CIVIL LIABILITY FOR AI CLINICAL TOOLS IN WESTERN AUSTRALIA: A CRITICAL OVERVIEW

KUEN YEI CHIN, MEREDITH BLAKE AND MARCO RIZZI*

Over the past 50 years, the use of AI clinical tools has experienced exponential growth, particularly in the realm of clinical judgements and assessment. While the use of AI clinical tools has been heralded as bringing the benefits of efficiency and accuracy to clinical practice, questions about how best to regulate the use of these tools persist, particularly in relation to the potential for patients experiencing harm. This article outlines the relevant legal frameworks in Western Australia, with a particular focus on civil liability regimes, analysing how these can adapt to the challenges of the use of AI in clinical settings. The analysis identifies possible gaps in the liability frameworks and provides an overview of select alternative liability models developed by legal scholarship, assessing their potential to address these gaps. The article concludes by highlighting the need for research into how health professionals navigate the risks of AI clinical tools, and the practical challenges they present, to inform the development of relevant liability frameworks in WA.

* Kuen Yei Chin is a PhD candidate in the Centre for Health Law and Policy at UWA Law School, The University of Western Australia; Professor Meredith Blake and Associate Professor Marco Rizzi are Co-Directors of the Centre for Health Law and Policy at UWA Law School, The University of Western Australia. The authors wish to thank the anonymous peer-reviewers and the student reviewer for their helpful comments, as well as the UWA Law Review team for their wonderful editorial work — with special thanks to Chantelle Naguib.

I INTRODUCTION

In 1955, John McCarthy introduced the term Artificial Intelligence ('AI'), a descriptor for the domain of science and engineering responsible for the development of intelligent machines and computer programs with the capacity to mimic human cognitive abilities, including thinking, acting, and decision-making.[1] There are currently several definitions of AI, each reflecting the diverse applications and interpretations of the technology.[2] For example, some definitions emphasise AI's ability to mimic human cognition, while others focus on its capacity to learn and improve autonomously.[3] For present purposes, however, the broad definition provided by the ISO/IEC 22989:2022 standard, as referenced in the Western Australian ('WA') Government AI Policy, will be adopted. This refers to AI as:

> An engineered system that generates predictive outputs such as content, forecasts, recommendations, or decisions for a given set of human-defined objectives or parameters without explicit programming. AI systems are designed to operate with varying levels of automation.[4]

The above definition more than hints at the complex issues which arise when deploying AI tools in the context of healthcare. The ethical, legal, and social quandaries which arise in connection with the increasing role of this technology in the provision of patient care cannot and should not be minimised. Particularly

---

[1] John McCarthy, 'What is Artificial Intelligence? Basic Questions', *Stanford Formal Reasoning Group* (Web Page, 12 November 2007) <http://www-formal.stanford.edu/jmc/whatisai/node1.html>.

[2] See generally Haroon Sheikh, Corien Prins and Erik Schrijvers, 'Artificial Intelligence: Definition and Background' in *Mission AI: The New System Technology* (Springer, 2023) 15, 15–16.

[3] Mara Graziani et al, 'A Global Taxonomy of Interpretable AI: Unifying the Terminology for the Technical and Social Sciences' *Artificial Intelligence Review* 56 (2023) 3473, 3474.

[4] Department of the Premier and Cabinet Office of Digital Government, *Western Australian Government Artificial Intelligence Policy* (Policy Paper, March 2024) 5.

difficult questions arise when considering the allocation of legal responsibility for patient harm, which occurs in association with the use of AI tools. These questions include, but are not limited to: who should bear the ultimate liability for any harm inflicted; and what is the most appropriate mechanism of liability allocation in the event of a patient being harmed?

These questions continue to form the focus of wide-ranging scholarship, from purely doctrinal analysis of liability to more high-level theoretical discussions, and, to a lesser extent, contextual and socio-legal aspects of the issue.[5] A full review of existing literature is beyond the scope of this article, but it is critical for legal scholarship to interrogate how liability models function in real-world settings, governed by specific and locally variable legal frameworks. This article provides a jurisdiction-specific analysis of the strengths and limitations of available liability models for health professionals who use AI clinical tools in diagnosing, treating, and advising patients, as well as a critical overview of possible alternatives.

The focus on the civil liability landscape of WA, as applied to the use of this technology in a local setting, facilitates the examination of these practical challenges. It questions whether the current legal frameworks are sufficiently nuanced and adapted to the use of AI clinical tools in the event that a patient suffers harm. The analysis underscores the need for empirical research to explore the experiences of health professionals in using AI clinical tools, particularly in connection with the identified questions of allocation of liability.

