consideration for the national Australian citizens' jury, reflected in 8 of 15 recommendations.[14] Jurors emphasised that all people, including people from minority backgrounds, should benefit from AI, that exacerbation of inequity should be avoided, diverse values should be respected in AI development and use, and that training data should be representative. This strong focus on fairness suggests that bias is a core concern of informed Australians, and should be a central consideration in the deliberations of the Select Committee.

Safety and evidence base

While studies have shown some AI systems outperform clinicians in some clinical tasks (e.g., reading radiographic images), there are considerable limitations in providing clinical evidence for safety and effectiveness of these AI systems. Limitations include use of small and highly selected datasets to train AI, limited independent validations, and few real-world trials.[15] With limited validation and real-world trials, it is difficult to establish evidence that the AI system can be properly integrated into a clinical workflow or effectively used by clinicians and healthcare workers. These limitations need to be addressed to ensure that AI systems are both safe and effective to protect welfare of patients and consumers. Further, evidence to support claims made by AI developers regarding model performance requires careful evaluation (as reflected, for example, in the Therapeutic Goods Administration's Software as Medical Device regulation process).[16]

Deskilling and dehumanisation

Our studies show both members of the public and health professionals are concerned about the risk of deskilling of healthcare workers and dehumanisation of healthcare work. In healthcare, deskilling may result in deterioration of clinical skills that compromise decision making across various stages of clinical management, and potentially undermine patient safety. We found that experts were particularly concerned about automation of certain healthcare tasks, especially care tasks that require relational skills and human-to-human contact. Care tasks are a strong feature of healthcare work that is not present in many industries that have embraced AI automation (e.g., banking and finance, manufacturing, and information technology). This demonstrates the need for nuance in the implementation of AI in different industries. Our systematic review of research with members of the public about the use of AI in healthcare demonstrates that people are strongly committed to retaining human decision makers in healthcare. Our own research with Australian women[17] about the use of AI in breast cancer screening revealed some reasons for this: clinicians are understood to have the expertise or skills that machines lack, including intuitive, holistic interpretive capabilities and the ability to deal with unusual decision-making challenges. Both the views of experts and the views of the public should be taken into consideration when implementing AI in healthcare to ensure that healthcare does not become over-automated to the detriment of quality of care.

c. Emerging international approaches to mitigating AI risks

Horizontal and vertical approaches

One way different jurisdictions vary in their regulation of AI to mitigate harms and risk is whether the approach is horizontal or vertical. A horizontal approach to regulation, such as the EU Artificial Intelligence Act, entails that the framework applies to all AI systems regardless of application or industry, generally employing a risk-based approach. A vertical approach to regulation entails creating regulation specific to the type of technology, or in which industry or sector the AI will be implemented. An example of the vertical

---

[14] "How should artificial intelligence be used in Australian health care? Recommendations from a citizens' jury." *Medical Journal of Australia* (2024), 220. https://doi.org/https://doi.org/10.5694/mja2.52283

[15] "AI for mammography screening: enter evidence from prospective trials." The Lancet Digital Health 5. https://doi.org/10.1016/S2589-7500(23)00176-0

[16] https://www.tga.gov.au/how-we-regulate/manufacturing/medical-devices/manufacturer-guidance-specific-types-medical-devices/regulation-software-based-medical-devices

[17] Australian women's judgements about using artificial intelligence to read mammograms in breast cancer screening. *Digital Health* (2023), https://doi.org/10.1177/20552076231191

approach is China's *Interim Measures for the Administration of Generative Artificial Intelligence Service*,[18] the scope of which is limited to providers of generative AI services offered to the public within China. The regulation does not apply to other types of AI technologies.

