Preamble

The Department of Health and Aged Care (the Department) welcomes the discussion paper released by the Department of Industry, Science and Resources (DISR) on Safe and Responsible AI in Australia, and provides the following submission:

Artificial Intelligence (AI) is an emerging capability that has the potential to transform wide areas of the economy and improve lives. It is expected to have significant impacts in the healthcare sector.

AI is advancing quickly and will likely generate disruptive innovation across many parts of society which has witnessed significant advancements and applications of AI in recent years. AI experts, journalists, policy makers and the public are increasingly discussing a broad spectrum of important and urgent risks from AI. The *Centre for AI Safety* recently called for global coordination to mitigate these risks[i] and the Digital Health Cooperative Research Centre having recently developed a comprehensive ethical framework for the responsible design, development, and use of generative AI technology in health and medicine[ii]. Any future regulation will need to balance the need to ensure patient safety and the need to maintain the security of protective sensitive health information and community trust with the health and economic benefits that may be realised from AI innovation.

The Department supports the development and implementation of policies and governance that promote safe and responsible AI in Australia. As healthcare delivery occurs at different levels of government, a national approach to AI governance, which includes sector-specific governance of AI in healthcare, is desirable to ensure alignment in policy and legislative development, clinical safety, and public health delivery prioritisation. The success of AI in healthcare will depend on national leadership to maintain trust and ensure these systems are safe, reliable, and understandable in how they work.

The responsible adoption of AI in government entails developing comprehensive policies, fostering trust and partnerships, and implementing strategies that communicate the benefits and regulate the use of AI effectively in various sectors. The Department would like to see a well-informed, ethical, and comprehensive approach to AI integration in the health sector. By understanding and managing risks, fostering inclusivity, and ensuring accountability. The responsible utilisation of AI technologies can contribute significantly to advancing healthcare and aged care services in Australia. This would be complemented by further education and training in relation to AI for the health workforce, policy makers, and the broader Australian public.

Australians have high expectations of the Department in the handling of sensitive information, including defining what data can be shared, with whom, and under what circumstances. Given this role in data sharing, the Department advocates for regulatory reforms and integration of AI-specific regulations within existing Acts. This would provide an avenue to harness the benefits of AI, while effectively managing its risks and protecting the integrity of healthcare information.

The Department also has a key role in providing equitable access to health interventions and services through programs such as Medicare and supporting the national health system in collaboration with states and territories. With a focus on keeping Australians healthy and safe, we recommend the regulatory approach to AI consider how AI impacts wider society as well as those who may be within vulnerable or marginalised communities. This requires approaches for addressing bias and fairness of AI technologies and associated data, and managing their ongoing use, to ensure AI is being used safely and

responsibly. Furthermore, transparency of AI in the delivery of health care is essential and this should be consistent across public and private sectors.

The Department supports greater ethics consideration when using data for AI purposes, particularly as it relates to health outcomes to enable maximum benefits while ensuring there is sufficient trust in the outcomes and how it affects individuals and society. For areas with direct impact on health outcomes of individuals, there is less tolerance for risk and the response should be proportional to the possible impact.

The Department supports the current approach by the Therapeutic Goods Administration (TGA) who regulate products that are intended for medical use including software (that incorporates AI), with a robust regulatory framework for software based medical devices. The framework addresses risks associated with AI and applies to any software included with, or that is a part of, a medical device that is used for diagnosis, prevention, monitoring, treatment, alleviation of disease, injury or disability.

The TGA regularly consults on its regulations to ensure it considers emerging technologies (and risks) to ensure the regulations remain fit for purpose and continue to safeguard Australian patients. The TGA publicly consulted on software including AI in 2019 and 2020 – and published updated specific guidance including clinical evidence and performance requirements in early 2021. Further information about the framework and risk classification with some examples is included in the attached Health response [Regulation of Software-based Medical Devices - Info sheet for DISR July 2023].

The Department does not recommend banning the use of high-risk AI applications, rather DISR may consider developing guidance on how to use controls to mitigate risk appropriately. The Department strongly advocates for a risk-based approach in relation to AI and recognises that it may need to be mandatory for moderate to high-risk applications in health and aged care. This provides flexibility to ensure regulatory burden and oversight align with the potential risk of a particular activity, and to reduce burden and promote innovation for low-risk AI applications. Key elements of a risk-based approach should include clear definitions of consequences of the risk and objective, clearly articulated criteria to determine the risk level and how to appropriately deal with the risk. Leveraging existing risk-based approaches, integrating AI-specific risks and controls into risk management, and employing a mix of regulatory and non-regulatory frameworks can support the development of a risk-based approach for addressing AI risks.

The Department recommends DISR considers, in partnership with appropriate regulators such as the National Data Commissioner and the Information Commissioner, the development of guidelines for Data Impact Assessments (DIA) as part of AI assessments. The DIA could be mandatory for organisations applying AI above a set impact threshold, similar to Privacy Impact Assessments. The DIA could take a multi-faceted approach taking into account the purpose, explainability, ethics, sensitivity, sovereignty, security, and impact to provide a holistic assessment of risk and need for regulation.

Additional input and detail on the discussion paper has been provided via direct responses to the 20 discussion questions. This input differentiates between issues directly related to the Department in comparison to the broader Australian Healthcare System.

[i] Center for AI Safety (CAIS)
[ii] https://www.thelancet.com/journals/ebiom/article/PIIS2352-3964(23)00077-4/fulltext and https://www.thelancet.com/journals/ebiom/article/PIIS2352-3964(23)00237-2/fulltex

Safe and Responsible AI in Australia – Discussion Paper August 2023

	Question	Department of Health and Aged Care	Australian Healthcare System
	DEFINITIONS		
1	Do you agree with the definitions in this discussion paper? If not, what definitions do you prefer and why?	Whilst a high-level definition could be useful, the way AI is used is context specific, as different sectors have differing needs. The Therapeutic Goods Administration (TGA) already has established definitions related to AI for medical (therapeutic) use which are aligned with the definitions in "Machine Learning-enabled Medical Devices: Key Terms and Definitions" published by the International Medical Device Regulators Forum (IMDRF) in May 2022. The Australian Commission on Safety and Quality in Health Care (ACSQHC) supports the use of the International Organisation for Standardization definitions. It is recommended to also include a definition for Automated Decision Making (ADM) to avoid misinterpreting ADM as fully autonomous. From a legislation perspective, it is recommended to keep the definition as technologically neutral as possible and consider future use-cases for AI to help maintain relevance.	No objection.
2	What potential risks from AI are not covered by Australia's existing regulatory approaches? Do you have suggestions for possible regulatory action to mitigate these risks?	The Department suggests that there is no clear pathway for the sharing of sensitive unit record health and aged care data with commercial entities within our current legislative frameworks. The existing regulations only permit the disclosure of critical departmental data, such as the Australian Immunisation Register, Medicare Benefits Scheme, and Pharmaceutical Benefits Scheme data, to a limited number of trusted Commonwealth agencies, including the Australian Bureau of Statistics (ABS) and the Australian Institute of Health and Welfare (AIHW). Similar disclosures to commercial entities are not permitted under either the primary legislation or the Data Availability and Transparency Scheme (which does not cover private sector firms). To effectively manage risks and protect the integrity of healthcare information while harnessing the benefits of AI, the Department advocates for comprehensive regulatory reforms and proposes integrating AI-specific regulations within existing Acts. These reforms would necessitate mandatory training for healthcare professionals and adherence to specific AI-related professional standards, ensuring that AI is utilised responsibly and ethically within the health sector.	In considering the development of AI regulation we need to consider a wide array of clinicians and professional bodies across the health and aged care sectors and their risk appetites. Over the past few years, lots of work has been undertaken to strengthen regulation and safeguards of Australia's critical infrastructure. Due diligence to ensure that there is no erosion of other forms of legislation would be essential.