To provide clarification around terminology, the following table provides definitions of key terms used in this article.

[5] See Helen Smith and Kit Fotheringham, 'Artificial Intelligence in Clinical Decision-Making: Rethinking Liability' (2020) 20(2) *Medical Law International* 131; Megan Prictor, 'Where Does Responsibility Lie? Analysing Legal and Regulatory Responses to Flawed Clinical Decision Support Systems when Patients Suffer Harm' (2022) 31(1) *Medical Law Review* 1; Paul Nolan and Rita Matulionyte, 'Artificial Intelligence in Medicine: Issues When Determining Negligence' (2023) 30(3) *Journal of Law and Medicine* 593; Søren Holm, Catherine Stanton and Benjamin Bartlett, 'A New Argument for No-Fault Compensation in Health Care: The Introduction of Artificial Intelligence Systems' (2021) 29(3) *Health Care Analysis* 171; Frank Griffin, 'Artificial Intelligence and Liability in Health Care' (2021) 31 *Health Matrix* 66; Dane Bottomley and Donrich Thaldar, 'Liability for Harm Caused by AI in Healthcare: An Overview of the Core Legal Concepts' (2023) 14 *Frontier Pharmacology* 1.

Table 1: Definitions

Term	*Definition*
Algorithms	Refers to a 'set of instructions that is designed to accomplish a task. Algorithms usually take one or more inputs, run them systematically through a series of steps, and provide one or more outputs.'[6]
Neural networks	Refers to algorithms where the input data connects to many 'hidden layers' before ultimately connecting to the output layer.[7]
Black Box	Refers to 'systems that hide their internal logic from the user'. They have been described as 'ubiquitous decision systems'.[8]
Machine Learning ('ML')	Refers to a subfield of artificial intelligence that involves the development of algorithms that allow computers to learn from data.[9] It uses statistical models and algorithms to analyse and identify patterns in data, and make predictions or decisions based on that data.[10]
Human-in-the-loop ('HITL')	'The term human-in-the-loop generally refers to the need for human interaction, intervention, and judgment to control feedback and change the outcome of a process. It is a practice that is being increasingly emphasised in ML and generative AI.'[11]
Clinical practice	Refers to the practical execution of skills and tasks in the provision of healthcare to patients by healthcare professionals.
Health professional	Refers to professionals who are delivering healthcare services to patients in the course of providing treatment, in this context particularly diagnostic services.
AI clinical tools	In this article, this term refers to clinical tools that use AI technology. Clinical tools in this context are tools used by health professionals to diagnose patients' conditions, predict patients' risk of disease onset, progression, or complications, and suggest treatment decisions.[12]

[6] *National Library of Medicine: Data Glossary* (online at 16 November 2025) 'Algorithm' <https://www.nnlm.gov/guides/data-glossary/algorithm>.
[7] See Jenna Burrell, 'How the Machine "Thinks": Understanding Opacity in Machine Learning Algorithms' (2016) 3(1) *Big Data & Society* 1, 5-6.
[8] Riccardo Guidotti et al, 'A Survey of Methods for Explaining Black Box Models' (2018) 51(5) *ACM Computing Surveys* 1, 1.
[9] See Ally S Nyamawe et al, 'Fundamentals of Machine Learning' in *Practical Machine Learning: A Beginner's Guide with Ethical Insights* (CRC Press LLC, 2025) 16, 16.
[10] Ibid.
[11] Xiao-Li Meng, 'Data Science and Engineering with Human in the Loop, Behind the Loop, and Above the Loop' (2023) 5(2) *Harvard Data Science Review* 1, 2.
[12] See Khaled Ouanes and Nasren Farhah, 'Effectiveness of Artificial Intelligence (AI) in Clinical Decision Support Systems and Care Delivery' (2024) 48(74) *Journal of Medical Systems* 1, 1–2.

II RECALIBRATING THE RISKS IN THE AI SPACE

As noted above, this article seeks to assess the legal liability frameworks in WA by reference to health professionals who interact with AI clinical tools. In traditional clinical settings, health professionals need to evaluate the benefits and risks of a particular type of treatment, procedure, or medication, as part of their professional and legal duties to avoid foreseeable harm to the patient.[13] This involves an exercise of professional autonomy whereby the health professional assumes responsibility in relation to the patient's care, including making decisions about the patient's diagnosis and treatment options.[14] Applying this approach to the use of AI clinical tools for diagnosis and associated treatment options, health professionals should similarly assess the risks and benefits of these tools, given the possibility that the patient may suffer harm should the risk materialise. However, the complexity of the processes associated with the application of AI clinical tools adds a dimension to the traditional approach to determining liability, raising questions around fault, causation of harm, and ultimately the fair allocation of liability for patient harm.

