

How should artificial intelligence be used in Australian health care? Recommendations from a citizens' jury

Stacy M Carter^{1,2} , Yves Saint James Aquino^{1,2} , Lucy Carolan^{1,2} , Emma Frost^{1,2} , Chris Degeling^{1,2} , Wendy A Rogers³ , Ian A Scott^{4,5} , Katy JL Bell⁶ , Belinda Fabrianesi^{1,2} , Farah Magrabi⁷ 

The known: Artificial intelligence (AI) will transform health care. Guidance regarding its use and governance is urgently needed, and should reflect public expectations about the technology.

The new: In a robust citizens' jury process, a diverse sample of Australian citizens recommended a national charter for health care AI and an independent decision-making body. They also emphasised that rigorous evaluation, fairness and patient rights, clinical governance and training, technical and data requirements, and community education and involvement were also critical areas requiring attention.

The implications: Australians welcome clinical applications of AI, provided that strong governance is in place. A coherent national approach is needed, as well as training, evaluation, and oversight in clinical practice.

In January 2024, the Australian government published its interim response to a consultation on “safe and responsible” artificial intelligence (AI) in Australia.¹ The consultation had the aim of determining how to govern this transformational technology in a manner that preserves public trust, mitigates risk, and supports safe and responsible practices. In clinical care, AI could bring great benefits and serious risks.² Australia currently lags behind other countries in health care AI development, deployment, and governance,³ and health care-specific strategies are needed,^{4,5} as recognised by the Australian Medical Association.⁶

Governance of rapidly emerging health technologies such as AI is at a crossroads.⁷ Traditional governance is slow; the speed and global diffusion of technological development are continuously increasing. Traditional governance paradigms focus on individual risk, but novel technologies can pose significant societal risks (eg, exacerbating inequality, workforce disruption). Traditional governance strategies exclude many of the people affected, including technology users and communities.⁷ New approaches are needed to complement existing governance strategies.

Deliberative democratic methods, such as citizens' juries, enable community members to influence health policy making.⁸ These robust methods share certain characteristics: participants are selected to reflect population diversity; they are asked to make recommendations regarding a specific question; they are provided high quality relevant information and have extensive opportunities to ask questions; and they work together to reach recommendations that take trade-offs between competing advantages, disadvantages, and values into consideration.⁸

Until 2023, no deliberative process with national representation had considered how AI should be used in health care. We therefore convened a national citizens' jury to discuss the use of artificial intelligence in Australian health care.

Abstract

Objective: To support a diverse sample of Australians to make recommendations about the use of artificial intelligence (AI) technology in health care.

Study design: Citizens' jury, deliberating the question: “Under which circumstances, if any, should artificial intelligence be used in Australian health systems to detect or diagnose disease?”

Setting, participants: Thirty Australian adults recruited by Sortition Foundation using random invitation and stratified selection to reflect population proportions by gender, age, ancestry, highest level of education, and residential location (state/territory; urban, regional, rural). The jury process took 18 days (16 March – 2 April 2023): fifteen days online and three days face-to-face in Sydney, where the jurors, both in small groups and together, were informed about and discussed the question, and developed recommendations with reasons. Jurors received extensive information: a printed handbook, online documents, and recorded presentations by four expert speakers. Jurors asked questions and received answers from the experts during the online period of the process, and during the first day of the face-to-face meeting.

Main outcome measures: Jury recommendations, with reasons.

Results: The jurors recommended an overarching, independently governed charter and framework for health care AI. The other nine recommendation categories concerned balancing benefits and harms; fairness and bias; patients' rights and choices; clinical governance and training; technical governance and standards; data governance and use; open source software; AI evaluation and assessment; and education and communication.

Conclusions: The deliberative process supported a nationally representative sample of citizens to construct recommendations about how AI in health care should be developed, used, and governed. Recommendations derived using such methods could guide clinicians, policy makers, AI researchers and developers, and health service users to develop approaches that ensure trustworthy and responsible use of this technology.

Methods

We convened a national citizens' jury to discuss the question: “Under what circumstances, if any, should artificial intelligence be used in Australian health systems to detect or diagnose disease?” (Supporting Information, part 1). The aim of deliberative democratic methods, developed in political science and government, is to enhance democracy by involving communities in developing the laws or policies that affect them. Deliberative recruitment and sampling methods have a political rather than an epidemiological logic; the aims are to provide all members of a community equal opportunity to participate and to reflect community diversity. These aims are typically achieved by random ballot invitation followed by stratified selection according to demographic criteria to select a mini-public, or diverse small group, that is asked to make decisions on behalf of the broader public (Supporting Information, part 2).

