The Royal Australian and New Zealand College of Radiologists®

5 April 2019

Therapeutic Goods Administration

By email: devicereforms@tga.gov.au

RANZCR Response to TGA on SaMD

About the Royal Australian and New Zealand College of Radiologists

The Royal Australian and New Zealand College of Radiologists (RANZCR) is the peak body advancing patient care and quality standards in the clinical radiology and radiation oncology sectors. It represents over 4,000 medical members in Australia and New Zealand.

The Faculty of Clinical Radiology is the bi-national body for setting, promoting and continuously improving the standards of training and practice in diagnostic and interventional radiology for the betterment of the people of Australia and New Zealand.

Clinical radiology relates to the diagnosis or treatment of a patient through the use of medical imaging. Diagnostic imaging uses plain X-ray radiology, computerised tomography (CT), magnetic resonance imaging (MRI), ultrasound and nuclear medicine imaging techniques to obtain images that are interpreted to aid in the diagnosis of disease. Interventional radiologists treat as well as diagnose disease using imaging equipment.

The Faculty of Radiation Oncology is the peak bi-national body advancing patient care and the specialty of radiation oncology through setting of quality standards, producing excellent radiation oncology specialists, and driving research, innovation and collaboration in the treatment of cancer.

Radiation oncology is a medical specialty that involves the controlled use of radiation to treat cancer either for cure, or to reduce pain and other symptoms caused by cancer. Radiation therapy is an effective, safe and cost effective method of treating cancer, and is involved in 40% of cancer cures. Unfortunately, while one in two cancer patients would benefit from radiation therapy, in Australia only about one in three will actually receive the treatment. The reasons for this underutilisation are a complex mix of lack of awareness of radiation therapy as a viable treatment option, physical access to a treatment centre, and patients not being provided with comprehensive information about all possible treatment options.

Introduction

RANZCR welcomes this consultation and the consideration being given by the TGA to the regulation of software as a medical device, and specifically artificial intelligence. Software used in medicine has advanced significantly in recent decades, a trend we expect to accelerate. There are complex interactions in decision-making between the clinician, the service provider and the software and the health system needs to ensure each component is regulated appropriately in line with their contribution to the service being provided.

Clinical radiology and radiation oncology are two areas of medicine that are data rich and already using advanced technologies and informatics software. Because of this, both are ready to adopt artificial intelligence (AI) and machine learning (ML). RANZCR believes that AI has enormous potential but could also do significant harm if left unregulated or operating autonomously in direct patient care.

RANZCR welcomes the approach taken by TGA considering the implications of AI and how regulatory mechanisms need to be revised with changes to technology. We would however caution against the creation of a route to market for all AI and machine learning tools in medicine when there is limited understanding of how this space will evolve in the coming years.

RANZCR commenced working on AI in 2016. Our initial focus was to understand the landscape and to inform our membership of advances in artificial intelligence and machine learning and to prepare the ground the significant changes we foresee. In November 2018, RANZCR organised Australia's first AI in healthcare conference called Intelligence18, which brought together international experts in AI to discuss the latest developments and implications for privacy and the practice of medicine.1 Also in 2018, the Faculty of Clinical Radiology established an Artificial Intelligence Working Group (AIWG) to consider the implications of artificial intelligence and machine learning on the potential implications for patients and on the disciplines of clinical radiology and radiation oncology and we plan a response that includes:

- how this technology can be applied appropriately and judiciously in the best interests of patients.
- appropriate education for members, trainees, stakeholders and the public

The AIWG has three main streams of work relating to AI ethics, the development of AI usage standards, and the skills that clinical radiologists will need to have to thrive in the future. The Faculty of Radiation Oncology is also giving consideration to the work of the AIWG and how it might be applied in radiation oncology.