		The Department emphasises the need to implement flagging mechanisms or additional security checks for medical professionals and researchers seeking access to sensitive data. This would ensure accountability and reduce potential security risks associated with the access and use of sensitive health information. Addressing bias in AI algorithms is of utmost importance to avoid disproportionate impacts on vulnerable populations, including First Nations and CALD and LGBTQIA+ and people with a disability. Data sets on which AI tools are trained, do themselves have inherit bias, (ie male skewed, no comprehensive data on women, gender reduced to the binary) making some groups "invisible" to the algorithm. Prejudice cannot be coded out of a model, but proper representation for minority groups can be taken into considered in setting up AI modelling. To support the empowerment of First Nations communities, The Department stresses the significance of adhering to Priority Reform 4 of the Closing the Gap Agreement. This reform aims to grant First Nations people the ability to collect, analyse, and use data in meeting their community's unique needs and priorities. Respecting data sovereignty rights and fostering genuine partnerships between the government and First Nations people are critical principles that must be incorporated into the development of AI technologies and regulatory frameworks. Complying with the CARE principles of Indigenous Data Governance further reinforces the commitment to fair and ethical AI practices. Having approaches for ongoing monitoring and potential re-training of AI technologies and models is essential to ensure their relevance and performance over time. Implementation of an AI tool without ongoing review can result in performance deterioration over time due to data drift, where the data used to train the model is different to the data where the model is being applied. This could have serious implications if the AI tool is being used in a high-risk sector such as healthcare, and evidence suggests it is already a concern in medical machine learning deployment[1]. The Department would like the longer-term management and use of AI tools be considered at project inception, including responsibilities for managing the tool, and mechanisms for detecting and mitigating data drift and performance degradation. The Department emphasises the significance of AI applications being available in multiple languages, ensuring inclusivity and accessibility for diverse populations, particularly culturally and linguistically diverse (CALD) communities. Addressing potential racial discrimination in AI,	

[1] https://www.birpublications.org/doi/10.1259/bjr.20220878

		particularly concerning recidivism predictions, demands meticulous examination and regulatory intervention to uphold fairness and justice. The TGA have been regulating products that are intended for medical use including software (that incorporate AI) since 2002, using a robust regulatory framework for software based medical devices. The framework addresses risks associated with AI, and applies it to any software included with, or that is a part of, a medical device that is used for diagnosis, prevention, monitoring, treatment, alleviation of disease, injury or disability. Regulatory requirements are technology agnostic and apply regardless of whether the product incorporates components like AI, chatbots, cloud, mobile apps or other technologies. The TGA regularly consults on its regulations to ensure it considers emerging technologies (and risks) to ensure the regulations remain fit for purpose and continue to safeguard users. The TGA publicly consulted on software including AI in 2019 and 2020 – and published updated specific guidance including clinical evidence and performance requirements in early 2021. The TGA has also consulted with specific groups such as MSIA and relevant health professional colleges on specific types and uses of software. The TGA has a range of regulatory actions it takes when software or AI is not performing as intended or if a product is being supplied without appropriate regulatory approval. Further information about the framework and risk classification with some examples is included in Attachment C of the Departments response [Regulation of Software-based Medical Devices - Info sheet for DISR July 2023]. The National Mental Health Commission (NMHC) would like to ensure the development and utilisation of AI across Australian society does not result in people who experience mental ill-health being treated unfairly, or in other harms to the mental health and wellbeing of the Australian community. The Department highlights the specific risks associated with AI for individuals experiencing suicidality. Past incidents where AI inadvertently facilitated access to harmful information underscores the need for vigilance and prompt regulatory action to mitigate potential risks. The Department notes the discussion paper did not outline regulation regarding the use of AI in the health sector, including in supporting/replacing health workforce and points out there is a risk of AI decision making being relied upon in remote settings versus human decision making supported by AI in urban areas. Ensuring equitable access to healthcare, especially for rural and remote communities, requires the responsible integration of AI support. The Department	

		recommends regulations that cover various aspects of AI's impact, including clinical decision-making, medical report writing, pathology, and health administration. The Medical Workforce Policy and Strategy underscores the importance of flexible and adaptable regulations, considering the rapid pace of AI technology evolution. The Australian Digital Health Agency (ADHA) suggests that AI presents a risk to the maintenance of quality healthcare information, which is currently not covered by Australia's existing regulatory approaches. The maintenance of healthcare information is currently governed by professional standards and regulatory frameworks that promote accuracy, quality, and handling of personal health information, such as the *My Health Records Act 2012*, *Healthcare Identifiers Act 2010* and Privacy Act 1988. None of this legislation currently contemplates risks from AI nor the benefits, and although the My Health Records Act 2012 does allow for decision making using a computer program, there is no specific instruction on automated decision-making. The Government is considering reforms to each of these Acts so there are opportunities to include AI-specific regulation as necessary. Current mitigations would be limited to relying on healthcare professionals undergoing mandatory training and have them meet a set of professional standards. Another potential risk from AI that is not covered by Australia's existing regulatory approaches relates to the use of AI in policy development and administration in the Australian Public Service. Clarity and disclosure around the use of AI data and algorithms, and the limitations of these methods, is crucial when AI outcomes are used to develop policy. The Agency welcomes the recently published *Interim guidance for agencies on government use of generative AI platforms* and recommends the development of more detailed guidance in the future, particularly in relation to policy development and administration.	
3	Are there any further non-regulatory initiatives the Australian Government could implement to support responsible AI practices in Australia? Please describe these and their benefits or impacts.	The Department suggests establishing nationally agreed AI principles as well as nationally agreed ethical, clinical and technical standards for AI. It is important to develop these national principles and standards using a transparent, co-designed and consensus-based approach (and leveraging international standards where appropriate) to support community trust and confidence in AI. It acknowledges efforts by the NSW Government in developing an AI Assurance Framework which mandates ethical principles to govern bespoke AI systems. It would be appropriate to consider if the NSW Framework can be applied nationally. The Department encourages partnerships with academic researchers and centres of excellence, such as the National AI Centre, and close monitoring of the Responsible AI Adopt Program, CSIRO Data61, and UTS Australian AI Institute to facilitate innovation, knowledge sharing, and resource utilisation. It also encourages an approach to AI governance that considers publicly available	• Likely need for public Education campaigns to provide education of the risks and benefits of the use of AI to ensure community and clinician awareness and trust levels remain high. • Consider establishing controlled environments for developers to test AI systems to identify risks and mitigation strategies before AI systems are released for use by Australians.

		information both nationally and internationally. It will be important to learn from international examples such as the European Union (EU) and Canadian approach and incorporating international guidelines, standards, and certifications into a national AI framework will ensure responsible adoption of AI technologies. Supporting AI developers and users is a priority, and diligently investigating tools developed by countries like the US and Singapore to identify and mitigate AI-related risks effectively should be considered. The work of the ACSQHC may assist in supporting the operationalisation of the 5th principle of the *AI Ethics Framework - Reliability and safety* - in the context of healthcare safety and quality. As part of its work plan, the Commission is developing resources to assist health services to evaluate and assess AI before the widespread uptake of these technologies. The resources aim to enable the safe implementation of AI into clinical practice and drive measurable improvements in the quality of patient care and outcomes. The Department recognises the potential need for additional regulatory and governance responses to ensure appropriate safeguards are in place and suggests accrediting the overarching governance processes of vendors developing AI technology, along with their device and software offerings, to instil trust and confidence in AI applications. The use of software and AI that performs a medical purpose, can be enhanced through further education of relevant health professional colleges and boards, higher education systems and training. Broader benefits could be gained by providing more accessible consumer and other stakeholder education as part of the government response to AI practices. These communication activities should include ensuring those products that are used for medical purposes have relevant TGA approval and that consumers understand the implication of AI and use of their personal information. There are many new stakeholders entering the market who are unaware of existing regulatory obligations and who do not fully understand their ongoing responsibilities. The TGA partners with ANDHealth to deliver webinars and education initiatives targeted to new entrants including those seeking to commercialise their product. Within the Australian Public Service (APS), the Government could consider performing an audit of current automated processes that use AI in order to promote transparency and ensure AI is being used safely and responsibly within the APS. To effectively manage AI-related procurements, upskilling of government officials will be required and close monitoring of procurements will be essential to uphold compliance standards and mitigate potential risks associated with AI implementation.	• Suggest encouraging positive incentives for compliance – accountability often build on penalties but incentives to reinforce safe and responsible use of AI could be introduced.

