



Ethical principles for the application of artificial intelligence (AI) in nuclear medicine

Geoff Currie¹ · K Elizabeth Hawk² · Eric M. Rohren³

Published online: 11 January 2020

© Springer-Verlag GmbH Germany, part of Springer Nature 2020

Keywords Nuclear Medicine · Machine learning · Deep learning · Artificial intelligence · Synthetic intelligence · Medical ethics · Intelligent imaging

Introduction

The emergence of artificial intelligence (AI) in Nuclear Medicine, and the promise of synthetic intelligence, heralds an era of disruptive technology with the potential to re-invigorate the ecosystem of Nuclear Medicine and reengineer the landscape in which Nuclear Medicine is practiced. While AI is not new in Nuclear Medicine, more recent developments and applications of machine learning and deep learning create refreshed interest in the ethical issues associated with AI implementation in Nuclear Medicine. Insight into the architecture, operation and implementation of AI in Nuclear Medicine is beyond the scope of this discussion and has been reported elsewhere (1–3). Nonetheless, it is important to provide key definitions.

AI is a general term used to describe algorithms designed for recognition, problem solving and reasoning generally associated with human intelligence, essentially imitating some aspects of intelligent human behaviour (1, 3). An artificial neural network (ANN) is a subset of AI, and in medical imaging, an ANN is an image analysis algorithm composed of layers of connected nodes that simulate the neuronal connections of the human brain (2, 3). ANNs are designed to analyse data and recognise trends or patterns that inform predictions (e.g. classification of

disease). A convolutional neural network (CNN) is a type of ANN used for deep learning that employs a convolutional process to extract features from the image itself, while an ANN typically has feature data as the input (1, 2).

Machine learning (ML) is a subtype of AI that employs ML algorithms through data analysis without being explicitly programmed. ML tends to be associated with solving problems of logic after learning from human-defined teaching cases typical of an ANN (1, 2). Deep learning (DL) is then a sub-type of ML that adds a number of processing layers (depth) to detect complex features in an image typical of a CNN (1, 2). Synthetic intelligence (SI) provides authentic higher-order reasoning using technology like quantum logic (2), while AI simply imitates human thought. In healthcare delivery in general, AI has two different types of presence in the patient care experience: virtual and physical. Virtual applications are often thought of as software-type algorithms that integrate into the patient care episode often for the purposes of decision-making. Conversely, physical presence is often in the form of a material tangible solution such as a robot or present machine which can interact directly with the patient (4). In Nuclear Medicine, AI solutions are most commonly virtual. Nonetheless, as the field continues to evolve, it will be critical to consider ethical challenges associated with physically present AI solutions.

AI in Nuclear Medicine and Molecular Imaging ushers in an exciting era with reengineered and reimagined clinical and research capabilities. AI has the potential to improve workflow and productivity, reduce costs, increase accuracy and empower research and discovery. With this comes a duty of care to patients to ensure AI-augmented diagnosis or treatment provides the best outcomes for patients. Perhaps the most contentious topic to navigate in AI application to Nuclear Medicine is the ethical questions that arise when using human data to develop human-targeted applications. These ethical considerations relate to three distinct areas: the

Footnote

AI ethics in nuclear medicine

This article is part of the Topical Collection on Editorial

✉ Geoff Currie
gcurrie@csu.edu.au

¹ School of Dentistry and Health Sciences, Charles Sturt University, Wagga Wagga, Australia

² Stanford University, California, USA

³ Baylor College of Medicine, Houston, USA

data used, the algorithms applied and the manner in which they are applied to practice. A white paper from the French radiology community (5) and a joint statement of European and North America societies (6) also identify these three areas in the dynamic between ethical and social issues for AI in medical imaging (Fig. 1).