Juror recruitment

The independent, not-for-profit, deliberative democracy recruitment agency, Sortition Foundation (<https://www.sortitionfoundation.org>), recruited thirty Australian residents for this jury. To ensure that each Australian resident had an equal chance of being invited, Sortition Foundation mailed invitations to 6000 households randomly selected from the Australia Post database in February 2023. The invitation described the topic with a brief explanatory background, details about what would be required of participants, information about the nature of community juries, and a detailed participant information statement ([Supporting Information](#), part 3). The number of invitations sent to each state and territory was proportional to its population size. One adult (18 years or older) from each invited household was eligible for participation.

People with direct involvement in AI development or implementation, in clinical occupations, or unable to speak English in a group were excluded from selection. From 109 unique eligible respondents (response rate, 1.8%), Sortition Foundation used an algorithm⁹ for the stratified random selection of 31 participants according to gender, age, ancestry, highest level of education,

location of residence (state/territory; urban, regional, rural). After selection, two jurors opted to not participate; one replacement person was invited, for a total of 30 jurors.

Each juror received \$1015 as compensation for their participation, and we booked and paid for travel, accommodation, and all meals for the face-to-face meeting. Extensive efforts were made to enable participation, including lending computer devices, Zoom training, assisting with logistics, and providing funding for special travel needs.

Jury planning and procedure

The entire jury process took 18 days (16 March – 2 April 2023): fifteen days online and three days face-to-face in Sydney ([Box 1](#)).¹⁰ We shared video and documents via the secure VisionsLive bulletin board platform (<https://visionslive.com/online-bulletin-boards>), and jurors interacted via message boards. Synchronous online discussions were undertaken via Zoom. Facilitation was led by author SMC (an experienced deliberation facilitator); CD (experienced in deliberation) and LC, YSJA, EF, and BF (qualitative researchers with deliberation knowledge) also acted as facilitators.

1 “Under what circumstances, if any, should artificial intelligence be used in Australian health systems to detect or diagnose disease?”: jury schedule

Time	Activities	Core steps
Week 1		
Before first meeting	Participant booklet sent, video conference platform practice (if needed), computer devices sent (if needed)	Understand purpose, information inputs, build skills
Thursday evening: synchronous	Online plenary/small groups: using online platforms, creating ground rules, introductions, learning to ask critical questions	Understand purpose, build relationships, build skills
Friday to Sunday: asynchronous	Two expert videos online, online text-based discussion by jurors, online videos about four case studies	Information input
	Author FM evidence video: What is AI and how does it work in health care? Author KJLB evidence video: How do screening and diagnosis work now? What is evidence-based medicine?	Information input
Sunday afternoon: synchronous	Online plenary/small groups: learning about cognitive biases, discuss values and priorities, develop questions for experts	Build relationships, build skills, information input
Week 2		
Monday to Sunday: asynchronous	Researchers present jurors’ questions to experts and place answers online. Two expert videos online, with online text-based discussion between jurors.	Information input
	Author IS evidence video: The potential and proven benefits of health care AI Author WAR evidence video: The potential risks and harms of health care AI	Information input
Sunday afternoon: synchronous	Online plenary/small groups: sharing new insights, generating questions for experts, identifying remaining knowledge gaps.	Information input
Week 3		
Monday to Thursday asynchronous	Researchers: present jurors’ questions to experts and place answers online; provide information online to address gaps, including contacting additional experts for input	Information input
Friday afternoon/evening (3–7 pm): face-to-face	Opening ceremony with presentation from supporting organisations and review of deliberative process. Speed dialogue with four experts. Activities for relationship/skill building. Revisit ground rules. Welcome reception	Understand purpose, information inputs, build skills, build relationships
Saturday face-to-face (9–5 pm)	Discuss benefits, harms and bias in small groups, review who is important when discussing AI in health care, identify areas for recommendations, begin drafting recommendations in groups	Group dialogue and deliberation, group decision making
Sunday face-to-face (9–3 pm)	Finalise recommendations in small groups and together, identify spokespeople, practice presenting recommendations, closing ceremony; supporting organisations and experts in attendance	Group dialogue and deliberation, group decision making, presentation, closing

The procedure followed six core steps for deliberative processes: understanding purpose, relationship building, skill development, information inputs, group dialogue and deliberation, group decision making, and closing.¹¹ Some activities focused on process, such as structured greeting or reflection activities, learning about cognitive bias, and how to ask critical questions.¹¹ Plenary and small group activities alternated with one another; small groups were frequently randomly re-allocated for cross-fertilisation of perspectives.