		Engagement of the health and aged care sector will be crucial to ensure equitable access to training and education related to AI technology, to facilitate wider acceptance and understanding.	
4	Do you have suggestions on coordination of AI governance across government? Please outline the goals that any coordination mechanisms could achieve and how they could influence the development and uptake of AI in Australia.	Whilst there is considerable activity related to AI occurring across government, more consideration is required to understand how AI will apply in the health and aged care contexts as there are unique ethical, legal, and regulatory challenges that must be addressed. Existing regulatory frameworks and legislation are not sufficiently developed for the full utilisation of AI and are likely to require significant reform. This is particularly so in relationship to risks to human health and around privacy and trust in the release of sensitive health data for use in the development of AI tools. To achieve a cohesive approach, coordination mechanisms need to establish consistency and coherence in AI policies and regulations across government departments and agencies. The coordination of AI governance also facilitates the effective sharing of knowledge and resources and encourages inter-agency cooperation. Government agencies could leverage diverse expertise and experiences to address challenges and capitalise on AI's opportunities. This knowledge-sharing approach nurtures innovation, accelerates AI adoption, and ensures that the technology aligns with Australia's unique needs. Common principles and guidelines will minimise potential inconsistencies and create an overall strategy for AI adoption in Australia. Sector-specific governance of AI in healthcare is essential given the unique risks, challenges and opportunities posed by AI in healthcare. The Australian Alliance for Artificial Intelligence in Healthcare's *Roadmap for Artificial Intelligence in Healthcare for Australia* could help inform Australia's approach to managing the opportunities and risks that AI brings. Building public trust and confidence in AI technologies will be important to ensure the uptake of AI solutions. Streamlining procurement processes related to AI technologies is also another crucial outcome of coordination mechanisms. A cohesive framework could guide government officials in AI-related procurements, ensuring compliance with ethical standards and mitigating potential risks. As a result, the smooth implementation of AI applications becomes feasible, driving effective and efficient use of the technology across government operations. Coordination mechanisms open avenues for cross-sector collaborations, bringing together government, academia, industry, and other stakeholders. This inclusive approach fosters strategic partnerships and joint initiatives, leading to transformative research and the development of AI solutions tailored to diverse societal challenges.	• As indicated in the paper, there are a range of existing regulatory frameworks that are relevant to AI governance. However, there would be value in coordination and information sharing in response to related issues. For example, privacy, copyright and online safety issues associated with the data used to train models will likely have similar intelligence and inquiry needs and would benefit from relevant instrumentalities having clear and strong channels for information sharing and referral of issues.

		DISR may wish to consider options for governance that supports ongoing inter and intra-governmental engagement on matters such as AI that are not industry specific – such as privacy or cybersecurity. The Department and the TGA already works with other governance bodies such as Office of the Australian Information Commissioner (OAIC) and the Australian Cyber Security Centre respectively, to ensure a cohesive approach to regulation. There is an opportunity for the Commonwealth to co-design with jurisdictions a national AI ethical and governance framework, in consultation with industry experts and the public. The new national framework would refresh the existing *Commonwealth AI Ethics Framework* and could be expanded to cover more specific guidance for key sectors in the economy where AI is already having, and will have, a significant impact, including healthcare. Such a framework could form the basis to support for future self-regulation or government legislation.	
	RESPONSES SUITABLE FOR AUSTRALIA		
5	Are there any governance measures being taken or considered by other countries (including any not discussed in this paper) that are relevant, adaptable, and desirable for Australia?	The Department is actively involved in ongoing surveillance of the international landscape with the view to identifying new and emerging governance measures, particularly in the EU, UK, Canada, and the USA. There is a need to understand the contextual differences between what is acceptable in these counties versus in Australia and learn from their experiences and evaluation of new initiatives. The TGA maintains a close relationship with other comparable regulators to ensure harmonisation of approaches. New requirements for software as a medical device (including AI) are emerging in different jurisdictions including Europe, Canada, UK and the USA. The European Medical Device Regulations (EU MDR 2017/745) has also introduced new requirements, including classifications specific for software as a medical device as part of a risk-based approach. The European rule is consistent with the IMDRF recommendations. Since 2013, the IMDRF (covering 10 regulators covering all major markets globally) has had in place, a dedicated working group reviewing Software as a Medical Device (SaMD), and AI – and publishes technical documents. The IMDRF undertakes global public consultation on all its work and participates in International Standards Organisation (ISO) standards, including those relating to SaMD. A governance measure that is not discussed in the consultation paper that might help inform Australia's approach to AI is the United States' proposed Algorithmic Accountability Act 2022. This legislation aims to establish a regulatory framework for assessing and mitigating bias and	• DoHAC is actively involved in ongoing surveillance of the international landscape with the view to identifying new and emerging governance measures, in particular in the EU, UK, Canada and the USA. There is a need to understand the contextual differences between what is acceptable in these counties versus in Australia and learn from their experiences and evaluation of new initiatives. Consultation with the health and aged care sectors will be essential to ensure that governance measures are fit for purpose.

		discrimination in AI systems used by large entities. It contains principles that might help Australia address algorithmic bias and promote fairness in AI applications. Consultation with the health and aged care sectors will be essential to ensure that governance measures are fit for purpose.	
	TARGET AREAS		

6	Should different approaches apply to public and private sector use of AI technologies?
	
	If so, how should the approaches differ?
		No. There should be no differences between use in private or public sector for software and AI used for medical purposes.
		
		Under existing arrangements, the private sector cannot obtain health data for AI research in preparation for commercial purposes. The private sector can however with universities and they can justify public interest considerations.
		
		Government public health data is largely administered (ie. without consent to gather data for Medicare and PBS) and the use of this data in potential AI applications raises additional ethical considerations that warrant thorough review and evaluation. Despite the potential risks, the use of sensitive personal information in the health and aged care sectors can provide a significant public benefit by improving policy and service delivery. Balancing these benefits against the inherent risks becomes crucial in contemplating specific regulations for AI activities in this domain, ensuring that they do not excessively burden existing frameworks and hinder genuinely public-interested activities.
		
		There are already sophisticated frameworks applied to the public sector that aim to ensure its focus on the public interest (including reconciling conflicts between ends and means), for example, parliamentary oversight and legislation, audit and other investigating bodies, FOI and transparency regimes, and public sector codes of conduct. As a result, there appear to be differences in the uses broadly considered acceptable by the public sector, including allowing for the public sector to engage in some higher risk applications of these technologies that would not be generally considered acceptable by the private sector.
		
		Sensitive personal information can be used for the purposes of improving policy and service delivery in the health and aged care sectors, which delivers a public benefit that can offset the risks inherent in using such information. The key challenge in considering any specific regulation of AI activities in this context, is that they do not create an excessive cumulative regulatory
			• In broad terms, the approaches to public and private sector use of AI should be similar in that the overarching principles should be similar and the risks posed by each sector's use are similar. However, the regulatory framework should be alive to the different kinds of conflicts of interest each sector may have in using AI technologies. In the case of the private sector, the typical concern is that the profit motivation will lead to misuse of these technologies and regulatory arrangements are crafted accordingly.
			• In the case of the public sector, the concern is that there may sometimes be a conflict between ends and means, that is:
			  ○ on the one hand, a public body's objectives (ends), which should be, by definition, in the public interest, and
			  ○ on the other hand, the public interest in the ways (means) in which public bodies conduct themselves being fair, honest,

		burden atop existing frameworks such that it is difficult to pursue genuinely public interested activities. Maintaining consistent ethical principles across all areas where AI technologies are deployed is critical particularly in healthcare, given the frequent transitions of care between public and private health services, as well as between care settings. These transitions involve the transfer of sensitive health information, demanding careful consideration of the interoperability of AI technologies to facilitate these transfers securely and efficiently.	ethical and in line with natural justice principles.
7	How can the Australian Government further support responsible AI practices in its own agencies?	Australians have high expectations of the Department in the handling of public information and building public trust will be critical. This requires careful consideration of data sharing practices to determine what data can be shared, with whom, and under what circumstances. AI and ADM's reliability is heavily dependent on the quality and fairness of the data on which it is trained, and health and aged care data often exhibits strong gender and cultural biases, necessitating substantial work to improve data quality and comprehensiveness before it can be utilised in real-world applications. Efforts are underway to develop standard authorisation provisions that facilitate greater data sharing and access. While synthetic data is often suggested as a remedy for privacy concerns, it is essential to recognise that its creation is resource-intensive and may still require some real data. If The Department legislation requires reform, the agency is prepared to contribute to this process. With the speed of emerging technology, it will be critical for government to ensure that adoption of any new technology will not compromise public safety. Government can continue to support responsible AI practices in its own agencies through establishment of guidelines for best practice and communication across portfolios to ensure a common understanding and implementation of priorities, rules, best practice, and investment where required. When evaluating AI model performance, considerations for vulnerable groups and clinicians are vital to ensure fairness and equity in the outcomes. To continuously improve the AI systems' impact and performance on the health workforce, regular evaluation and monitoring are essential. This includes assessing the effectiveness of AI applications, identifying, and addressing biases or unintended consequences, and adapting regulatory frameworks as needed. To support health professionals in making ethical decisions and ensuring patient-centred care when using AI systems, the provision of ethical-decision making support tools and guidelines is imperative. These tools will aid in navigating the complex ethical considerations that arise in the application of AI in healthcare. Government agencies should consider the recommendations	• Provide funding for grants through the Australian Government's $20 billion Medical Research Future Fund (MRFF) Applied Artificial Intelligence Research in Health. • AMA reports there is currently no national framework for an AI ready health workforce. Use AI requires retraining of the workforce, retooling health services and transforming workflows. The health systems is already resource constrained and such changes will not happen without strategic investment (AMA Journal 13 June 23).