Ethical challenges

In imaging, the learning around ethics in AI is occurring in parallel to innovation and implementation (6). There is a duty of care to deeply understand the technology, its benefits and its risks. Many AI algorithms operate within a “black box” environment, where the underlying steps in the analysis are not transparent. The inability to achieve deep understanding in the face of rapidly evolving technology is a significant ethical challenge. AI confronts the basic ethical challenges associated with autonomy, beneficence, justice and respect for knowledge (7). One key concept that arises with initial handling of data is the difference between data privacy and confidentiality. Privacy refers to control over personal information, while confidentiality refers to the responsibility of maintain privacy (8).

There is increased demand for not only richer data (well-labelled) but commercial access to it driven by the potential to significantly improve health and well-being (6). There is a trade-off then between this beneficence and potential maleficence through commercial exploitation of data or doing actual harm to patients or the “common good” (6). The foundation of capability for ANNs is large data sets for training and validation. This requires access to “big data”. Commercialisation of ML and DL algorithms then requires export of data to third-

party vendors. Whether training, validation, research or clinical data, issues of privacy and confidentiality linked to data usage include whether the patient is aware their data is being used, to what extent and which aspects of their data are being used (9). Furthermore, patients should know who has access to their data and whether (and to what degree) their data has been de-identified (9). From an ethical perspective, a patient should be aware of potential for their data to be used for financial benefit to others and whether potential changes in legislation increase data vulnerability in the future, especially if there is any risk that the data could be used in a way that is harmful to the patient (9). There remains debate about who owns patient data and what is and is not allowed permitted to be done with that data. Five key aspects of ethical data handling in AI include informed consent, privacy and data protection, ownership, objectivity and inequity related to those that either have or lack the resources to process the data (10).

An important ethical and social perspective relates to the manner in which AI shapes human behaviour. The risk of weakening the social and moral responsibilities of individuals relying on AI toward their fellow human has been identified (5). Furthermore, this is accompanied by a risk that decision-making around diagnosis or management and, thus, responsibility could be diverted to autonomous AI systems (5). This is perhaps exacerbated by the more recent emergence of SI and associated logic and autonomous decision-making capability. Clear guidelines need to be developed when decisions are made based on the output of AI in terms of ethical and legal responsibility; is it the designer or the user that is responsible, where is the balance of responsibility if the answer is “both” and how do we reconcile that responsibility when algorithms are developed unsupervised? Ultimately, the physician takes responsibility for decisions, how information is used or weighted, and has stewardship of the knowledge economy broadly and for an individual patient. AI in and of itself has no responsibility which raises an interesting ethical dilemma.

As often depicted in movies, humans expect ethical/moral interactions with and from a non-human intelligent life forms. When dealing with super-logic and human or superior-to-human intelligence (e.g. humanoid, android, etc.), an expectation exists to apply the same social, ethical and moral norms as would be expected amongst humans. By extension, humanoid or android beings should be held to the same standards. In the case of responsibility and liability, AI is not currently accountable. If AI and SI learn unsupervised, the blurred borders between responsibilities will be challenged increasingly in the future. The Frankenstein paradox might be considered here; human regulation of science capable of superiority over humans with blurred boundaries between what is human, human-like and non-human. It is apt that the term Frankenstein is generally used, neither defining Dr. Frankenstein nor Frankenstein’s monster separately. Is the real question a question of consciousness? If the Turing test is