The answers to these questions are inextricably tied to the status allocated to the AI clinical tool in the health professionals' clinical responsibilities. The scholarship reveals several approaches to this issue. One approach is to regard the use of the AI clinical tool as merely an adjunct or clinical decision support tool, implying that it is the end user of the tool, the health professional, who should bear responsibility for any harm suffered by the patient.[15] Megan Prictor argues that this is the effect of the current Australian regulatory frameworks. Her work has analysed potential forms of liability in negligence, contract, and under statutory consumer law in Australia relating to unintended harm associated with the use of AI clinical tools in healthcare.[16] Her conclusion is that the contours of legal risk still centre upon the clinician's duty to exercise decisional autonomy and to intercept flawed recommendations generated by algorithmic errors within clinical decision support systems.[17]

---

[13] The requirement of foreseeability of harm is enshrined in the general legislative provision governing breach of duty: *Civil Liability Act 2002* (WA) s 5B(1)(a) ('*CLA 2002*').
[14] See Joseph Lee, 'The Standard of Medical Care under the Australian Civil Liability Acts: Ten Years On' (2014) 22 *Journal of Law and Medicine* 335, 336–9.
[15] See Prictor (n 5) 2, 7–8.
[16] Ibid 7–9.
[17] Ibid 14–15.

Another approach canvassed in the literature is to regard the relationship between the patient and the AI system as direct, since the patient's data is processed by the system itself, even if a clinician mediates the interaction.[18] This perspective resonates with the broader discourse on algorithmic contracts, which references the reliance on algorithmic decision-making and challenges the traditional place of human autonomy and responsibility. Natalie Skead and Marco Rizzi, for example, discuss the influential effect of algorithmic decision-making on contractual relationships in ways that erode free will or informed consent, raising concerns akin to undue influence.[19] In the clinical context, this would manifest through AI clinical tools shaping medical decision-making, either by subtly nudging health professionals toward particular recommendations, or by introducing a Black Box effect that obscures independent judgement. This may warrant reconsideration of liability allocation beyond the individual health professional to include AI suppliers or manufacturers.[20]

These different approaches to the essential role of AI tools in the decision-making process are implicitly bound up with risk determinants — both the degree of risk and the type of risk (the nature of the harm which may be experienced by patients). As such, it is important to outline the particular risks of harm associated with the use of AI in clinical practice before assessing the capability of current civil liability regimes available to patients in WA in dealing with these risks.

III WHAT RISKS OF WHAT HARM?

Any liability regime requires the identification of a harm of some kind as the necessary premise for the regime to be engaged. Simply put, without harm, there can be no liability.[21] Defining harm is not as easy as it may intuitively seem. What constitutes cognisable harm at law has been the subject of a long-lasting doctrinal debate in the Anglosphere.[22] If one were to cast a wider net to capture the

[18] Smith and Fotheringham (n 5) 137.
[19] Natalie Skead and Marco Rizzi, 'Algorithmic Contracts and the Equitable Doctrine of Undue Influence: Adapting Old Rules to a New Legal Landscape' (2020) 14(3) *Journal of Equity* 301.
[20] Ibid 314–16.
[21] See generally, eg, Jane Stapleton, *Three Essays on Torts* (Oxford University Press, 2021); Nicky Priaulx, 'Reproducing the Properties of Harms that Matter: The Normative Life of the Damage Concept in Negligence' (2017) 5 *Journal of Medical Law and Ethics* 17; Donal Nolan, 'Rights, Damage and Loss' (2017) 37(2) *Oxford Journal of Legal Studies* 255; Nicky Priaulx, 'Endgame: On Negligence and Reparation for Harm' in Janice Richardson and Erika Rackley (eds), *Feminist Perspectives on Tort Law* (Taylor & Francis Group, 2012) 36; Gemma Turton, *Evidential Uncertainty in Causation in Negligence* (Bloomsbury Publishing, 2016); Kumaralingam Amirthalingam, 'The Changing Face of the Gist of Negligence' in Jason Neyers et al (eds), *Emerging Issues in Tort Law* (Hart Publishing, 2007) 467.
[22] Examples of that debate can be found in the literature referenced at n 21.

