Australian College of Rural & Remote Medicine
WORLD LEADERS IN RURAL PRACTICE

College Submission
September 2022

Feedback on the RANZCR Draft Statement: Ethical Principles of Artificial Intelligence (AI) in Medicine

About the Australian College of Rural and Remote Medicine (ACRRM)

ACRRM's vision is ***the right doctors, in the right places, with the right skills, providing rural and remote people with excellent health care.*** It provides a quality Fellowship program including training, professional development, and clinical practice standards; and support and advocacy services for rural doctors and the communities they serve.

ACRRM is accredited by the Australian Medical Council to set standards for the specialty of general practice. The College's programs are specifically designed to provide Fellows with the extended skills required to deliver the highest quality Rural Generalist model of care in rural and remote communities, which often experience a shortage of face-to-face specialist and allied health services.

ACRRM has some 5000 rural doctor members including some 1000 doctors in training, who live and work in rural, remote, and indigenous communities across Australia. Our members provide expert front line medical care in a diverse range of settings including general practices, hospitals, emergency departments, Aboriginal Medical Services, and other remote settings such as RFDS and Australian Antarctic Division.

Initial Comments

The College welcomes the opportunity to provide input into the Royal Australian and New Zealand College of Radiologists review of the statement on Ethical Principles of Artificial Intelligence (AI) in Medicine.

The College agrees that as technology rapidly evolves, the importance of checks and balances to ensure patient safety becomes even more paramount. AI holds great promise for improving diagnosis, treatment, health research, and delivery of healthcare and medicine worldwide, but only if ethics and human rights are put at the heart of its design, deployment, and use.[1] Studies from developing

[1] https://www.who.int/publications/i/item/9789240029200

countries demonstrate that the application of AI medical techniques can improve healthcare outcomes in rural areas[2]. When considering the benefits for rural and remote patients and practitioners it is important that AI ethical frameworks are cognisant of rural and remote context and circumstances.

We have provided feedback in relation to the principles which we considered would benefit from amendment and provided recommendations for possible inclusions.

Response to Nine Ethical Principles for Artificial Intelligence

Principle One: Safety (lines 22-26)

Although AI tools have enormous potential, a range of new risks will emerge from AI tools or through their implementation.

The first and foremost consideration in the development, deployment or utilisation of AI tools must be patient safety and quality of care, with the evidence base to support this

Lines 22-26: the College agrees that patient safety and quality of care are paramount, however, AI tools must also be accurate and efficient. Quality of care should include quality control and quality improvement measures, which should be available to users.

Principle Two: Privacy and Protection of Data (lines 27-34)

Healthcare data is amongst the most sensitive data which can be held about an individual. Patient data must not be transferred from the clinical environment at which care is provided without the patient's consent, approval from an ethics board or where otherwise required or permitted by law. Where data is transferred or otherwise used for AI research, it must be de-identified such that the patient's identity cannot be reconstructed.

Every effort must be made to store a patient's data securely and in line with relevant laws and 33 best practices.

Line 31: the College would suggest adding a definition of "de-identified data". True de-identification can be problematic in small rural and remote communities.

Principle Three: Avoidance of Bias (lines 37-46)

AI tools are limited by their algorithmic design and the data they have access to making them prone to bias. As a general rule, AI tools trained on greater volumes and varieties of data should be less biased. Moreover, bias in algorithmic design should be minimised by giving conscious consideration to avoiding bias and involving a range of perspectives and skill sets in the design process.

The data on which AI tools are based should be representative of the target patient population on which the system or tool is being used. The characteristics of the training data set and the environment in which it was tested must be clearly stated when marketing an AI tool to provide transparency and facilitate implementation in appropriate clinical settings. Particular care must be taken when applying an AI tool to a population, demographic or ethnic group for which it has not been proven effective.

[2] https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6110188/

To minimise risk of bias, the process, training data set and outcome measures used during development must be transparently stated.

Line 42 – we would suggest adding a definition for 'target patient population'. How granular does that need to be? e.g., all Australians includes all Indigenous Australians, but the representation would be small, potentially impacting accuracy.