		provided in the CSIRO Data 61 paper. This approach will foster the establishment of ethical guidelines and practices for the responsible use of AI in the health sector. Collaborative efforts involving health care professions, AI experts, policymakers, and regulatory bodies are essential to develop comprehensive AI governance frameworks that align with the specific needs and challenges of the health sector. To achieve this, The Department suggests prioritising investment in training and education to raise awareness of the use of AI and mitigate potential risks of harm to humans. To optimise AI implementation, it is crucial to separate regulation from policy development, project management, and service delivery, as the utilisation of AI in these different areas varies significantly. In addition to the recently published Interim guidance for agencies on government use of generative AI platforms, consideration could be given to the following activities: • Invest in training and education programs. For instance, further to priority 5 in the Australian Alliance for Artificial Intelligence in Healthcare's Roadmap for Artificial Intelligence in Healthcare for Australia, the training workforce for the use of AI should also include policy makers. • Establish an AI ethics review board or committee to oversee AI projects within government agencies. • Regularly assess and audit AI systems used by government agencies to identify and mitigate any bias, discrimination, or unintended consequences. • Collaborate with international organisations and governments to align responsible AI practices and standards and foster knowledge-sharing. • Establish mechanisms for reporting and addressing concerns or complaints related to AI systems used by government agencies. Mandatory reporting of AI use and activities (where relevant) through each Commonwealth entity's corporate plan to enable monitoring, oversight, compliance, and insights into evolving AI technologies.	
8	In what circumstances are generic solutions to the risks of AI most valuable?	Generic solutions and approach to AI are not suitable for AI used for medical purposes (ie: in medical devices).	• There will be little scope for generic solutions to risks of AI in healthcare settings. It will be important to have

	And in what circumstances are technology-specific solutions better? Please provide some examples.	In healthcare alone there are a range of different settings including but not limited to public and private hospitals, primary care, aged care, acute care and National Disability Insurance Scheme. Further information about how risk based classification rules for software as a medical device apply is in Classification of active medical devices (including software-based medical devices).	technological and human based solutions that maximise the benefits while reducing the harm.
9	Given the importance of transparency across the AI lifecycle, please share your thoughts on: a. where and when transparency will be most critical and valuable to mitigate potential AI risks and to improve public trust and confidence in AI? b. mandating transparency requirements across the private and public sectors, including how these requirements could be implemented.	a) where and when transparency will be most critical and valuable to mitigate potential AI risks and to improve public trust and confidence in AI? Transparency in the delivery of health care is essential. The Australian Charter of Healthcare Rights states consumers should be given clear information about their condition, the possible benefits and risks of different tests and treatments, so they can give their informed consent. Consumers also need to be advised if their data will be used for any future purpose. To support health professionals, it's important for patients/consumers to have transparency regarding the use of AI in their clinical/treatment decisions. Transparency and the ability to explain the recommendations or outputs of AI and ADM systems affects the level of trustworthiness people have in the use of AI technologies. Transparency and explainability are at times used interchangeably, but it is useful to separate them out, and discuss their differences. The Best Practice AI Regulation Toolkit notes that explainability is a requirement that goes one step beyond mere transparency. It seeks to explain how AI and ADM has been applied and why a particular outcome has occurred. For medical purpose devices that contain AI, transparency is critical in: • autonomous use where decisions are made primarily based on the AI output without any other contextual information to verify accuracy • adjunctive use where other contextual information such as patient symptoms, physical examination, lab testing or imaging are considered together with the AI output. b) mandating transparency requirements across the private and public sectors, including how these requirements could be implemented. There should be transparency consistent across public and private sectors and places the consumers of AI products at the centre of any policy or mandate. In the context of medical devices, manufacturers of a programmed or programmable medical device including those that incorporate the use of AI, must be able to demonstrate compliance	• Transparency could be partly addressed through regulation and obtaining consent from people for the use of AI technology in their diagnosis and treatment. • Transparency is important to ensure people are aware of the presence and function of AI in the products they are buying and the risks this may carry. Both the general public and clinicians need to be equipped to understand these risks. • Community awareness and education is important and CALD communities will need advice on AI in acceptable languages. • Mental Health settings and guidance will also need special attention due to sensitive nature of mental health issues and potential impacts.

		with the essential principles for safety, quality and performance. This is a legislative requirement set out under the Therapeutic Goods (Medical Devices) Regulations 2002. The essential principles include specific requirements to ensure that medical device software is designed and produced in a way that ensures the safety, performance, reliability, accuracy, precision, usability, privacy, security, and repeatability is appropriate for the intended use of the device, and that suitable information on the labelling and instructions for use is provided to the user. Where AI is part of a consumer's care it should be declared by health service organisations. Consideration should be given to regulating the use of statements, similar to privacy statements. These statements could advise the consumer what part of their care is supported by AI, what type of AI is used – static or dynamic and what safeguards are in place to ensure user safety. It is important that consumers are given information about the source of decisions made in their care. If AI is used in pathways of decision-making, it is important that the relevance of that decision making to the outcome of care is made transparent. AIHW engages with the community regarding their trust or otherwise in complex work, such as data integration that necessitates use, at least initially, of identifiable data. A key feature of the response is dependent on the uses to which those techniques are being applied. It is possible that the risks of using AI technologies would be approached similarly by the community and therefore transparency and a clear articulation of the benefits of such technologies will be crucial to maintaining community trust. This would include, for example, clear explanations of: • Why such technologies are necessary to the application in question and what specific public benefits they deliver over other approaches. • How the technologies have been designed, developed, deployed, monitored, and maintained. • What human oversight and checking there is of the results of the use of these technologies. Mechanisms are needed to encourage AI vendors to share or publish their research on how their tools address bias. Regular reviews of algorithms are required to ensure biases don't increase over time (which they will if left unchecked). How often reviews of algorithms are required should be determined by the risks related to the application.	
10	Do you have suggestions for: a. Whether any high-risk AI applications or technologies	a) The Department does not support blanket bans on applications or technologies without review of potential benefits and risks against existing, or new, regulatory frameworks. The Department continues to support risk-based approach to any regulatory frameworks.	• Rigorous testing by clinicians is required where AI systems are allowed to impact clinical decisions.

	should be banned completely? b. Criteria or requirements to identify AI applications or technologies that should be banned, and in which contexts?	Banning AI may be required for ethical reasons in some contexts, but it is more likely to be about upholding existing content areas bans. In a healthcare setting, it is not appropriate for AI to both predict and respond to a health situation without human intervention. Treatments recommended by AI may not consider the holistic needs of patients, such as their values or preferences. For example, an AI algorithm may recommend healthcare based on what will lengthen a patient's life expectancy, not taking into an account their preference for at home care or an ethical objection to certain treatments. Regulation is required for high-risk AI technology applications in healthcare delivery. Consideration could be given to the European Union's proposed AI Act, which lists high-risk AI systems including those relating to healthcare i.e. the management and operation of critical infrastructure, software for managing public healthcare services and electronic health records b) Criteria or requirements to identify AI applications or technologies that should be banned, and in which contexts AI systems deemed to be high-risk should be inspected if they are going to be deployed and the creators of the system should have to show that it was trained on unbiased datasets in a traceable way and with human oversight. Regulatory frameworks and risk assessments of new AI algorithms in healthcare are required, particularly for those that have the possibility of doing harm to human health. These assessments could utilise many of the elements of existing health technology assessment mechanisms such as Pharmaceutical Benefits Advisory Committee and Medical Services Advisory Committee. In a healthcare setting, we recommend considering patient safety and privacy requirements when identifying whether any high-risk AI applications should be banned completely. AI regulation is an area where the precautionary principle should be applied where potential harm to individuals or society is present, and the likelihood of such harm to materialise even where there is a paucity of evidence.	
11	What initiatives or government action can increase public trust in AI deployment to encourage more people to use AI?	There is a growing demand for appropriate policies in relation to the use of AI in government, to ensure alignment with the Public Service Act, the Public Governance, Performance and Accountability (PGPA) Act, APS values, and a risk-based approach. To facilitate responsible AI adoption, whole-of-government regulations and guidelines for the release and use of AI tools, accompanied by widespread publicity should be considered.	• See response to Question 9.