Any approach to AI regulation and governance would need to interleave with existing regulatory mechanisms, and Australia can benefit from both approaches. A horizontal approach has the benefit of setting standards across all domains and levels of government. Australia has the benefit of a portfolio of prior work by the Australian Human Rights Commission which recommended a range of cross-sectoral interventions, including an independent audit of all Government use of AI to ensure human rights were being upheld. To complement this, policy specific to health sector may be warranted. As noted above, the AAAiH National Policy Roadmap recommended National AI in Healthcare Council. The Australian Citizens' Jury on AI in Healthcare began their recommendations by specifying that there must be a charter for AI in the Australian health system and services, managed by an independent decision-making body. This recommendation was domain specific because the jury was asked to think about health applications of AI.

It is feasible that a broader national charter for AI could harmonise regulatory mechanisms across all sectors or domains. Nonetheless, a horizontal approach should be complemented by sector- or industry-specific regulatory mechanisms. Sectors and industries vary in their norms, relationships and tasks. Sector-specific regulation should take into account perspectives of the professional and public stakeholders within that sector to understand the priorities, assumptions and concerns specific to that sector.

Legislative and non-legislative instruments

Different approaches to regulation range from legislative to non-legislative instruments. Some jurisdictions such as the EU and the US are leading legislative approaches by creating/proposing/enacting new laws to govern AI. Other jurisdictions such as Australia, New Zealand and the UK rely on non-legislative instruments (or soft law) to govern AI, typically in the form of ethics frameworks or principles.

Both legislative and non-legislative instruments play different roles in the governance of AI. A thorough analysis should clarify the strengths and gaps in existing laws that govern technologies, whether new legislation is warranted, or whether non-legislative instruments may be more appropriate.

In healthcare, identifying relevant legislative and non-legislative instruments to govern AI requires coordination among various government agencies responsible for different aspects of healthcare. These agencies include medical device regulators, health ministries or departments, academic institutions, and professional accreditation bodies.

d. Opportunities to adopt AI in ways that benefit citizens, the environment and/or economic growth, for example in health and climate management

Healthcare is a field where the careful implementation and adoption of AI could have meaningful benefits to the community. In many healthcare subspecialties, AI is already being used for certain clinical and administrative tasks.

Many publications, including the Australian Government's Safe and Responsible AI whitepaper,[19] identify and problematise a high level of public distrust in AI. In contrast, our own research on AI in breast screening has found that Australians are supportive of the use of AI, with reservations and conditions.[20] To cultivate trustworthy AI implementation, we advocate for a more active involvement of citizens in AI design, implementation and governance.[21] A more active approach to public engagement would involve citizens in deciding what 'trustworthy' AI looks like, and prioritising the development and deployment of trustworthy AI tools. The benefits of public engagement programs are twofold: programs allow citizens to become more informed about AI and develop a sense of familiarity with otherwise complex and inaccessible technology, and citizens have the opportunity to provide their own perspective, and make recommendations that may have not otherwise been considered.

---

[18] "Do not go gentle into that good night: The European Union's and China's different approaches to the extraterritorial application of artificial intelligence laws and regulations." Computer Law & Security Review 53 (2024): 105965.

[19] https://consult.industry.gov.au/supporting-responsible-ai

[20] Australian women's judgements about using artificial intelligence to read mammograms in breast cancer screening. *Digital Health* (2023), https://doi.org/10.1177/20552076231191

[21] "Facilitating public involvement in research about healthcare AI: A scoping review of empirical methods." *International Journal of Medical Informatics* (2024): https://doi.org/10.1016/j.ijmedinf.2024.105417.

Our Community Jury on AI in healthcare, for example, generated a series of recommendations reflecting how citizens want AI to be implemented in healthcare. These recommendations included,

- clinical training and accreditation to ensure that clinicians are knowledgeable about the use and limits of AI systems;
- the provision of quality information to patients to ensure that they are appropriately informed about when and how AI is used in their care; and
- regular auditing of AI tools to ensure that patient outcomes are being improved by the implementation of AI.