Each juror spent at least six hours on jury-related activity across the fifteen-day online period; most contributed much more than six hours. Online activity included watching information videos, asking the experts questions, receiving answers, and interacting with other jurors in three 90-minute online meetings and on the dedicated private bulletin board. Materials generated for and by the process are available online¹⁰ and in the [Supporting Information](#), parts 4 to 8, including a participant booklet sent to participants before the jury process (background information and four diagnostic or screening case studies), and four 10–15-minute online video presentations by four content experts (authors FM, KJLB, IAS, WAR), the drafts of each of which were reviewed by the three other content experts ([Box 1](#)). All jurors watched all four videos. Questions for the experts were developed by jurors online and answered by the experts online. After ten days, jurors identified remaining knowledge gaps, and the research team located appropriate resources for closing them (eg, systematic reviews, websites).

During the three-day face-to-face meeting, jurors met for about 18 hours in total ([Box 1](#)). Observers from several organisations with a professional interest in AI or consumer engagement were present for the three-day face-to-face meeting; a formal protocol and agreement minimised their influence on deliberations. During face-to-face small group discussions, jurors recorded their deliberations in templates. The four speakers directly answered final questions at the end of the first face-to-face day. On the second day, a world café-style session¹² helped jurors discuss and record their insights about the benefits, harms, and bias and fairness of AI in health care.¹³ Jurors then developed a list of questions that might require recommendations, which the research team sorted into draft categories; the entire jury finalised the category list together.

The jury then drafted recommendations in their own words in each of the revised categories, working in self-selected working groups (four to seven people) and drawing on written records of their earlier discussions. All jurors provided input through iterative cycles of plenary feedback, re-drafting, and voting. A recommendation was included in the report if at least 24 jurors supported it. A subgroup of jurors presented the final recommendations to the observers and experts in a closing ceremony.

Analysis

Recommendations and reasons were transcribed and are reported as supplied by the jury; we have added minor edits in square brackets to ease reading. Data recorded in templates during the world café conversations were transcribed into Excel by author LC; SMC and YSJA applied inductive qualitative analysis to independently develop, name, and apportion data to clusters of the jurors' main concerns, resolving differences via consensus. Our report complies with CJCheck guidelines.¹⁴

Ethics approval

This project was approved by the University of Wollongong Health and Medical Human Research Ethics Committee (2022/314).

2 Demographic characteristics of the thirty jurors

Characteristic	Number	Australian population reference
Gender ^{15,*}		
Women	14 (47%)	50.7%
Men	15 (50%)	49.3%
Other	1 (3%)	—
Age group (years) ^{15,†}		
18–24	4 (13%)	10.9%
25–39	8 (27%)	27.4%
40–54	8 (27%)	24.6%
55–74	8 (27%)	27.5%
75 or older	2 (7%)	9.6%
Ancestry ^{15,‡}		
European (British or Irish)/North American	10 (33%)	53.0%
Asian	4 (13%)	17.5%
European (continental)	3 (10%)	19.6%
Aboriginal or Torres Strait Islander	1 (3%)	3.2%
African/Middle Eastern	1 (3%)	4.6%
Latin American	1 (3%)	0.8%
Multiple ancestries/cannot pick one	5 (17%)	32.5%
Other	5 (17%)	—
Highest level of education ¹⁶		
Postgraduate degree	4 (13%)	8.9%
Undergraduate degree	7 (23%)	17.4%
Trade certificate	7 (23%)	28.4%
School certificate or other	12 (40%)	45.3%
State/territory of residence ¹⁷		
New South Wales	10 (33%)	31.4%
Queensland	7 (23%)	20.5%
Victoria	7 (23%)	25.5%
Western Australia	2 (7%)	10.7%
Australian Capital Territory	1 (3%)	1.8%
Northern Territory	1 (3%)	1.0%
South Australia	1 (3%)	7.0%
Tasmania	1 (3%)	2.2%
Remoteness ¹⁸		
Major cities	20 (67%)	72.1%
Other	10 (33%)	27.8%

* Population data is for sex. † Population data proportions are for people aged 18 years or older. ‡ Census has "Australian" as a response option (30% of respondents); we assumed that this category included people with British or Irish ancestry and multiple ancestry. ◆

Results

The demographic characteristics of the jury were similar to those of the Australian population (Box 2). Two jurors participated online but could not attend the face-to-face

meetings because of acute illness; 28 jurors participated in the final deliberations.