RANZCR is proud that we are the first professional body in healthcare to have developed a set of ethical principles for AI. Having reviewed the literature globally and discussed the issues with experts in AI in medicine, the AIWG developed a set of eight ethical principles which are intended to ensure that AI and ML tools at all times reflect the needs of patients, their care and their safety, and they should respect the clinical teams that care for them. The ethical principles cover eight areas which include:

- Safety
- Avoidance of Bias
- Transparency and Explainability
- Privacy and Protection of Data
- Decision-Making on Diagnosis and Treatment
- Liability for Decisions Made
- Application of Human Values
- Governance.

Our draft ethical principles have recently been released for feedback from our members and stakeholders.[2] **RANZCR strongly believes that the regulation of AI needs a strong ethical underpinning and that developers of AI technologies in medicine should demonstrate in principle adherence to an ethical framework.**

One important issue that the AIWG encountered when considering the forthcoming changes is the importance of clearly defining AI, ML and associated terms to ensure there is a common

[1] https://www.eiseverywhere.com/ehome/index.php?eventid=349139&
[2] https://www.ranzcr.com/whats-on/news-media/307-ethical-principles-for-ai-in-medicine-consultation

understanding for discussion on this topic. We have produced a short paper to inform members about the current status of AI, called the State of Play which we will share with the TGA once finalised.

RANZCR is also considering what principles should guide the development of a regulatory structure around AI. Furthermore, we are revising our curricula to incorporate learning outcomes for our trainees to equip them to work alongside AI. We will also develop upskilling opportunities for practicing clinical radiologists and radiation oncologists.

RANZCR would appreciate an opportunity to meet with the TGA to hear your perspective on our work on AI and discuss how we might collaborate on this important matter.

Consultation Questions

1. **Do you support the proposal to change the way medical device software is regulated? Why or why not? If you do not support the proposal, do you have any suggestions for an alternative that would be acceptable to you?**

Overall, RANZCR supports the intent and direction of the proposed changes but has some concerns and recommendations regarding some specific areas including: classification; importation of software; the essential principles; safe deployment and monitoring; the definition of SaMD; and the transition.

Classification rules

RANZCR strongly supports the proposed changes to the classification rules, agreeing with the principle that the degree of scrutiny and evidence required to demonstrate the safety of a medical device should be commensurate to the risk that it poses to patients.

In particular, RANZCR agrees that *decision-making* devices are inherently different than other forms of medical devices, and can pose unique risks to patient safety. The degree of risk depends on several factors, including the **clinical importance** of the decision as described in the proposed changes. RANZCR supports this approach to determining risk.

The degree of **autonomy** of the *decision-making* device is also a key factor in determining the risk, and RANZCR supports the incorporation of this concept in the proposed changes. If we consider a device that can diagnose a life-threatening condition, but does so under the supervision of an expert trained in that diagnosis, this may actually represent a lower risk to the patient than a system that autonomously produces a low to medium risk diagnosis. This concern is not far-fetched, given the FDA approval in 2018 of an autonomous screening tool for the assessment of diabetic retinopathy [3].

However, the classification of *all* devices that "aid a clinician in making a diagnosis" as low to moderate risk (Class IIa) is potentially problematic. In radiology, we have had extensive experience with this type of device, commonly referred to as "computer aided detection or diagnosis" systems. Historically, these systems have been widely used in the United States of America for screening mammography, approved by the FDA using a similar "low to moderate" risk classification. Unfortunately, the evidence that was used to prove safety (multireader

[3] https://www.fda.gov/NewsEvents/Newsroom/PressAnnouncements/ucm604357.htm

studies) was inadequate, and has potentially lead to patient [4] [5]. The interaction *between* humans and decision-assisting computer systems has been implicated, as prompts from computers have been shown to bias human decisions [6]. The validity of controlled multireader experimental testing has also been called into question.

The second factor that should be considered is the potential **scale** of the system. Software devices are rapidly scalable (i.e. they can be deployed quickly across multiple systems), with the potential to impact the care of a large population of patients in a short space of time. This is particularly relevant for devices that are intended for screening for common conditions, due to the disproportionate effect they can have at the population level.