		Trust is a crucial issue, particularly concerning AI technologies, for First Nations peoples, who have a history of data collection that has often failed to benefit or, worse, negatively impacted their communities. Establishing trust with First Nations communities necessitates careful consideration of the CARE (CARE Principles for Indigenous Data Governance | ARDC) and data sovereignty (Delivering Indigenous Data Sovereignty | AIATSIS) principles. To foster genuine partnerships with First Nations people, decisions around data sharing must be made collaboratively and respectfully, acknowledging and respecting First Nations data holders' decisions not to share data in AI development It is possible to introduce AI technology implementation with lower risk or lower impact applications first and demonstrate the benefits of these initial applications to the public. This could help shift the prevailing discourse from a narrow focus on risks to a more balanced consideration of both risks and benefits. Successful applications of AI already in use can serve as examples. The National Science and Technology Council has recently provided the government with advice on the opportunities and risks associated with current AI technologies, underscoring the need for a thoughtful and well-regulated approach. To ensure the responsible and effective integration of AI in various sectors, the government can involve clinicians and consumers in the design and evaluation of AI services and products can enhance their acceptance and usability. Large-scale communication about the benefits of AI applications can help build public confidence. Incorporating AI into clinician education and skilling up the workforce, including requirements in clinical qualifications and undergraduate degrees, can bolster the capacity to leverage AI effectively. Accreditation or endorsement from trusted sources, such as TGA approval or meeting the Australian Digital Health Agency's conformance profiles, can instil confidence in AI applications. The government can also develop resources that operationalise regulatory or ethical principles to provide clearer guidelines for AI use. In addition to clinical applications, promoting AI utilisation for non-clinical purposes can also be beneficial. For example, leveraging AI to assist health services in evidence gathering for quality improvement, service planning/modelling, and accreditation can lead to more efficient and effective healthcare practices The government could encourage professional peak bodies and with patient-facing clinical organisations to increase familiarity, knowledge and skills of using AI for practitioners and for	

		patients – and encourage use of software and AI products that have relevant regulatory approval.	
	IMPLICATIONS & INFRASTRUCTURE		
12	How would banning high-risk activities (like social scoring or facial recognition technology in certain circumstances) impact Australia's tech sector and our trade and exports with other countries?		The Australian health system is primarily government funded and banning these high-risk activities would ensure its integrity remains intact.
13	What changes (if any) to Australian conformity infrastructure might be required to support assurance processes to mitigate against potential AI risks?	Entities that engage in the development of Australian infrastructure must report any use of AI technology in any stage of their processes. A declaration could be considered from an entity that all processes would not impact our national security or individuals. The Digital Health Agency supports establishing nationally agreed AI principles as well as nationally agreed ethical, clinical and technical standards for AI. These non-regulatory frameworks could help unlock benefits of AI in healthcare delivery, harness opportunities for innovation and promote safer and more secure data sharing practices. It is important to develop these national principles and standards using a transparent, co-designed and consensus-based approach (and leveraging international standards where appropriate) to support community trust and confidence in AI. In the digital health space, the Agency's Connecting Australian Healthcare – National Healthcare Interoperability Plan 2023-2028 outlines a national vision to share consumer health information in a safe, secure and seamless manner and identifies 44 actions across five priority areas relating to identity, standards, information sharing, innovation and measuring benefits. Priority area 2 references clinical decision support, a form of AI in its implementation. Exploration of how AI systems could support and enhance interoperability between clinical systems could be beneficial.	The health system is looking to AI in conformity infrastructure and so testing standards before deployment would be required to ensure against unintentional perverse outcomes within the sector. It also needs to meet requirements given that the health system is part of Australia's critical infrastructure.
	RISK BASED APPROACHES		

14	Do you support a risk-based approach for addressing potential AI risks? If not, is there a better approach?	The Department strongly advocates for this risk-based approach and the TGA currently has a risk-based approach to assessing and approving AI and any potential risks. A risk-based approach provides enough flexibility to ensure that both complex/sophisticated technologies and more simple technologies can both comply with the regulations. The regulations could be significantly strengthened by standards and accreditation, in alignment with regulation. A risk-based approach could apply to all AI applications, with increased risk based on potential harm to members of the public, or where there is no professional oversight. These frameworks should be continually reviewed to ensure they remain fit for purpose as technology emerges. Further information about risk-based classification of medical devices and classification rules with some examples are provided in the Health response attached [Regulation of Software-based Medical Devices - Info sheet for DISR July 2023]. The involvement of healthcare safety experts becomes crucial in understanding the clinical risks associated with the implementation of various technologies. To ensure its successful implementation, APS staff will require further training to effectively assess risks associated with AI applications. It will be essential to take into account the relevant concerns addressed in the Five Safes approach concerning data usage.	
15	What do you see as the main benefits or limitations of a risk-based approach? How can any limitations be overcome?	Risk-based approaches ensure that regulatory burden aligns with the potential risk of a particular activity for the Australian public and the oversight is proportionate to the level of risk. A strategic and methodical approach that is inclusive of key stakeholders will ensure limitations are actively managed and efforts are made to maximise benefits and reduce risk. **Benefits** • Improved patient safety: AI tools should be thoroughly tested, monitored, and reviewed. • Enhanced decision making: AI can aid clinical diagnosis and treatment. • Resource allocation: AI applications can utilise resources more efficiently if put in place safely and appropriately. • Regulatory compliance: ensuring ethical, safe, transparent, and accountable application. **Limitations** • Ensuring risk ratings are being accurately applied and assessed (the higher the risk the more regulation that is required). • Bias and discrimination can be perpetuated by AI systems leading to discriminatory, unethical and flawed outcomes.	

		• Underdeveloped risk assessments that do not accurately quantify risk, particularly due to the complexity and uncertainty of the technology. • Lack of standards, regulation, and guidance.	
16	Is a risk-based approach better suited to some sectors, AI applications or organisations than others based on organisation size, AI maturity and resources?	Risk-based approaches in health care may harness the resource-saving utility of AI while mitigating avoidable harm. A risk-based approach to applications of AI is suitable for the APS. These frameworks should be continually reviewed to ensure they remain fit for purpose as technology emerges. A risk-based approach is ideal for high-risk sectors with regulatory requirements, such as aged care, healthcare and critical infrastructure. As healthcare delivery occurs at different levels of government and within the private sector, it is highly desirable to have a consistent approach to public and private sector use of AI technologies to ensure consistent health outcomes for all patients.	• A risk-based approach to applications of AI in Healthcare is critical and essential. • Many of the risks associated with healthcare and human health will be 'high risk' due to patient safety, data sensitives and data sharing, use of diagnostic tools.
17	What elements should be in a risk-based approach for addressing potential AI risks? Do you support the elements presented in Attachment C?	The elements in Attachment C are largely supported. A risk-based approach requires clear definitions of the consequences of the risk and objective, clearly articulated criteria to determine the level of risk and how it is dealt with. Where possible, the criteria should be written in plain, non-technical English language that an ordinary person can understand. The elements presented in Attachment C, while serving as a foundation for risk assessment, are lacking the necessary level of detail to be effectively applicable to the health care sector and possibly other industries. In the context of healthcare, it becomes crucial to conduct a thorough review to determine if the implementation of AI has resulted in the replacement of human activities. If such a replacement has occurred, it becomes imperative to further review the AI system to ascertain whether it has brought measurable improvements and benefits to the safety and quality of the tasks or processes it is involved in. Furthermore, alongside the elements outlined in Attachment C, an essential aspect to consider in developing, scoring, and evaluating risk-based approaches is the established accuracy and effectiveness of the predictive modelling that underpins an AI system.	• Patient safety and the ethical use of AI is a key issue in considering risk identification, mitigation and management.