European continental tradition, the results would be even more fragmented. Broadly speaking, liability regimes tend to be engaged either when the claimant has suffered a form of quantifiable loss (the dominant approach at common law), or when an identifiable right or legitimate interest has been violated (a common approach in the civil law tradition).[23] The emergence of AI is further complicating this debate, and Australian jurisdictions are not immune to the challenge. The *WA Government Artificial Intelligence Assurance Framework* identifies several potential risks of harm associated with AI technologies, noting that physical harm, psychological harm, unauthorised use of health or other sensitive information, incorrect advice or guidance, and inconvenience or delay are particularly relevant to the use of AI clinical tools.[24] Adverse health outcomes such as physical harm, for example, can result from misdiagnosis or inappropriate treatment recommendations associated with data biases embedded in AI clinical tools (discussed below).[25]

In considering how the use of AI clinical tools produces risks of these sorts of harms, it is essential to have regard to the operational premise underpinning these tools. Their function is intrinsically founded in the application of algorithms. AI algorithms in AI clinical tools are designed to enhance decision-making processes in clinical settings by analysing vast amounts of data to provide insights, predictions, and recommendations.[26] This involves using predictive analytics to identify correlations and patterns that contribute to disease development and progression.[27] This critical element of predictive analysis is inseparable from the flaws that are more generally associated with

---

[23] See, eg, James Gordley, 'The Architecture of the Common and Civil Law of Torts: An Historical Survey' in Mauro Bussani and Anthony J Sebok, *Comparative Tort Law: Global Perspectives* (Edward Elgar, 2021) 160.

[24] Department of the Premier and Cabinet Office of Digital Government, *WA Government Artificial Intelligence Assurance Framework* (October 2025) <https://www.wa.gov.au/government/publications/wa-government-artificial-intelligence-policy-and-assurance-framework>. The main risk factors for individuals or communities encompass physical harm, psychological harm, environmental harm or harms to the broader community, unauthorised use of health or other sensitive information, impact on right, privilege or entitlement, unidentified identification or misidentification of an individual, misapplication of a fine or penalty, other financial or commercial impact, incorrect advice or guidance, and inconvenience or delay.

[25] For a general discussion see Cakesha Hardin, 'Data-Driven Technology Medical Malpractice: A Narrative Review on the Legal Implications in Clinical Settings' in Darrell Norman Burrell (ed), *Organizational Readiness and Research Security, Management, and Decision Making* (IGI Global, 2025) 373.

[26] Vikas Kumar Kharbas, Y R Meena and Praveen Kumar Thakur, 'Enabling Clinical Decision-Making in Smart Healthcare through AI/ML Algorithms' (Conference Paper, Proceedings of the 15th International Conference on Computing Communication and Network Technologies (ICCCNT), 24–28 June 2024) 2.

[27] Ibid 3.

system-generated decisions.[28] The conclusions offered by predictive analytic software make it nearly impossible to analyse and understand the basis for a diagnosis or treatment.[29] This inherent complexity means that the flaws are deeply embedded within the system's operation as the logic and inputs may not be fully explained to the user.[30]

Before diving into an analysis of how civil liability models in WA can (or could) respond to cases where harm may ensue for patients, the following paragraphs succinctly unpack algorithmic flaws and their related risks as explored in the literature, grouping them around three themes: bias, error, and opacity.

A *Bias*

Bias in AI is a systematic error in decision-making arising from data bias, algorithm bias, or user bias.[31] Bias here refers to the situation where people with particular characteristics such as race, gender, or socioeconomic status are disproportionately affected in the provision of healthcare. The bias underpinning the manifestation of these effects is generated by AI systems which create misleading results that compromise the accuracy of decisions in healthcare contexts.[32] For example, a study by Ziad Obermeyer et al, examining bias in a widely used United States healthcare risk-prediction algorithm, revealed that Black patients were more likely to present with higher needs on a range of common diseases such as severity of diabetes, high blood pressure, renal failure, cholesterol, and anaemia than White patients.[33] It was demonstrated that this occurred because the algorithm used past healthcare spending as a proxy for health status. This measure is influenced by structural inequities in access to care, resulting in fewer White patients being flagged for high-risk care management and revealing systemic racial bias in the tool's design.[34]