Jurors hoped that AI might make health care more efficient, improve systems performance and outcomes and therefore

3 Summary conversations about benefits, harms, and fairness of AI in health care that underpinned recommendation development

Theme cluster	Juror concerns
How important are the potential benefits of using AI for screening and diagnosis? What benefits are most important? Why are those benefits important?	
Cluster 1: Increased access, greater productivity, and reduced costs	Greater productivity through streamlined workflows and automation. Reduced pressure on health services, better allocation of clinician time for delivering higher quality care, increased access to care, including in rural communities, after infrastructure is established, reduced costs, less invasive testing, enabling more testing, easier access to necessary tests.
Cluster 2: Improved clinician performance and care outcomes, increasing confidence in health care	More timely and accurate diagnosis and better prevention and cure by AI-enabled systems. AI could mitigate human bias. Improved clinician performance would improve patient care and caregiver experience, build confidence, and support greater trust in health care and in AI itself.
Cluster 3: Support information sharing, resource allocation, research	Data-rich health services could promote a culture of data sharing, support information sharing and knowledge, identify new causes of disease, and better direct resources and research.
How important are the potential harms or dangers of using AI for screening and diagnosis? What harms are most important? Why are those harms important?	
Cluster 4: Alienation, dehumanisation, and distrust	Reduced human contact and empathy and inability to replicate complex human responses in health care, seeding patient distrust in health care. Population distrust of AI systems, reduced confidence, effect on doctor-patient relationships and flow-on effects of distrust on others. Patients may miss out on beneficial AI-supported health care because of mistrust.
Cluster 5: Governance, commercial, and systems risks	Lack of transparency and review, commercial ownership restricting access to information and reducing public control, unclear lines of responsibility, greater dependence on data accuracy, insurance risks (eg, premium increases), increased costs, more brittle health systems, broader social harms.
Cluster 6: Human costs of poor AI performance	AI errors resulting in psychological and physical harm to patients because of deficiencies in training data, failure to communicate decisions probabilistically, and false screening results (eg, false positive results leading to unnecessary alerts or recalls).
Cluster 7: Job loss, deskilling, and automation bias	Loss of clinical skills, clinician complacency about AI failings and reliance on AI, unrealistic expectations of AI performance, deterioration of health systems because of automation, potential job losses.
Cluster 8: Performance limitations of AI	Concerns about AI mistakes, unknown outcomes, changes over time, inability to synthesise information in the way humans can.
Cluster 9: Algorithmic bias and inequity	Narrow training sets, decreased equity in access and outcomes, over-reliance on incomplete or outdated data.
How can we respond to the potential for bias or unfair outcomes from AI for screening and diagnosis? What principles should guide our responses?	
Cluster 10: Bias in human/AI systems	Sources of bias (eg, developers, coders, data, evidence), humans are also biased, bias is hard to detect and define
Cluster 11: Performance and validation	Need for ongoing testing and validation, including renewing data sources and testing using local data.
Cluster 12: Equity/diversity concerns	Larger, robust, local, and diverse training data, robust research design, diverse developers, and equity of access to AI.
Cluster 13: Transparency regarding limitations of AI/data	Making data and AI shortcomings transparent, ensuring clinicians understand the limitations of AI, making training data transparent.
Cluster 14: Data quality and management	
Managing the quality of data used, maintaining data sources and ensuring data appropriate for the question being asked	
Cluster 15: Principles, solutions and need for guidance	Other possible actions/principles in response to bias. Dominated by the need for strong and proactive governance (prior to implementation) and accountability. Other principles included the need for AI-supported systems to perform at least as well as humans do now, effective advocacy and inclusion of patient perspectives, complete separation from the insurance industry, safeguards against commercial in-confidence algorithmic systems, and ensuring that misuse of private data is prosecuted.

increase trust in health care, and strengthen knowledge and research. Jurors were concerned about dehumanisation of health care, negative effects on clinicians, automation and algorithmic biases, physical and psychological harm to people arising from AI error, and governance risks. Jurors were concerned about difficulties in recognising bias and advocated measures for mitigating bias, including optimal data management, transparency, and strong governance (Box 3).

The jury made fifteen recommendations in ten categories (Box 4). While the evidence provided, the question framing, and jurors' discussions of benefits, harms and bias all focused on diagnosis and screening, many of the final recommendations were more general.

The first recommendation concerned the need for an overarching, independently governed charter and framework. The other nine recommendation categories concerned balancing benefits and harms; fairness and bias; patients' rights and choices; clinical governance and training; technical governance and standards; data governance and use; open source software; AI evaluation and assessment; and education and communication. Jurors endorsed a responsive and sustainable approach to governing health care AI that served the national interest (recommendation 1) and processes for ongoing evaluation (recommendations 3, 8, 10, 13, 14) (Box 4).

The jury understood that health care AI could cause harm, but was not prohibitionist, instead asserting the right of all Australians to access to AI (recommendation 4) and proposed conditions for its legitimate use, including the need to balance harms and benefits (recommendation 3) (Box 4).

Each recommendation achieved support from at least 24 jurors; all but recommendations 4 and 11 achieved unanimous support. Two jurors expressed concern about extending rights beyond Australian citizens and residents (recommendation 4); one juror objected to making heterogeneous datasets mandatory (recommendation 11) because specialised datasets could be appropriate for people from minority groups. This latter disagreement reflected a shared commitment to promoting health equity, but different views on how it should be achieved.