		In the pursuit of risk assessment for AI in healthcare, it is crucial to take into account the existing risk-based approaches applied to other health technology assessments. By building on these existing frameworks and regulations, it is possible to establish a comprehensive and coherent structure for AI risk evaluation, avoiding unnecessary duplication and streamlining the assessment process. To ensure a thorough and comprehensive risk stratification, it is essential to involve health care safety experts in the identification and assessment of clinical risks associated with AI technologies. Their expertise can significantly contribute to a more informed and nuanced evaluation of the potential risks and benefits. Fortunately, Australia has well-established assessment bodies, frameworks, and regulations in place. Leveraging these existing structures as much as possible can provide a strong foundation for developing a robust and tailored risk assessment framework for AI in healthcare. This approach allows for the incorporation of industry-specific nuances while benefiting from the knowledge and experience gained from previous health technology assessments.	
18	How can an AI risk-based approach be incorporated into existing assessment frameworks (like privacy) or risk management processes to streamline and reduce potential duplication?	This could be achieved by integrating AI-specific risks, controls, and references into risk management frameworks and privacy impact assessments, and fostering collaboration and training between relevant teams, such as privacy, risk, AI development, and cyber security. In respect to therapeutic goods, requirements for privacy are already incorporated into the medical devices regulatory framework (through the essential principles for safety, quality and performance). For medical devices incorporating AI, clinical and technical evidence needs to demonstrate the safety and performance of products to the same standard as any other (non-AI) medical devices. For higher risk products, clinical and technical evidence requirements are more stringent. The manufacturer of a medical device must be able to demonstrate the safety, quality and performance by providing documentary evidence which shows the medical device is designed and produced in a way that ensures the risks associated with the use of device are removed or minimised as far as practicable. The manufacturer is also required to ensure that privacy of the data or information is maintained. Any risks associated with the use of the device must be	

		acceptable when weighed against the intended benefit to the patient. Evidence to support this requirement must be available when requested.	
19	How might a risk-based approach apply to general purpose AI systems, such as large language models (LLMs) or multimodal foundation models (MFMs)?	A well-structured model may consider specific use-cases and assign them a higher-level of risk accordingly. For example, Privacy Impact Assessments typically identify higher risks in situations where there is a greater amount of personal information being utilised, thus providing a starting point for analysing such risks. The risk-based model could also consider how the LLM is being used. For instance, if the LLM is employed to generate clinical notes from a recorded medical encounter, then regulations should be based on the inherent risks involved in this particular use-case. To mitigate these risks, controls could be put in place such as a review of any information generated. When applying risk-based approaches to general purpose AI systems then it is important to consider data (including sovereignty), cybersecurity, misinformation and technology dependent risks. A privacy-preserving-based service architecture, e.g. federated learning, is urgently needed to be integrated with LLM and smartphone APPs to better protect user's privacy. Similar to other kinds of AI for software with an intended medical purpose, LLMS and MFMS should be subject to a risk-based approach that considers the consequence of using the product (i.e. the risk of harm and the need for safety and accuracy). When LLMs or MFMs have a medical purpose, they may be subject to TGA approval. Regulatory requirements are technology-agnostic for software-based medical devices and apply regardless of whether the product incorporates components like AI, chatbots, cloud, mobile apps or other technologies. In these cases, where a developer adapts, builds on or incorporates a LLM into their product or service offering to a user or patient in Australia, the developer is deemed to be the manufacturer and has obligations under section 41BD of the *Therapeutic Good Act 1989*. Technical information and clinical evidence must be available to the Australian regulator to demonstrate the safety and performance of the product using the LLM to the same standard as other medical devices – for higher risk products, clinical and technical evidence are required to be more stringent. Further information is published by Health on Artificial Intelligence Chat, Text, and Language.	**Clinicians** • Communication risk. People may use large language model-based chatbots (e.g. ChatGPT) for asking questions about diseases and treatments. How to align the health worker's explanation with the patient's knowledge that is mainly sourced from such models and other Internet channels. • Mental health risk. Does chatting with these models impact the mental health of teenagers or patients? **Public** • Digital divide risk. The advancement of technology has the potential to aggravate the Digital gap or Digital divide. • It is necessary to invest more resources in training and assisting Australians, especially ageing people to live with AI. • Employment risk. The latest AI technology will reshape many industry sectors. It is critically important to invest resources to re-train and re-employ people. • Privacy issues. The conversation between end-users and large language model-based chatbots will be sent to the server for processing. How these data will be processed and stored is

			unclear. Moreover, with the development of many start-up companies, the large language model-based chatbots has been extended to many other applications that might be further integrated with our smartphones, wearable devices, and computers.
20	Should a risk-based approach for responsible AI be a voluntary or self-regulation tool or be mandated through regulation? And should it apply to: a. public or private organisations or both? b. developers or deployers or both?	Decisions around voluntary or self-regulation versus mandated regulation should be based on the level of risk associated with a product or activity. For therapeutic goods, the regulatory framework for medical devices applies to all software-based products that meet the definition of a medical device, whether obtained from, or used within public or private organisations. Higher risk uses should be subject to a more substantial regulatory component. To establish a risk-based approach in healthcare, it is essential to implement a mandatory framework with limited or no voluntary options. One potential source of guidance for determining risk ratings could be the EU AI Act, which offers valuable insights in this area. However, while embracing this approach, it is crucial to carefully consider the cost implications associated with its implementation. If made mandatory, the government may need to allocate adequate resources to ensure that all sectors of the Australian community can participate equally. Particular attention must be given to supporting vulnerable groups, such as First Nations, people with disabilities, those in aged care, individuals with long-term health conditions, those facing mental health challenges, residents of rural areas, and those with limited economic resources. It is important to note that the extent to which the risk-based approach needs to be enforced could depend on the specific application of the technology. For instance, for less critical uses like an app or simple chatbot on a website, it might be more feasible to have a voluntary or self-regulated approach. However, in cases where the AI applications are utilized for clinical diagnosis and decision-making tools, a more rigorous and mandatory regulatory framework may be necessary. Throughout the process, various stakeholders play essential roles. Initially, developers are responsible for creating the product, and then deployers come into the picture when it comes to applying and using the technology. Establishing a risk-based approach in healthcare requires a well-balanced combination of mandatory regulations, cost considerations, and application-specific assessments, while	

		involving relevant stakeholders at each stage of development and deployment. By carefully linking these key ideas, we can foster a more efficient and inclusive healthcare ecosystem. The Department recommends employing a mix of regulatory and non-regulatory frameworks to enable a holistic approach in managing the risk of adverse consequences of AI in healthcare, as well as to harness any benefits and promote innovation. Non-regulation The Department supports establishing nationally agreed AI principles as well as nationally agreed ethical, clinical and technical standards for AI.[1] These non-regulatory frameworks could help unlock benefits of AI in healthcare delivery, harness opportunities for innovation and promote safer and more secure data sharing practices. It is important to develop these national principles and standards using a transparent, co-designed and consensus-based approach (and leveraging international standards where appropriate) to support community trust and confidence in AI. In the digital health space, the Australian Digital Health Agency's Connecting Australian Healthcare – National Healthcare Interoperability Plan 2023-2028 outlines a national vision to share consumer health information in a safe, secure and seamless manner and identifies 44 actions across five priority areas relating to identity, standards, information sharing, innovation and measuring benefits. Priority area 2 references clinical decision support, a form of AI in its implementation. Exploration of how AI systems could support and enhance interoperability between clinical systems could be beneficial. Regulation The Department supports having patient safety and data security regulation from the outset and measuring risk before deployment. This is important to manage risks of poor clinical outcomes for patients due to misapplication of AI or bias in training AI which can lead to patient harm and/or misdiagnosis. Regulation is required for certain high-risk AI technology applications in healthcare delivery. Consideration could be given to the European Union's proposed AI Act, which lists high-risk AI systems including those relating to healthcare i.e. the management and operation of critical infrastructure, software for managing public healthcare services and electronic health records. AI applications such as large language models (LLMs) which have a medical purpose may be subject to medical device regulations for software and require approval by the Therapeutic Goods Association (TGA). The TGA's Software as a Medical Device outlines how the TGA regulates software based medical devices.	

SOFTWARE BASED MEDICAL DEVICE REGULATIONS

Software is becoming increasingly important in medical devices and digital adoption more broadly. In addition, it is becoming more important as a medical device, in its own right.

Rapid innovation in technology has driven significant changes to software functionality and adoption, giving rise to a larger number of devices able to inform, drive or replace clinical decisions, or directly provide therapy to an individual. Advances in computing technology and software production have led to a large increase in the number of software-based medical devices available on the market, requiring the implementation of regulatory reforms to ensure patient safety.