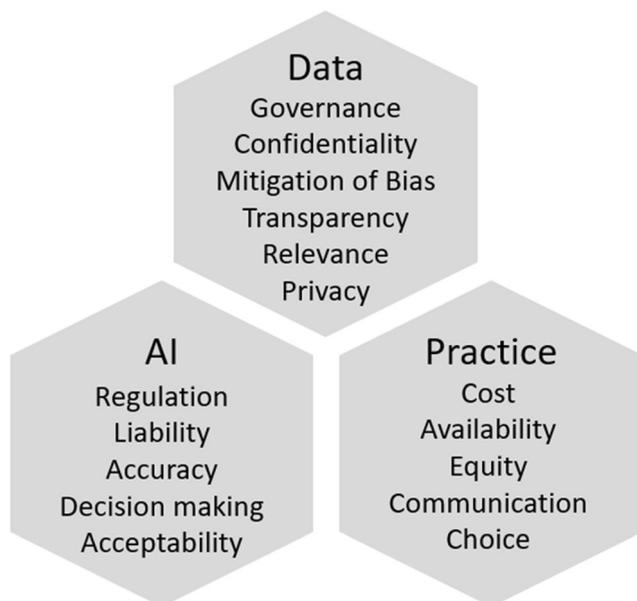


Fig. 1 Ethical and social considerations for AI in Nuclear Medicine

designed to assess the ability of AI to think like a human, we may be concerned about reaching a point where AI aces the Turing test, but some consideration needs to be given to reaching the point where SI consciously fails the Turing test. Does this redefine our legal, moral and ethical guidelines?

Does a patient deserve the right to refuse diagnosis or treatment augmented by AI (11)? Morally and ethically, is this any different to refusing any form of care? One presumes the implementation of AI in a patients' care extends some enhanced benefit for patient outcomes. The "first do no harm" mantra equally applies to creating harm by introducing AI as it does create harm by withdrawing it, indeed, creating harm by having inequitable access to it. Nonetheless, patients have the general right to have healthcare individualised to accommodate cultural or philosophical preferences. This may include AI driven healthcare but should be distinguished from a right to have decision-making made by the human physician. Can AI or SI be the subject of discrimination? If AI or SI are capable of super-logic and synthetic consciousness, given the discussion above with respect to social norms, would refusal to receive care from an AI-/SI-based system on those grounds alone be any different to refusing care from a healthcare practitioner based on their gender or ethnicity? Such questions need to be considered.

With specific reference to the application of commercial algorithms, care needs to be taken to ensure the reference population and subsequent data analysis has ecological validity. A very robust DL algorithm trained and validated on a population with specific characteristics may not be externally valid to another population. Potential bias, errors and over fitting could result (5). Insufficient representation of a particular population (e.g. minority, vulnerable group, pathology subtype, etc.) in the training and validation data may not only create bias and error in predictions but could contribute to widening of the gap in health equity. This requires a number of key considerations. Firstly, before utilising an AI algorithm, one must have some sense of the training data relative to the data being entered. Secondly, consideration needs to be given to ecological validity when considering the output from an algorithm. Thirdly, currency of algorithms needs to be maintained to ensure the historical training data remains current for today's data. Finally, consideration should be given to technology-based bias or error where there is a difference between training data and the equipment used for the actual patient data. This is especially important given that the neural network function is often like a magician's box; we see what goes in and what comes out but are not really sure what happens in the box itself. Indeed, the real magic may lie outside the box. Ethically this raises questions about transparency, justification and knowledge sharing. With unsupervised learning in particular, the scope of operation and conditions associated with that are not defined by the users but rather it is extracted from the data itself. The ability to audit this process is critical to instil

confidence in the AI outputs, to improve quality assurance of the AI algorithm and to enable human learning from the AI.

The group AI4People, a committee of the Atomium-European Institute for Science, Media and Democracy (EISMD) outlined a framework in 2018 for ethical application of AI generally in society (12). The general principles include:

- Beneficence to promote well-being, preserve dignity and for sustainability
- Non-maleficence to provide privacy, security and "do no harm"
- Autonomy associated with the power to decide or otherwise through human communication
- Justice to ensure equity and fairness associated with accessibility and outcomes
- Explicability to enable the principles above and create intelligibility and accountability

Focused to the field of Nuclear Medicine, therefore, some consideration needs to be given to the disruptive nature of AI with respect to human relationships (staff/staff and staff/patients), health inequity (positive or negative), decision-making (human, AI or hybrid) and regulation. Furthermore, data use, storage and sharing, algorithm transparency and reliability and the necessity (medical benefit and patient preferences) are key ethical and social considerations for AI use in Nuclear Medicine. These are largely deconstructed and captured in the ethical principles below.