Discussion

We report the first nationally representative deliberative democratic process for developing general recommendations about the use of AI in health care. The recommendations provide decision makers a clear indication of the values and priorities of a well informed and diverse Australian mini-public. Our study illustrates the feasibility of robust public engagement and deliberation for guiding AI development and implementation.

Health care decision makers and clinicians should pay attention to the core features of the recommendations and reasons advanced for them, particularly the two most frequent concerns: evaluation, integrity and transparency; and fairness. Jurors called for mandatory reporting of unfavourable outcomes, performance, misuse, and benefits, robust data and evidence, and ongoing evaluation to guarantee safety, effectiveness, appropriate scope of application, and training data selection, and to ensure that benefits outweigh harms and health system performance is preserved (recommendations 3, 6–10, 12–14). Jurors emphasised that all people, including people from minority backgrounds, should benefit from AI, that exacerbation of inequity be avoided, diverse values be respected, and training data be representative (recommendations 1–5, 11, 13, 15).

Five further principles informed several recommendations: making decision makers accountable for the performance of AI health care systems (recommendations 7–9, 15); supporting community understanding of and involvement in the governance of health care AI (2, 12, 15); preserving choice, rights and autonomy in health care systems (3–5); managing conflicts of interest and ensuring independence in health care AI governance and implementation (2, 12); and support and training for clinicians in the use of AI (3, 6).

The few previous studies similar to ours were all undertaken in the United Kingdom. In 2019, two five-day, 18-person citizens' juries in Manchester and Coventry discussed the question, "Should AI systems give an explanation, even if that results in less accurate decisions?"; jurors expressed a preference for accuracy only in health scenarios.¹⁹ In 2018, a four-day, 29-person citizens' jury from England and Wales deliberated the question, "Under what conditions, if any, is it appropriate to use automated decision systems?";²⁰ in 2020, a 50-person Citizens' Biometrics Council from Bristol and Manchester discussed (for 60 hours over nine months) "What is or isn't OK when it comes to the use of biometric technologies?"²¹ Jurors in the latter two discussions emphasised the need to avert bias, and called for robust frameworks for responsibility, oversight, and accountability, independent evaluation, monitoring and auditing, and consent (eg, the option of declining the use of biometric technologies).^{20,21} Although these processes were not focused on health, their recommendations resonate with those of our jury.

The most fundamental recommendation in our study was the call for a health AI charter and an independent decision-making body. This is more ambitious than a framework or code of conduct, and would provide AI-specific oversight in health. There are other examples of AI-specific legislation or regulation, most notably the European Union AI Act.²² Implementing this recommendation would require identifying potential system barriers and developing an operational plan and supportive policy. Some elements recommended by the jury (eg, evaluation of training data) are currently undertaken within the "software as a medical device" approach to AI regulation of the Therapeutic Goods Administration.²³ However other elements, such as examining the effects of AI systems on patient outcomes, clinicians, and health systems, should be incorporated into health care quality and safety processes and governance processes.²

Our jury proposed responsibilities for people across the health care system, including:

- individual clinicians: understanding and evaluating AI as used in health care, including its shortcomings, and ensuring that training data are relevant to local people;
- clinical training and accreditation bodies: ensuring that clinicians are knowledgeable about the use and limits of AI systems;
- patients' representatives: advocating patients' rights, the provision of quality information to patients, and standards for AI use, as well as holding decision makers to account;
- health care organisations and service providers: auditing AI systems for integrity, performance, and bias in local populations before procurement, managing conflicts of interest, considering the use of open source software, ensuring the ongoing monitoring of overall health system performance;