At a population level, a high-risk but small footprint device only poses a modest risk to patient safety. Conversely, a low or medium-risk device which will be applied to millions of patients can cause an enormous amount of harm, even if the risk of any single individual is low.

Regarding devices which make **therapeutic decisions**, or **provide direct-to-consumer advice**, RANZCR supports the proposed changes.

Page 8 of the consultation documents outlines a proposed approach to classification of software. It is unclear whether there is expected to be medical oversight in each of the scenarios outlined e.g. for making a diagnosis or screening patients. We strongly believe that fully autonomous AI that can make clinical decisions without expert oversight constitutes a major risk to patients safety and should be treated with extreme caution, particularly given the current lack of experience and evidence supporting the safe use of AI in clinical practice. **RANZCR's fifth ethical principle outlines that liability for decisions made about patient care rest principally with the responsible medical practitioner in conjunction with the patient themselves after full and complete discussion.**

We would like to see explicit reference being made to medical (or if appropriate health practitioner) oversight in the variety of scenarios presented on page 8. We recognise that there may be instances where the medical practitioner is not directly involved, for example when the patient is monitoring their own condition, which also needs explicit categorisation.

Recommendation 1

RANZCR supports the TGA proposal to consider the spectrum of autonomy as a major factor when determining risk, but recommends that care be taken when classifying systems that aid human decision-making, as the historical evidence and human-computer interface literature suggests that the risks may be higher than expected.

Recommendation 2

RANZCR also recommends that the intended scale of the system is taken into account when determining the risk of decision-making devices. In practical terms this may mean devices which would be classified as IIa or IIb under the proposed changes, but have the potential to affect the care of a large group of patients, may need to be considered higher risk.

[4] Lehman CD, Wellman RD, Buist DS, Kerlikowske K, Tosteson AN, Miglioretti DL. Diagnostic accuracy of digital screening mammography with and without computer-aided detection. JAMA internal medicine. 2015;175(11):1828-37

[5] Fenton JJ, Taplin SH, Carney PA, Abraham L, Sickles EA, D'Orsi C, et al. Influence of computer-aided detection on performance of screening mammography. New England Journal of Medicine. 2007;356(14):1399-409

[6] Coiera E, Ash J, Berg M. The unintended consequences of health information technology revisited. *Yearbook of medical informatics* 2016; (1): 163

Importation of software devices

RANZCR supports the proposed changes around the importation of software devices, although in the field of radiology the end-user will almost always be a clinician rather than a patient, and so the risk is much lower.

One area that may need further consideration is the use of open-source software; that is, software which may be of clinical benefit but lacks a financial incentive for a sponsor or importer to apply for regulatory approval. A clearly defined exception for the use of this software in research or for safety and quality activities is recommended, to prevent the impression of a "ban" on the unregulated importation of this software for non-commercial purposes.

Recommendation 3

RANZCR supports the proposed changes to the regulation of the importation of software devices, but suggests that clearly defined exceptions are included for the use of open-source software in research and quality improvement activities.

Changes to the essential principles

RANZCR supports the proposed clarification of the "minimum requirements" of medical software, but is concerned specifically about the stated support for the updating of decision-making software, as well as the lack of a focus on traceability.

Traditionally, medical devices have been functionally "static"; that is, the effect on the patients does not change after marketing. Updating machine learning models to improve performance or remediate problems is likely a necessary part of the use of these systems, however the consequences of doing so can unpredictably alter the properties of the algorithmic decision maker and hence affect the clinical impact. It is an unsolved question whether an update of this kind can be safety neutral, or whether it may instead pose an unacceptable risk to patients. Options to deal with this may include: 1) simply allowing updates, 2) requiring close monitoring, or 3) requiring a new regulatory submission for each "new" version of the software. The latter method is currently in place in most jurisdictions, as any functional change to an approved device is usually considered to produce a new or different device.