Role of the Therapeutic Goods Administration (TGA)

Software based medical devices are medical devices that incorporate software or are software, including software as a medical device (SaMD), or software that relies on particular hardware to function as intended, and are regulated in Australia by the Therapeutic Goods Administration (TGA). Software (including mobile apps) is a medical device when the manufacturer (or developer) intends the product to be used for diagnosis, prevention, monitoring, treatment, alleviation of disease, injury or disability. Specific medical device regulation including for software is established through the *Therapeutic Goods (Medical Device) Regulations, 2002*.

In 2021, the TGA refined and clarified the regulatory requirements for software and depending on the intended purpose, a particular product could be:

- Software as a medical device (SaMD) – regulated by the TGA; or
- SaMD that is “carved out” from TGA regulation if the device presents a low risk to safety or if alternative oversight schemes are in place; or
- Consumer health software – not regulated by the TGA

Compliance with the regulatory requirements is neither optional nor voluntary.

Regulatory requirements

Software that is regulated by the TGA includes:

- ✓ Digital – software on any computing platforms (computers, tablets, smartphones, browsers)
- ✓ Software that is part of a medical device – is regulated as part of that device
- ✓ Apps that control a medical device – are regulated as an accessory or a device
- ✓ Apps that rely on medical device hardware in addition to a general computing platform (eg: sensors) – are part of a medical device.

Regulatory requirements are technology-agnostic for software-based medical devices and apply regardless of whether the product incorporates components like AI, large language models (LLMs) such as ChatGPT, other chatbots, cloud, mobile apps or other technologies – the regulations apply to these products when they are intended for medical purposes.

Companies who wish to supply a medical device in Australia must apply to the TGA to have their device included in the Australian Register of Therapeutic Goods (ARTG),

PO Box 100 Woden ACT 2606 ABN 40 939 406 804
Phone: 1800 020 653 or 02 6289 4124 Fax: 02 6203 1605
Email: Digital.Devices@tga.gov.au https://www.tga.gov.au

unless the device is exempt or excluded from that requirement. Any company who has an approved device included in the ARTG also has post-market reporting obligations to the TGA.

The level of scrutiny the TGA applies to a medical device before it can be included in the ARTG and made available to Australian consumers depends on its risk classification and the level of risk posed to a patient.

Devices with a higher risk classification (for example, software that makes a diagnosis for, provides information about, or recommends treatment options for a patient with a serious disease) must have very detailed evidence available to demonstrate they are safe and fit for their intended purpose. Less detailed evidence is required for devices with a lower classification (because they present a lower risk of harm). The manufacturers of all medical devices must hold evidence of compliance with the Essential Principles for safety, quality and performance, and where relevant undergo third-party certification of their manufacturing processes and their technical files.

Some examples of software, software devices and apps that diagnose and monitor illness, states of health or vital physiological processes regulated by the TGA include:

- Class I: phone apps for in-home monitoring of long-sightedness or for screening shingles recovery using uploaded images
- Class IIa: physiological monitoring software for patients not in immediate danger – respiration, heart rate, ECG, blood gases, blood pressure monitoring, body temperatures, EEG
- Class IIb: sleep apnoea monitoring that alerts a carer of life-threatening episodes
- Class III: app that analyses images of moles uploaded by a user to screen for malignant melanoma, without further input from a health care provider.

In addition to software incorporated in a medical device or SaMD, the TGA regulates software when it is included in an in vitro diagnostic (IVD) device.

Guidance regarding the regulatory requirements for SaMD can be found on the TGA's website and includes information about data management (privacy, collection, use), cybersecurity, algorithm and model description and validation, bias, integration with other data or devices/systems.

What is "carved out"?

Some low-risk products have been excluded from medical device regulation and therefore are not subject to any TGA regulatory requirements and do not need to be included in the ARTG. Examples include:

- Consumer health products – health prevention and management devices that do not provide specific treatment suggestions (eg: consumer products for monitoring heart rate or rhythm solely for general wellness or fitness purposes)
- Enabling technology – for telehealth, remote diagnosis, healthcare or dispensing
- Digitisation – simple dose calculators and electronic patient records
- Analytics – population based
- Laboratory information management systems
- Some aspects of clinical decision support software – eg: if they are not intended to replace health professional judgement in making a diagnosis or treatment decision.

Regulatory Guidance

Guidance published by the TGA provides useful information for companies seeking regulatory approval. There are general requirements and specific requirements depending on the risk classification and type of product. The guidance documents set out what must be included in technical files and evidence. Including use of real world evidence. This applies whether the software development methodology is agile (or a variant of agile) or other methodology.

When assessing a product, the TGA considers the following:

- Software architecture and design, physical and logical.
- Validation artefacts– overall test strategy and approach, test cases, requirements traceability matrix, test data, test results and defect rates
- Defect management process
- Human factors – showing how usability and accessibility have been incorporated into the design and take account of the needs of the users as general population who are not technically or medically trained
- Cybersecurity risks and how they have been addressed
- Data privacy - how it has been managed as it relates to patient safety and Australian privacy and data protection law.
- Clear instructions for use on how to use the device with accuracy and metrics disclosed

In addition to general software requirements, for software that uses AI or machine learning (ML), the manufacturer is required to show evidence that is sufficiently transparent to enable evaluation of safety and efficacy of the product.

TGA guidance has been developed with input from software industry and other relevant stakeholders. It includes flowcharts and examples to demonstrate the regulatory requirements and rules. An on-line classification tool available on the TGA website is available to assist companies to determine if their software is regulated by the TGA.

The following guidance documents are published on the TGA website at Regulation of software based medical devices.

- Is my software regulated?
- How the TGA regulates software based medical devices
- Regulatory changes for software based medical devices
- Examples of regulated and unregulated software (excluded) software based medical devices
- Clinical decision support software
- Exemption for certain clinical decision support software - Guidance on the Exemption Criteria
- Real world evidence (RWE) and patient reported outcomes (PROs)
- Artificial Intelligence Chat, Text, and Language
- Medical device cyber security guidance for industry
- Software as in vitro diagnostic medical devices (IVDs)

International alignment

Where possible the TGA contributes to and aligns with, guidance developed by the International Medical Device Regulators Forum - International Medical Device Regulators Forum (IMDRF)

IMDRF comprises medical device regulators from around the world who develop guidance to accelerate harmonisation of regulation. A SaMD Working Group is currently updating guidance on AI and TGA participates along with regulators from the USA, Canada, EU, UK, Brazil, China, Korea, Singapore.

APPENDIX A

Summary of classification rules for software based medical devices

		Diagnosing and/or recommending treatment or intervention for a disease or condition	
		Provides information to an individual	**Provides information to a health professional**
Risk to individual or public health	**Death/severe deterioration/high public health risk**	Class III	Class IIb
	Serious disease or condition/otherwise harmful/moderate public health risk	Class IIb	Class IIa
	Any other case	Class IIa	Class I
		Screening and/or specifying a treatment or intervention for a disease or condition	
	Death/severe deterioration/high public health risk	Class III	
	Serious disease or condition/otherwise harmful/moderate public health risk	Class IIb	
	Any other case	Class IIa	
		Monitoring the state/progression of a disease or condition	
	Immediate danger to a person/high public health risk	Class IIb	
	Other danger to a person or another/moderate public health risk	Class IIa	
	Any other case	Class I	
		For providing therapy through provision of information	
	May result in death/ severe deterioration	Class III	
	May cause serious harm	Class IIb	
	May cause harm	Class IIa	
	Any other case	Class I	