4 Final recommendations of the jury*

Category/Recommendations	Reasons
Overarching charter and framework	
<p>1. We must have a charter for AI in the Australian health system and services. The charter must include (but not be limited to) the following:</p> <ul style="list-style-type: none"> ▶ Underrepresented people, including Aboriginal and Torres Strait Islander people and [people from] minority populations ▶ Rural and remote [populations] ▶ Sustainability and environment ▶ Australian security and sovereignty ▶ Ethics and human rights 	<ul style="list-style-type: none"> • It is important that the charter responds to dynamic environments to ensure an ethical and moral pathway to the development and deployment of AI in health diagnosis and detection.
<p>2. There must be an independent decision-making body to manage the charter. We recommend representation from across all stakeholder groups. We recommend the board chair is independent of the health system and investors to avoid bias.</p>	<ul style="list-style-type: none"> • Reasons for independent board: <ul style="list-style-type: none"> ▶ Adequate representation of all concerned. ▶ Create fairness and equity across the board. ▶ Guide future direction. ▶ Community acceptance. • Reasons for independent chair: <ul style="list-style-type: none"> ▶ To avoid specification bias ▶ Instil neutrality ▶ Balancing vested interests of board members ▶ Free from conflict of interest
Balancing benefits and harms	
<p>3. Our recommendation in the application of AI in health care is that it must be continually evaluated to ensure the benefits to patients and health care professionals outweigh the harms</p>	<ul style="list-style-type: none"> • To address potential harms and benefits to underrepresented groups. • To ensure people are able to make an informed choice with regard to their health care in partnership with healthcare professionals. • The charter must include potential benefits and harms to all patients, health care professionals and underrepresented groups.
Fairness and bias	
<p>4. Our recommendation is that access to AI in health care must be the universal right of all Australians</p>	<ul style="list-style-type: none"> • To ensure equitable access free from limitations and bias with specificity to addressing underrepresented groups.
Patient rights and choice	
<p>5. There must be a guideline for patient rights. It is important to have guidelines that are inclusive of and non-discriminatory [with respect] to: individual values/beliefs, choice, accessibility, respecting underrepresented peoples, and being culturally appropriate.</p>	<ul style="list-style-type: none"> • Every patient has their own morals, values [and] beliefs that must always be considered and respected when making health care choices (that are outside, inside, or around AI). • Allowing patients to have their own choice in denying, accepting, or knowing alternatives to the use of AI and health care in general. • It is important that patients understand the procedures being performed in simple language. • Accessibility is a right that must be respected for people who are underrepresented, people from all cultures and backgrounds, people with disability as well as people living in rural areas. • AI use [should] have an overall fairness and inclusivity, in terms of cost, quality of service, and respecting the wide range of differences in reference [to] (but not only inclusive of) age, gender, ethnic background, and identity preference
Clinical governance and training	
<p>6. We recommend that health care workers must be trained in AI systems that are to be implemented to their practice environment before clinical use.</p>	<ul style="list-style-type: none"> • Safe and effective delivery of health care in line with evidence-based practice.
<p>7. We recommend that professional bodies must have clear directions regarding the use and intended outcomes of AI in the domains for which they are responsible.</p>	<ul style="list-style-type: none"> • Accountability for potential misuse of AI. • [Background notes for this recommendation emphasised the need for clinicians to receive high-quality training about the shortcomings of AI systems, as some AI models have significant limitations that clinicians need to understand; eg, models that can identify large but not small vessel occlusion stroke]
<p>8. We recommend that monitoring, auditing, and reporting be made mandatory to the appropriate governing body [or] bodies. Such processes should include but are not limited to unfavourable outcomes, performance, misuse and any benefits to the patients, clinicians, and health care systems.</p>	<ul style="list-style-type: none"> • To ensure patient safety and accountability.

Continues

4 Continued

Category/Recommendations	Reasons
Technical governance and standards	
9. Upon submission to the regulator, an AI system must provide information on its intended purpose and efficacy, its training dataset, flaws and limitations of use.	<ul style="list-style-type: none"> To make clear to all involved on what AI does and doesn't do.
10. For AI systems to be approved in Australia, they must perform equal to or better than current standard health care practice.	<ul style="list-style-type: none"> To ensure accuracy and specificity of the detection performance of an AI system. Any approved system needs to meet a high standard and threshold. This provides measurable standards which can be applied across all future and proposed AI systems entering clinical settings. This standard must be maintained and re-evaluated at regular intervals once in use.
Data governance and use	
11. It is important that AI training datasets must strive to be adequately representative and inclusive to capture Australia's multiculturalism and diversity.	<ul style="list-style-type: none"> Australia has a wide variety of cultural, gender, and ethnic groups. The representation of these groups should be captured in these datasets to train and set up parameters of an AI.
Open source software	
12. Encourage and consider having AI software in health be free and open source software to ensure transparency, public ownership, financial integrity, collaboration, security, privacy and trust.	<ul style="list-style-type: none"> Transparency and quality control: The technology should be transparent in its inner workings, flaws and limitations, changes over time now and in the future. Public ownership/intellectual property: The technology should be owned by the public, not private companies. We should avoid creating and supporting monopolies. Financial integrity: The technology should avoid relying too much on companies to maintain financial integrity, to avoid being dictated by financial motivations. Collaboration: The technology should benefit from the improvements, reduced cost and reduction of bias that collaboration can provide. Security, privacy and trust: free and open source software is known to be highly secure when implemented properly and helps privacy/trust.
Evaluation and assessment	
13. We recommend that research used to underpin the use of AI in health care must be peer-assessed in an unbiased, independent, and robust manner. Australian data, with a sample representative of the population, should be used, but overseas data can be used when justified.	<ul style="list-style-type: none"> To confirm [or] verify developers' claims of AI system performance. To maintain a standard of quality for healthcare in Australia
14. Research assessing the performance of AI screening tools should reflect real world clinical practice and follow standardised procedures in trial design. Data analysis and reporting should be transparent, and conclusions should reflect system performance.	<ul style="list-style-type: none"> The evaluation process should be transparent to ensure validity
Education and communication	
15. We recommend that there is a comprehensive and fully funded community education program. This will ensure that the community is brought along with developments in and the application of AI in health. This should be located within a broader program of general digital health literacy that recognises particular community needs such as age, gender, ethnicity etc.	<ul style="list-style-type: none"> To ensure the community is informed and educated on current AI developments. Also, children are exposed to AI and digital health through school-based learning programs. This will aid understanding of future development of AI in health, ensuring greater participation. Community education and awareness raising will ensure community can hold authorities such as regulators to account, as they will be knowledgeable about raising these concerns and reporting instances of non-compliance. A community that is educated about AI in health might have less fear and be able represent their individuals more effectively.