RANZCR also believes that **traceability** should be an expected component of any medical software. Comprehensive record keeping of both the operation of the software as well as user choices should be an essential principle of safe practice, and every effort should be made to provide information to investigators when something goes wrong. This is essential to ensure that patient care can be improved and to prevent further patient harm. This record keeping should also include "model snapshots" which record the exact software used to make any clinical decision, so traceability is not impaired if the software is changed or updated.

Regarding future changes to the essential principles, it would be useful to see the revised principles in their entirety to judge their suitability.

As noted in the introduction, RANZCR believes that the manufacturer of the SaMD (that includes AI) should demonstrate that they have taken an ethical approach to develop the AI components for example by explaining how they have minimised the potential for bias or ensured that a discerning doctor can explain how a decision was made by an AI tool, in line with RANZCR's Ethical Principles for Use of AI in Medicine. In order to achieve this, we would like to see the following added to the essential principles:

"Artificial intelligence and machine learning tools should be developed in line with ethical principles" (such as RANZCR's Eight Ethical Principles on Use of AI in Medicine).

Recommendation 4

RANZCR recommends that the TGA considers the possibility of "non-static" software in the proposed changes; that is, software which can change over time, resulting in unpredictable changes in the safety profile of the device.

Recommendation 5

RANZCR also recommends that traceability is made a core principle of safe software design, such that any failure can be investigated to determine the contribution of each component.

Recommendation 6

A further essential principle should be added to state that:

"AI and machine learning tools should be developed in line with ethical principles (such as RANZCR's Eight Ethical Principles on Use of AI in Medicine).

Concerns regarding safe deployment and monitoring

RANZCR has specific concerns regarding the deployment of decision-making software in clinical practice. Given the reliance of machine learning systems on their training data, it is unclear how to ensure safety in clinical practice. It is unlikely that clinical populations will be sufficiently similar to training populations, and clinical populations rarely remain stable over time as demographics and disease distributions change (population drift). As such, the judgement that a machine learning system is safe at the time of regulatory approval does not guarantee the system will be safe in a new population, or even in the same population in the future.

There are several possible approaches to take with this issue. The first is to require demonstrations of safety to be performed across multiple sites, to demonstrate the performance of the system across different populations.

The second approach could be to require the demonstration of safety *locally* for each deployment. This is a more conservative approach, and carries a significant compliance burden for vendors.

The third option would be to require comprehensive post-market monitoring and ongoing review of safety following deployment. While this also increases compliance costs, it also allows vendors to deploy their devices prior to extensive local testing. This may be the most palatable option, and appears to be the current approach that the FDA and other similar organisations are taking.

If the third option were to be adopted, then it would be important to require clarity around the monitoring responsibilities of vendors and users (i.e., who will be legally responsible to ensure ongoing local safety), as well as having clearly defined failsafe mechanisms. For example, if a device is shown to be unsafe during monitoring then a well-defined "shutdown" process would be required, as well as a mechanism to remediate any problems.

Recommendation 7

RANZCR recommends that it be a requirement of regulatory approval that all decision-making devices have explicitly defined monitoring procedures, with clearly defined responsibilities for the vendors and users, with explicit failsafe mechanisms to prevent ongoing patient harm if safety targets are not met.

Concerns regarding the definition of SaMD

RANZCR supports the specific regulation for SaMD products, but is concerned that the TGA differentiates SaMD from embedded software; that is, software that is part of a physical device.

RANZCR recognises many situations where decision-making software is likely to be marketed embedded *in* physical devices, for example where machine learning algorithms form part of the traditional software that operates medical imaging systems such as MRI scanners or radiographic workstations. In this setting, it is plausible that the software could constitute a higher risk to patient safety than the device would otherwise, and it would be important that the software is regulated under the new guidelines. Closing this loophole would be important to prevent exposing patients to risks of software that is excluded from these regulations simply because it is bundles with a physical machine.

Recommendation 8

RANZCR recommends that caution is taken when differentiating between SaMD and "traditional" software on the basis of whether it is embedded or stand-alone. Any software that can make decisions should be regulated according to the proposed changes.