APPENDIX B

"Carve out" from regulation – examples of what's in and what's out

	Carved out - examples	Remaining regulated - examples
Consumer health life-cycle prevention, management and follow up		
(a) SaMD [not hardware] intended for self-management of an existing disease or condition that is **not serious** (without providing specific treatment or treatment suggestions)	The information is intended to be shared with a healthcare provider as part of a pre-diabetes management plan An app or wearable that monitors sleep and movement to assess and report on quality and quantity of sleep.	Software tool to organise and track a person's health information **and** gives a diagnosis for diabetes Software that monitors sleep and predicts risk of sleep apnoea.
(b) Consumer health and wellness products (may be software or a combination of non-invasive hardware and software), excludes serious conditions	A wearable that allows the wearer to track their heart rate for fitness An app on a smartphone that measures a physiological function such as oxygen saturation and makes no claims about serious diseases or conditions. An app that records and tracks physiological measurements such as blood pressure, blood test results as part of a personal health record.	A wearable that analyses the wearer's cardiac rhythm for the purpose of screening for a serious heart condition (which may include atrial fibrillation, heart attack risk among others) - the data collection component (the sensor) and the software are regulated. An app on a phone or tablet that analyses blood pressure and diagnoses hypertension. An app that uses the microphone to analyse sounds for the purpose of monitoring or diagnosing asthma. An app that analyses temperature, movement or oxygen saturation to diagnose COVID risk.
(c) Behavioural change or coaching software for improving general health parameters (for example weight, exercise, blood pressure, salt intake).	A 'sun smart' app that gives user alerts for UV protection to minimise skin cancer risk A consumer cognitive behavioural therapy (CBT) app	

	Carved out - examples	Remaining regulated - examples
(d) PROMs (patient recorded outcome measures) and patient surveys (including those that form part of an electronic health record)	An app that digitises an established PROM questionnaire (e.g., to assess the quality of life of a patient undergoing cancer treatment), similar to a paper-based version.	
(e) Digital mental health tools	Software that replicates paper-based mental health assessments in electronic format. The information must be from authoritative medical sources, as recognised by the relevant field or discipline, and must be cited in the software. The results can be independently reviewed by a health professional.	A patient questionnaire app that analyses the responses using a novel, unpublished algorithm to predict the risk of depression or an anxiety disorder. The software provides a diagnostic output that the health professional would otherwise not have access to.
Enabling technology for telehealth, remote diagnosis, health care facility management		
(a) Communication software that enables telehealth consultations or supports [a clinician in making] remote diagnosis	Video conference with a medical practitioner, with a waiting room facility. Communication of information, for example, non-urgent test results.	Software that records and communicates readings from a patient monitor to allow the patient's condition to be monitored from a remote location. The software generates real time feedback based on measured signals and generates alerts if signals are outside an established range.
(b) Software intended to administer or manage health processes or facilities, rather than patient clinical use cases	Processing of financial records, claims, billing, appointment schedules, business analytics, admissions, practice and inventory management, utilisation, cost effectiveness, health benefit eligibility, population health management, and workflow.	

	Carved out - examples	Remaining regulated - examples
(c) Systems that are intended only to store patient images	Medical image storage and retrieval device, or medical image communication between devices.	Software that records an image directly from an MRI scanner. Software that analyses an MRI scan to automatically identify potential tumours.
(d) Software intended to be used by health professionals to provide alerts or additional information. The health professional can exercise their own judgement in determining whether to action the alert or information.	Pharmacy dispensing systems and prescribing software used by GPs, also Clinical decision support software – these are not intended to be used by laypeople and are not themselves acting as a de facto decision maker.	
(e) Software embedded in delivery of health services	Clinical workflow and support – including display medical information about a patient or peer-reviewed clinical studies and clinical-practice guidelines	
(f) Middleware that does not recommend a diagnosis or treatment decision and that do not message IVD instruments or other medical devices	Laboratory software that facilitates the electronic transfer of data between medical devices. The software does not control a medical device, or analyse the data transferred in any way.	Software that operates an IVD instrument. Software that combines IVD results to calculate and report a result for clinical purposes. For example, software that interprets results from a first trimester screening assessment for foetal risk of trisomy 21.
Digitisation of paper based or other published clinical rules or data		
(a) Simple calculators	Software that calculates drug dosing based on a published clinical standard. The user inputs the parameters (e.g., age, gender, weight) and can independently review the calculation.	An automated insulin bolus calculator that controls the dose delivered by an insulin pump

	Carved out - examples	Remaining regulated - examples
(b) Electronic Patient Records (EMRs) and Electronic Health Records (EHRs)	Receive, collect, store, manage, display, output, and distribute data, within or between healthcare facilities, to manage patient clinical data. It typically enables healthcare providers to review and update patient medical records, place orders (e.g., for medications, procedures, tests), and view data from many specialties.	A module integrated into an EMR that directly records readings from a patient monitor to allow the patient's condition to be monitored remotely. Apps that connect to an EHR and analyse patient data to screen for high risk of a specific condition
Population based analytics		
Data analytics that are class or group based rather than individual patient based	Analysis on a population who are asked via email reminders to report via a website on fever, cough, days off work, and vaccination status. The results are used to generate population statistics and track infections, which may be of use in studying and controlling epidemics. The information is not used to inform interventions for any of the individuals involved.	Analytic app using aggregated population data for a particular condition to make inferences about the most appropriate treatment options for an individual Analytic app that extracts groups from a EHR database and combines data with other sources to identify high risk of a disease that leads to action for an individual
Clinical decision support systems		
A clinical decision support system is exempt if: • it is not intended to acquire, process, or analyse a medical image or a signal from a hardware medical device or an in vitro diagnostic device, and • it is intended for the purpose of displaying, analysing, or printing medical information about a patient or other medical information (such as peer-reviewed clinical studies	Software that displays information about a patient and other medical information (such as peer-reviewed clinical studies and clinical-practice guidelines). The software takes this information and presents a diagnosis and relevant treatment recommendation, along with the rationale for this, to a health professional for the purposes of assisting them in determining a diagnosis	Software that obtains data from a closed-loop blood glucose monitor, analyses the data to provide early diagnosis of a diabetic emergency such as serious hypoglycaemia. When a patient experiences a diabetic emergency, the software will alert the treating health professional and use the results of the analysis to make treatment decisions based on the patient's unique health profile.

	Carved out - examples	Remaining regulated - examples
and clinical practice guidelines), and • it is intended only for the purpose of supporting or providing recommendations to a health professional about prevention, diagnosis, or treatment of a disease or condition, and • it is not intended to replace the clinical judgement of a health professional to make a clinical diagnosis or treatment decision regarding an individual patient.	or treatment for their patient.	
Laboratory information management systems (LIMS) and Laboratory information systems (LIS)		
	Software that automates workflows, integrates instruments, manages samples, reports results of assays - but does not recommend a diagnosis or treatment.	A LIMS software module that performs a manipulation on the data that affects the interpretation of results or generates new diagnostic data/information.

APPENDIX C

Examples of medical devices incorporating AI regulated by the TGA

Medical device/SaMD incorporating ML/AI	Risk classification as per Australian legislation
Diagnostic digital imaging system workstation application software Software as a Medical device incorporating AI for full-field digital mammography (FFDM) and digital breast tomosynthesis (DBT). Software is a computer-assisted detection and diagnosis (CAD) Artificial Intelligence (AI) software device intended to be used concurrently by physicians while reading FFDM and DBT exams from compatible FFDM and DBT systems. The system detects soft tissue densities (masses, architectural distortions and asymmetries) and calcifications in the FFDM images and DBT slices. The detections and Certainty of Finding and Case Scores assist interpreting physicians in identifying soft tissue densities and calcifications that may be confirmed or dismissed by the interpreting physician.	Class IIb
Radiology image processing application software A radiological computer aided triage and notification software indicated for use in the analysis of Chest and Thoraco-abdominal CT angiography. The device is intended to assist hospital networks and trained radiologists in workflow triage by flagging and communicating suspected positive findings of Chest CT angiography for Pulmonary Embolism (PE) and Chest or Thoraco-abdominal CT angiography for Aortic Dissection (AD). The device uses an artificial intelligence algorithm to analyse images and highlight cases with detected PE or AD on a standalone Web application in parallel to the ongoing standard of care image interpretation. The user is presented with notifications for cases with suspected PE or AD findings.	Class IIb
Retinal optical coherence tomography interpretive software Software including artificial intelligence that assists medical personnel by inputting the patient's fundus image and indicating the suspected symptoms of eye diseases (cataract, glaucoma, retinal disease) and predicting cardiovascular risk.	Class IIa
Software or app for use with in vitro diagnostic (IVD) tests that are intended to analyse and enable the interpretation of the test result Software (incorporating ML/AI) that analyses/interprets results from COVID-19 rapid antigen self-tests or Software that allows a user to combine their test result with other symptoms to provide an indication or likelihood of having COVID-19.	Class 3 IVD
Gene sequencing platforms for diagnostic use Software intended to analyse next-gene sequencing data and molecular testing data of tumour tissue using ML/AI approaches to recommend treatment options personalised to a patient's tumour characteristics.	Class 3 IVD