* Recommendations and reasons were transcribed and are reported as supplied by the jury; we have added minor edits in square brackets to ease reading. ◆

- researchers and evaluators: auditing datasets for representativeness, rigorously and independently evaluating AI system performance in clinical care, and embedding ongoing monitoring and feedback; and
- health departments and agencies: building public understanding of health care AI and incorporating public voices into decision making about AI in health care.

The jurors emphasised collective concerns related to system integrity, fairness, accountability, and community involvement, reinforcing the need for governance that considers societal aspects beyond risks to individuals.⁷ They also emphasised rigorous evaluation and fairness, aspects that may be neglected by

commercial producers of health care AI. Reported breakthroughs in health care machine learning have often not been supported by more methodologically rigorous scrutiny,¹⁹ and evaluations of health care AI have often focused on overall accuracy rather than bias or fairness.¹³ The jury's recommendations suggest that a well informed public might reject these approaches as unjustifiable.

Limitations

Best practice methods were applied to recruiting and selecting jurors (invitation by random ballot, stratified sampling according to selected population demographic characteristics). However, as deliberative democratic processes require substantial interest and commitment from participants, selection bias was

inevitable; people who agreed to participate may have been more civic-minded and interested in the discussion topic than Australians in general. However, all jurors actively participated, and the diversity of views expressed reflected the diversity of the jury. Our jury size was adequate for effective deliberation; 20 to 50 people is typical for this type of study.²⁴ While larger juries can seem more representative, they require more resources, individual jurors may be less active because they assume others will represent their views, blocs can form, and effectively including everyone in deliberations becomes more difficult²⁰ (Supporting Information, part 2).

The focus of the study question was screening and diagnosis, but the jurors expressed final recommendations regarding AI in health care generally, although the evidence they were provided was more limited. Jurors considered four case studies about how AI might be used in medical practice; their judgements may have differed had they been presented different cases. The jurors' recommendations are reported verbatim, and reflect the limited time available to prepare their wording.

Conclusion

A nationally representative citizens' jury can express informed community views about how AI in health care should be

developed, used, and governed. Few deliberative democratic processes have considered such questions, but these methods could guide clinicians, policy makers, AI researchers and developers, and health service users to develop approaches that can support the trustworthiness of this technology.

Acknowledgements: This study was supported by the National Health and Medical Research Council (1181960).

Open access: Open access publishing facilitated by University of Wollongong, as part of the Wiley – University of Wollongong agreement via the Council of Australian University Librarians.

Competing interests: No relevant disclosures.

Data sharing: Individual deidentified participant data will be partially shared. The ethics approval for the study stipulated that transcripts of recordings of the jurors' deliberations would remain confidential because of the risk of individual identification. Our study did not involve data dictionaries. Extensive information about the study protocol, and data generated for and in the study (including descriptions of the process, the expert witness videos, questions generated by the jury, and answers provided by the experts) are available at https://uow.info/TAWSYN_JURY. ■

Received 30 June 2023, accepted 6 November 2023

© 2024 The Authors. *Medical Journal of Australia* published by John Wiley & Sons Australia, Ltd on behalf of AMPCo Pty Ltd.

This is an open access article under the terms of the [Creative Commons Attribution License](#), which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