In addition to these recommendations, our evaluation of the jury found that the process improved participants' perceived knowledge of AI, as well as their overall support for the use of AI in healthcare. By the end of the process, all except 2 of the jurors supported or strongly supported the use of AI in healthcare (with 17 jurors increasing the extent of their support compared to baseline). In developing an approach to regulating and managing AI in healthcare, Australia has an opportunity to meaningfully engage citizens in the process of creating frameworks for designing, implementing and monitoring AI.

e. Opportunities to foster a responsible AI industry in Australia

Although AI in Australia is regulated through various technology- and industry-agnostic legislative and non-legislative instruments, it is our position that Australia currently does not have the necessary controls in place to ensure that AI is implemented safely and effectively in healthcare. For example, the clinical use of publicly available generative AI tools like ChatGPT is not formally regulated at present, because it falls outside the remit of the Therapeutic Goods Administration's Software as a Medical Device regulations. This could cause significant harms if clinicians are not sufficiently trained or knowledgeable to understand the risks and their responsibilities in relation to the use of these tools, and has led some states (e.g. Victoria) to introduce a ban on the use of generative AI in their healthcare systems until such time as risks can be governed appropriately. In addition, current regulatory approaches in Australia do not address ways of managing adaptive AI tools, where the model is designed to be updated post-market to 'adapt' to new contexts and circumstances.

Our Community Jury recommended the development of an overarching charter and independent decision-making body to oversee the use of AI in healthcare. Australia could follow the example of New Zealand in developing a charter for the use of AI across all industries.[22] Any governance approach to AI – in healthcare and otherwise – will need to be committed to staying up to date with industry developments to develop proactive regulatory tools.

Governance of AI should consider specific rules for **transparency** about the use of AI. Concern about transparency (and the lack thereof) is a common finding across our consumer-facing studies [23] as well as in the published literature on public views on AI. [24] Transparency is prominent in the EU approach,[25] as well as in local guidelines including the NSW Artificial Intelligence Assurance Framework[26] and the Automated Decision-making Better Practice Guide.[27] Governance of AI should require developers to transparently report certain metrics to describe the design and performance of their tool; transparency requirements should cover any ADM/AI-related recommendation or decision with direct implications to clients or consumers.

In the healthcare context, members of the public have expressed interest in receiving information about:

- How accurate or effective the AI system is at performing tasks
- To what extent a recommendation by an AI system is overseen by a person

---

[22] "Algorithm charter for Aotearoa New Zealand" (2020). https://www.data.govt.nz/toolkit/data-ethics/government-algorithm-transparency-and-accountability/algorithm-charter/

[23] "Australian women's judgements about using artificial intelligence to read mammograms in breast cancer screening". *Digital Health* (in production).

[24] "Facilitating public involvement in research about healthcare AI: A scoping review of empirical methods." *International Journal of Medical Informatics* (2024): https://doi.org/10.1016/j.ijmedinf.2024.105417.

[25] See https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A52021PC0206

[26] See https://www.digital.nsw.gov.au/policy/artificial-intelligence/nsw-artificial-intelligence-assurance-framework

[27] See https://www.ombudsman.gov.au/__data/assets/pdf_file/0029/288236/OMB1188-Automated-Decision-Making-Report_Final-A1898885.pdf

- Which data sources were used to train the AI system
- If and how the AI systems collect and store user data

Transparency should be meaningful and understandable for clients/consumers at point of use. To ensure that all health service users, regardless of literacy levels, can access advice or explanation about AI use, information should be provided at the point of service using plain language. Our colleagues at the Sydney Health Literacy Lab provide strategies (e.g., testing for readability) to make complex and technical information understandable to the public.[28]

Complete banning of certain high-risk AI applications should be considered to protect the safety and welfare of Australians. We note that this approach has been taken in some jurisdictions (e.g., the EU AI Law). Risk evaluations that could lead to the banning of AI come down to people's values and priorities. Thus, engagement with the public will be important to determine what applications should and should not be prohibited in the Australian context.

We are grateful for the opportunity to provide a submission to the Select Committee, and would be pleased to expand on any of the above if useful to the Committee's deliberations.

Sincerely,

Dr Yves Saint James Aquino

Emma Frost

Prof Stacy Carter

---

[28] See https://www.sydneyhealthliteracylab.org.au/tips-and-tricks