Rural and remote practitioners and patients must be involved in the design, application and evaluation of AI tools, and systems must consider the range of clinical settings in which the AI will be utilised/applied.

Line 46: we would recommend that where a demographic is underrepresented, efforts should be made to identify appropriate research and datasets to enable populations.

Principle Four: Transparency and Explainability (lines 49-56)

AI tools can produce results which are difficult to interpret or replicate. When used in medicine, the medical practitioner must be capable of interpreting the basis on which a result was reached, weighing up the potential for bias and exercising clinical judgement regarding findings.

AI tools should ideally employ explainable AI (XAI) techniques to justify the underlying basis for decision-making in a way that is understandable to humans.

When designing or implementing an AI tool, consideration must be given to how a result that can impact patient care be best understood and explained by a medical practitioner.

Line 53: where the products of an AI method are not 'explainable' efforts should be made to validate the derived algorithm/s and proceed to clearly label the unexplainable / black box AI and how it will be utilised.

More generally, sufficient information should be published or documented prior to design and/or deployment to assist medical practitioners in providing a suitable explanation of results which a patient can easily understand.

Principle Five: Application of Human Values (lines 57-67)

The development of AI tools for medicine should ultimately benefit the patient and society. ML and AI are programmed to operate in line with a specific world view, however the use of AI tools should function without unfair discrimination and not exacerbate existing disparities in health outcomes. Any shortcomings or risks in AI tools should be considered and weighed against the benefits of enhanced decision making for specific patient groups.

The medical practitioner must apply humanitarian values (from their training and the ethical framework in which they operate) to any circumstances in which AI tools are used in medicine but must also consider the personal values and preferences of their patient in this situation. Entities developing AI tools must demonstrate an understanding of ethical principles and human values.

Suggested edits to Lines 58-60:

"The development of AI tools for medicine should ultimately benefit the patient and society. ML and AI are programmed to operate in line with a specific world view, however the use of AI tools should function without ~~unfair~~ discrimination and not exacerbate existing disparities in health outcomes."

Delete reference to "unfair" where it appears in Line 60.

Principle Six: Decision-making on diagnosis and treatment (lines 68-75)

Fundamental to quality healthcare is the relationship between the medical practitioner and the patient. The medical practitioner is the trusted advisor on complex medical conditions, test results, procedures and treatments who then communicates findings to the patient clearly and sensitively, answers questions and agrees on the next treatment steps.

While AI tools can enhance decision-making capability, final decisions about care are made after a discussion between the medical practitioner and the patient taking into account the patient's presentation, history, options, and preferences.

Lines 73-75 - consider adding reference to the patient or his/her/their carers when the patient is not able to make decisions. This could alternatively be added by way of a definition on patient in Appendix 1 Definitions.

By way of general feedback on this principle, the World health Organisation (WHO) guidance on *Ethics & Governance of Artificial Intelligence for Health*[3] Reference states human autonomy as the first principle of AI in healthcare. Humans should remain in control of health care systems and medical decisions and patients must give valid informed consent through appropriate legal frameworks for data protection. Amending this principle to state the principal of patient autonomy in decision making would bring the Draft Statement in line with WHO recommendations.

Principle Eight: Responsibility for Decisions Made (lines 76-94)

Responsibility for decisions made about patient care rests principally with the medical practitioner. Medical practitioners need to be aware of the limitations of AI tools and must exercise solid clinical judgement at all times. However, given the multiple potential applications of AI tools in the patient journey, there may be instances where responsibility is shared between:

- ***The medical practitioner caring for the patient;***
- ***The hospital or practice management who took the decision to deploy the systems or tools; and***
- ***The manufacturer that developed the ML system or AI tool.***

Although the prime responsibility regarding patient care remains with the medical practitioner, when using AI tools, the responsibility is also shared by the managers of the healthcare environment and the manufacturers and developers of AI tools. This potential for shared responsibility when using AI tools must be identified, recognised by the relevant party, and recorded upfront when researching or implementing AI tools.