Concerns regarding the transition between regulatory regimes

RANZCR supports the consultative approach the TGA is taking, but also recognises that the field of medical software moves rapidly, and suggests that interim solutions may be required while robust new regulations are developed. We do not wish to stymie the development of systems that will improve care and management, and we can see a course that allows development while getting the principles and ethics firmly embedded.

In particular, RANZCR recommends the TGA reviews the approach being taken by the NZ Ministry of Health in their revised Therapeutic Products Bill.[7] We would particularly recommend looking at the new 'Type 4 product', which ought to be regulated but does not fit any of the existing categories of medicine or medical device. This allows the regulator to put some controls in place for such products that protect the public whilst regulation is being revised to account for innovative new products.

Other Comments

On page 5 of the consultation document, three key issues relating to the regulation of software are outlined. RANZCR believes there is a fourth dimension to consider in respect of liability for decisions made, particularly where AI can operate with autonomy. RANZCR's fifth ethical principle sets out that liability for decisions made should principally lie with the responsible medical practitioner, however this may be partly shared with health service executives or AI developers.

Post market surveillance is increasingly essential as SaMD becomes more and more sophisticated, with the potential of artificial intelligence to refine its outputs based on the case mix it is present with. This can lead to the performance of the same AI tool diverging between sites at which it is deployed. While that may be beneficial to the care needs at various locations, it also presents a risk that the AI tool will not perform as was expected when

[7] https://www.health.govt.nz/our-work/regulation-health-and-disability-system/therapeutic-products-regulatory-regime

deployed. The FDA in the USA has taken the approach of licensing a particular version of an AI tool when placed on the market, and requiring further versions or upgrade to be reassessed. RANZCR would like the TGA to give consideration to this.

RANZCR recommends the TGA reviewing the approach being taken by the NZ Ministry of Health in their revised Therapeutic Products Bill.[8] We would particularly recommend looking at the new 'Type 4 product', which ought to be regulated but does not fit any of the existing categories of medicine or medical device. This allows the regulator to put some controls in place for such products that protect the public whilst regulation is being revised to take account of the innovative new product.

2. **What do you consider to be the benefits and disadvantages of the particular proposals for change?**

Please refer to our response to Question 1.

3. **Do you believe there will be any unintended consequences arising from the proposed changes?**

Please refer to our response to Question 1, particularly in relation to the software being used having a higher risk category than the medical device it is embedded within.

RANZCR is also concerned that SaMD is evolving rapidly, particularly in relation to AI and machine learning. At this point, it is difficult to anticipate how these tools will evolve and how healthcare services will transform as a result. RANZCR feels it would be prudent to review these changes within a short timeframe such as two years.

4. **What changes would you need to make (if any) to meet the new arrangements? If not, what are the impediments?**

Not applicable to our response.

5. **What financial impact (both costs and savings) would implementing the proposed amendments have for you? If possible please provide a breakdown of the impacts. This information will be used to quantify the financial impact to all affected stakeholders.**

Not applicable to our response.

6. **What period would be needed for your organisation to implement the proposed changes? This information will be used to inform any transitional arrangements.**

We believe that these changes to the regulatory environment are very important and should not be rushed. RANZCR is available to assist the TGA with any further consideration of the issues we have raised above.

Regarding artificial intelligence, it is important that standards are in place and the workforce is prepared for the deployment of AI within medicine. As noted in the introduction, RANZCR has a broad program of work through 2019 to develop and publish standards of practice relating to the use of AI in clinical radiology and radiation oncology. We hope to have these in place by early 2020. RANZCR is also working to revise our curricula for clinical radiology and radiation

[8] https://www.health.govt.nz/our-work/regulation-health-and-disability-system/therapeutic-products-regulatory-regime

oncology. Updating our curricula is a complex task involving multiple stakeholders. We are aiming to have these published in 2021.