Ethical principles in AI and SI

Based on the discussions above, we propose 16 ethical principles hereby outlined, in order to guide development and implementation of AI solutions in Nuclear Medicine research and clinical practice. These 16 Principles in Intelligent Imaging Ethics (16PI²E) provide the standards by which the needs and safety of both the patient and the health practitioner are respected and preserved.

While these principles set a standard for development and implementation of AI solutions in Nuclear Medicine, they are not designed to be interpreted independently of the Nuclear Medicine professional ethical codes, codes of practice, patients' bill of rights, regulatory/legislative requirements, nor general ethical principles in health and medicine. Rather, these principles are a conduit between the opportunities of emerging disruptive technologies and those other well-established ethical principles. Recognition must be made that, depending on specific activities around AI, some principles may be of more immediate importance than others amongst individuals, but all principles should be considered at all times. The following 16PI²E (Fig. 2) should be applied in consideration of the 23 AI principles developed at the Future of Life Institute Asilomar Conference on Beneficial Application of AI (13).



Fig. 2 Summary of ethical principles (16PI²E) that should guide the use of AI/SI in Nuclear Medicine

Beneficence

AI/SI solutions should be designed and implemented for the common good to benefit humanity by some measure.

Non-maleficence

When designing and implementing AI/SI solutions, outcomes, care and treatment (including costs) should not be worse with the introduction of AI/SI.

Fairness and justice

AI/SI solutions should be designed and implemented with processes in place to ensure that algorithms treat all patients fairly and equitably.

Safety

When designing and implementing AI/SI solutions, priority must be given to maintaining and demonstrating evidence of patient safety and quality of care. Healthcare providers must be properly trained and enabled to safely integrate AI/SI solutions into their practice.

Reliability

AI/SI solutions must be designed to be reliable and reproducible when implemented, including ecological validity. AI/SI mechanisms should have methods in place for quality assurance and evaluation of accuracy of performance.

Security

When designing and implementing AI/SI solutions, all data must be stored and moved securely within the scope of regulatory requirements. Data should not be transferred outside the physical/electronic bounds of the healthcare provider without patient consent and ethics approval.

Privacy and confidentiality

AI/SI solutions should be designed and implemented in a manner that allows all data to be de-identified, maintaining privacy and confidentiality for big data, cross institutional collaboration and commercial application of AI algorithms.

Mitigation of Bias

When designing and implementing AI/SI solutions, rigorous, evidence-based clinical trials must be applied, the data on which algorithm training and validation are based must be transparently valid for target populations, and all limitations and potential bias must be transparently reported.

Transparency and visibility

AI/SI solutions should be designed and implemented to provide transparency to patients and service consumers with respect to the use of and reliance on and input from AI/SI solutions and how predictions are convolved in the algorithm (magician's box). Patients should maintain autonomy to understand how AI/SI solutions are being used for their care and in decision-making.

Explainability and comprehensibility

When designing and implementing AI/SI solutions, any impact on patient care, diagnosis or treatment must be understood and clearly explainable/justified. AI/SI systems should be designed to explain their reasoning and allow humans to interpret their outputs.

Human values

AI/SI solutions should be designed and implemented with a human-in-the-loop process incorporated, where reasonable, to apply humanitarian values, accommodate patient values and preferences (including social and cultural norms) and to augment AI/SI predictions.

Autonomy, judgement and decision-making

When designing and implementing AI/SI solutions, a human-in-the-loop process must be incorporated to ensure judgement,

and decision-making in relation to patient care accommodates the patient presentation, history, findings and preferences following a conversation between patient and healthcare provider.

Collegiality

AI/SI solutions should be designed and implemented for the optimisation of outcomes through multidisciplinary collegiality, collaboration and leveraging the unique capabilities of team members in the AI/SI pipeline.

Accountability

When designing and implementing AI/SI solutions, recognition must be made of the shared accountability amongst stakeholders and documented prior to development of AI/SI solutions; accountability should not lie on the end user interpreting the AI output alone.