Line 82: in a shared care arrangement, the term "medical practitioner" must be defined so it is clear which provider is responsible. We would suggest using terminology already widely used elsewhere,

[3] *Ibid.*

such as “the patient’s nominated healthcare provider” in My Health Record guidelines, which generally refers to the patient’s GP.[4]

Principle Nine: Governance (lines 95-102)

ML and AI are fast moving areas with the potential to add great value but also to do harm. The 96 implementation of AI tools requires consideration of a broad range of factors including how the ML or AI will be adopted across a hospital or practice and which patient groups will be affected and how it might align with patients’ goals of care.

A hospital or practice using or developing AI tools for patient care applications must have accountable governance to oversee implementation and monitoring of performance and use, to ensure the practice is compliant with ethical principles and standards.

AI technologies are generally designed to perform specific tasks. Hospitals or practices using or developing must ensure that AI is used (i) under appropriate conditions and (ii) by appropriately trained people.

Governance systems must ensure effective mechanisms for recourse and/or redress for users and/or patients who may be adversely impacted by implemented AI technologies, their results, or their outcomes.

Broader Ethical Frameworks (lines 103-115)

Other ethical frameworks cover the expected approach and behaviour of medical practitioners when delivering care to patients and provide general guidance relating to the development and adoption of 105 new technologies in medicine.

Medical practitioners in Australia are expected to practise in accordance with the Medical Board of Australia’s Good Medical Practice: A Code of Conduct for Doctors in Australia and the Australian Medical Council’s Good Medical Practice.

Medical practitioners in New Zealand are expected to practise in accordance with the New Zealand Medical Council’s Good Medical Practice and the Code of Ethics set by the New Zealand Medical Association. Medical Practitioners in New Zealand must also comply with the Code of Health and Disability Services Consumers’ Rights.

RANZCR has also developed a more explicit Code of Ethics for clinical radiologists and radiation oncologists.

Australia’s Artificial Intelligence Ethics Framework[5] guides responsibly designed, developed, and implemented AI, and a guiding principle of the framework is that AI systems should be used to benefit all human beings, including future generations. Ensuring that AI systems respect human rights, diversity and the autonomy of patients are paramount in the development of strong ethical frameworks.

Appendix one – Definitions (lines 118 to 176)

[4] https://www.myhealthrecord.gov.au/for-healthcare-professionals/howtos/shared-health-summaries
[5] https://www.industry.gov.au/data-and-publications/australias-artificial-intelligence-ethics-framework/australias-ai-ethics-principles

Line127 - Definition of Explainable Artificial Intelligence (XAI): it would be useful to include a statement on unexplainable AI, which could be used in the context of education and training processes for other doctors.

Is anything important missing?

The WHO guidance (*Ethics & Governance of Artificial Intelligence for Health*) referenced earlier in this submission contained six guiding principles for the design and use of AI in health in 2021. These recommendations ensure the governance of AI for health and ensures the accountability and responsiveness of technologies to healthcare workers, communities, and individuals.

RANZCR may wish to consider amending the Draft Statement to include reference to the following aspects of the WHO guiding principles:

WHO Principle 5 - aims to ensure inclusiveness and equity, recommending that AI for heath be designed to encourage the widest possible equitable use and access irrespective of age, sex, gender, income, race, ethnicity, sexual orientation, ability of other characteristics protected under human rights codes. Consider including a similar guiding principal in the RANZCR statement.

WHO Principle 6 – promotes AI which is responsive and sustainable. Designers, developers, and users should continuously assess AI applications during use to determine whether AI response adequately and appropriately to expectations and requirements. Systems should be designed to minimise environmental consequences and increase energy efficiency. Consideration should be given to amending the Draft Statement to incorporate reference to minimising environmental impact and maximising efficiency.

College Details

Organisation	Australian College of Rural and Remote Medicine (ACRRM)
Name	Marita Cowie AM
Position	Chief Executive Officer
Location	Level 1, 324 Queen St, PO Box 2507 Brisbane Qld 4001
Email	m.cowie@acrrm.org.au
Phone	07 3105 8200

ACRRM acknowledges Australian Aboriginal People and Torres Strait Islander People as the first inhabitants of the nation. We respect the Traditional Owners of lands across Australia in which our members and staff work and live and pay respect to their Elders past present and future.