- 1 Australian Department of Industry, Science and Resources. Supporting responsible AI: discussion paper. Government's interim response. 17 Jan 2024. <https://consult.industry.gov.au/supporting-responsible-ai> (viewed Mar 2024).
- 2 Lyell D, Wang Y, Coiera E, Magrabi F. More than algorithms: an analysis of safety events involving ML-enabled medical devices reported to the FDA. *J Am Med Inform Assoc* 2023; 30: 1227-1226.
- 3 Wadie J. A roadmap for artificial intelligence in healthcare for Australia news]. *Australian Alliance for Artificial Intelligence in Healthcare*, 1 Dec 2021. <https://aihealthalliance.org/2021/12/01/a-roadmap-for-ai-in-healthcare-for-australia> (viewed Oct 2023).
- 4 Coiera EW, Verspoor K, Hansen DP. We need to chat about artificial intelligence. *Med J Aust* 2023; 219: 98-100. <https://www.mja.com.au/journal/2023/219/3/we-need-chat-about-artificial-intelligence>
- 5 Pearce C, McLeod A, Rinehart N, et al. Artificial intelligence and the clinical world: a view from the front line. *Med J Aust* 2019; 210 (6 Suppl): S38-S40. <https://www.mja.com.au/journal/2019/210/6/artificial-intelligence-and-clinical-world-view-front-line>
- 6 Moodie C. Australian Medical Association calls for national regulations around AI in health care. *ABC News (Australia)*, 28 May 2023. <https://www.abc.net.au/news/2023-05-28/ama-calls-for-national-regulations-for-ai-in-health/102381314> (viewed Oct 2023).
- 7 Mathews DJH, Balatbat CA, Dzau VJ. Governance of emerging technologies in health and medicine: creating a new framework. *N Engl J Med* 2022; 386: 2239-2242.
- 8 Degeling C, Carter SM, Rychetnik L. Which public and why deliberate? A scoping review of public deliberation in public health and health policy research. *Soc Sci Med* 2015; 131: 114-121.
- 9 Flanigan B, Gözl P, Gupta A, et al. Fair algorithms for selecting citizens' assemblies. *Nature* 2021; 596: 548-552.
- 10 Australian Centre for Health Engagement Evidence and Values. Artificial intelligence in health: community jury. Undated. https://uow.info/TAWSYN_JURY (viewed Oct 2023).
- 11 White K, Hunter N, Greaves K. Facilitating deliberation: a practical guide. Melbourne: MosaicLab, 2022.
- 12 Brown J, Isaacs D. The world café: shaping our futures through conversations that matter. San Francisco: Berrett-Koehler, 2005.
- 13 Aquino YSJ, Carter SM, Houssami N, et al. Practical, epistemic and normative implications of algorithmic bias in healthcare artificial intelligence: a qualitative study of multidisciplinary expert perspectives. *J Med Ethics* 2023; jme-2022-108850.
- 14 Thomas R, Sims R, Degeling C, et al. CJCheck Stage 1: development and testing of a checklist for reporting community juries: Delphi process and analysis of studies published in 1996-2015. *Health Expect* 2016; 20: 626-637.
- 15 Australian Bureau of Statistics. Snapshot of Australia, 2021. 28 June 2022. <https://www.abs.gov.au/statistics/people/people-and-communities/snapshot-australia/2021> (viewed Oct 2023).
- 16 Australian Bureau of Statistics. Education and training: census, 2021. 28 June 2022. <https://www.abs.gov.au/statistics/people/education/education-and-training-census/latest-release> (viewed Oct 2023).
- 17 Australian Bureau of Statistics. National state and territory population, December 2023. 28 June 2022. <https://www.abs.gov.au/statistics/people/population/national-state-and-territory-population/dec-2021> (viewed Oct 2023).
- 18 Australian Institute of Health and Welfare. Profile of Australia's population. 6 June 2023. <https://www.aihw.gov.au/reports/australias-health/profile-of-australias-population> (viewed Oct 2023).
- 19 van der Veer SN, Riste L, Cheraghi-Sohi S, et al. Trading off accuracy and explainability in AI decision-making: findings from 2 citizens' juries. *J Am Med Inform Assoc* 2021; 28: 2128-2138.
- 20 Forum for Ethical AI. Democratising decisions about technology: a toolkit. 24 Oct 2019. <https://www.thersa.org/reports/democratising-decisions-technology-toolkit> (viewed Oct 2023).
- 21 Ada Lovelace Institute. The Citizens' Biometrics Council London: report with recommendations and findings of a public deliberation on biometrics technology, policy and governance. 30 Mar 2021. <https://www.adalovelaceinstitute.org/report/citizens-biometrics-council> (viewed Oct 2023).
- 22 European Commission. Proposal for a regulation of the European parliament and the council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain union legislative acts (COM/2021/206 final). 21 Apr 2021. <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52021PC0206> (viewed Oct 2023).
- 23 Therapeutic Goods Administration. Regulation of software based medical devices. 28 Sept 2023. <https://www.tga.gov.au/how-we-regulate/manufacturing/medical-devices/manufacture-guidance-specific-types-medical-devices/regulation-software-based-medical-devices> (viewed Oct 2023).
- 24 Street J, Duszynski K, Krawczyk S, Braunack-Mayer A. The use of citizens' juries in health policy decision-making: a systematic review. *Soc Sci Med* 2014; 109: 1-9. ■

Supporting Information

Additional Supporting Information is included with the online version of this article.