Governance

AI/SI solutions should be designed and implemented within a framework of overarching governance to ensure compliance with ethical principles, regulatory requirements and professional standards.

Inclusiveness

When designing and implementing AI/SI solutions, there should be engagement and empowerment of all stakeholders and, in the case of disruptive technology, minimisation of impact or displacement on workers.

Conclusion

The most profound impact of AI applications in Nuclear Medicine with respect to the patient experience arise from the capacity of AI to enable deeper, more meaningful interactions between the patient and the physician. The paradigm shift created by applications of AI in Nuclear Medicine has the potential to inspire a culture shift where emphasis is increased on the value of the interaction between the patient and the provider. AI offers a powerful toolkit for the rapid and safe automation of tedious or repetitive tasks and deep analysis beyond the capability of the human mind and, in doing so, provides time and energy for the physicians to more expertly review patient data and to communicate with the patient (5). Integrating AI solutions into the patient care solution has the potential to impact the patient/physician trust dynamic (14). From data privacy and security, through potential misuse to an

augmented balance of shared accountability and risk, use of AI in practice alters the patient physician dynamic. AI, while transformative, presents a number of ethical or social challenges that require careful attention and formulation of guidelines and policy documents.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Ethical approval This article does not contain any studies with human participants or animals performed by any of the authors.

References

1. Currie G. Intelligent imaging: artificial intelligence augmented nuclear medicine. *Journal of Nuclear Medicine Technology*. 2019;47: 217–22.
2. Currie G. Intelligent Imaging: anatomy of machine learning and deep learning. *Journal of Nuclear Medicine Technology*. 2019;47(4):273–81.
3. Currie G, Hawk KE, Rohren E, Vial A, Klein R. Machine learning and deep learning in medical imaging: intelligent imaging. *Journal of Medical Imaging and Radiation Sciences*. 2019;50(4):477–87.
4. Hamlet P, Tremblay J. Artificial intelligence in medicine. *Metabolism clinical and experimental*. 2017;69:S36–40.
5. SFR-IA Group, CERF. Artificial intelligence and medical imaging 2018: French radiology community white paper. *Diagn Interv Radiol*. 2018;99:727–42.
6. Geis JR, Brady A, Wu C, et al. Ethics of artificial intelligence in radiology: summary of the joint European and north American multisociety statement. *Insights into Imaging*. 2019. <https://doi.org/10.1186/s13244-019-0785-8>.
7. Jalal S, Nicolaou S, Parker W. Artificial intelligence, radiology, and the way forward. *Can Assoc Radiol J*. 2019;70:10–2.
8. Balthazar P, Harri P, Prater A, Safdar NM. Protecting your patients' interests in the era of big data, artificial intelligence, and predictive analytics. *J Am Coll Radiol*. 2018;153:580–6.
9. Jaremko JL, Azar M, Bromwich R, et al. Canadian Association of Radiologists white paper on ethical and legal issues related to artificial intelligence in radiology. *Can Assoc Radiol J*. 2019;70:107–18.
10. Kohli M, Geis R. Ethics, artificial intelligence, and radiology. *J Am Coll Radiol*. 2018;15:1317–9.
11. Ploug T, Holm S. The right to refuse diagnostics and treatment planning by artificial intelligence. *Health Care and Philosophy: Medicine*; 2019. <https://doi.org/10.1007/s11019-019-09912-8>.
12. Floridi L, Cowls J, Beltrametti M, et al. AI4People – an ethical framework for a good AI society: opportunities, risks, principles, and recommendations. *Mind Mach*. 2018;28:689–707.
13. Future of life institute. Asilomar AI principles, Asilomar Conference on Beneficial AI. 2017. <https://futureoflife.org/ai-principles/>.
14. Reddy S, Allan S, Coghlan S, Cooper P. A governance model for the application of AI in health care. *J American Medical Informatics Association*. 2019. <https://doi.org/10.1093/jamia/ocz192>.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.