---

[28] Ibid.
[29] Christophe Lazaro and Marco Rizzi, 'Predictive Analytics and Governance: A New Sociotechnical Imaginary for Uncertain Futures' (2023) 19(1) *International Journal of Law in Context* 70, 85.
[30] Ibid 73.
[31] Emilio Ferrara, 'Fairness and Bias in Artificial Intelligence: A Brief Survey of Sources, Impacts, and Mitigation Strategies' (2024) 6(3) *Sci* 1, 2.
[32] Ninareh Mehrabi et al, 'A Survey on Bias and Fairness in Machine Learning' (2021) 54(6) *ACM Computing Surveys* 115, 3–9. In the 'Taxonomy of Bias' section, where they provide detailed explanations of data bias, algorithm bias, and user bias, the descriptions often include examples and implications for fairness, specifically in real-world applications.
[33] Ziad Obermeyer et al, 'Dissecting Racial Bias in an Algorithm Used to Manage the Health of Populations' (2019) 366(6464) *Science* 447, 450.
[34] Ibid 453.

Data bias can stem from limitations in input data, such as imbalanced data sets, which fail to accurately represent real-world populations.[35] Research has established that one cause of data bias is a distributional shift, which occurs when ML systems encounter new contexts or data that differ from their training.[36] This shift is observed when systems trained on specific devices or populations perform poorly in broader applications.[37] Examples of distributional shifts can be identified in the clinical use of ML algorithms. One study in relation to Epic Sepsis — a widely used proprietary ML tool designed to predict sepsis risk in hospitalised patients — found that it performed significantly worse in real-world settings compared to its reported accuracy.[38] The model struggled to identify many sepsis cases when rolled out in clinical settings, highlighting the challenges of deploying AI-driven clinical decision support tools in a dynamic and evolving healthcare environment.[39] Another study reported bias in skin disease triage tools which classify dermatological images to prioritise which patients require urgent specialist evaluation.[40] The study revealed that these tools performed less accurately on individuals with darker skin tones and in detecting uncommon diseases like malignant skin cancer.[41]

A separate but related type of bias is algorithmic bias, which occurs as a result of biases inherent to the ML process born out of assumptions embedded in the system's design.[42] The tool which the Obermeyer et al study examined is a good example of algorithmic bias.[43] The nature of the ML process, particularly when based on artificial neural networks, poses a significant obstacle to the detection of bias (and ensuing errors), as its predictions are intrinsically opaque.[44]

---

[35] Robert Challen et al, 'Artificial Intelligence, Bias and Clinical Safety' (2019) 28(3) *BMJ Quality & Safety* 231, 233.
[36] Ibid 232–3.
[37] Ibid; Larry Han, 'Addressing Distribution Shift for Robust and Trustworthy Prediction and Causal Inference in Clinical AI Settings' (2025) 8(6) *JAMA Network Open* 1, 1.
[38] Andrew Wong, Erkin Otles and John Donnelly, 'External Validation of a Widely Implemented Proprietary Sepsis Prediction Model in Hospitalised Patients' (2021) 181(8) *JAMA Intern Med* 1065, 1068–9 .
[39] See ibid 1069.
[40] See Andre Esteva et al, 'Dermatologist-Level Classification of Skin Cancer with Deep Neural Networks' (2017) 542(7639) *Nature* 115.
[41] Ibid 118.
[42] Ferrara (n 31) 2.
[43] Obermeyer et al (n 33) 453.
[44] Challen et al (n 35) 233; Adam Andreotta, Nin Kirkham and Marco Rizzi, 'AI, Big Data, and the Future of Consent' (2022) 37(4) *AI & Society* 1715, 1719.

So far, we have identified bias-affected inputs, which affect the front end of the generative process. Other forms of bias are based in the users of the tools, which arguably present more of a remedial challenge. General user bias arises when individuals interacting with AI systems introduce conscious or unconscious biases, which can influence the system's behaviour as it iteratively learns from experience.[45] Hierarchical dynamics in hospitals, disparities in access to care, and cultural norms around diagnosis and treatment all shape how users engage with AI.[46] For example, if a hierarchical dynamic in a clinical setting leads to a culture of deference to senior staff involved in the recognition of diagnostic patterns, then the AI tool, depending on its nature, may be prompted to replicate the pattern and embed whatever bias underpins it. This example reflects a view that user bias is closely linked to the sociotechnical context in which an AI tool operates, that is, the interaction between technical systems and the social, cultural, and institutional environments they are embedded in.[47] It casts general user bias as a complex and multilayered problem, emphasising the dangers of approaching the design of AI tools as a purely technical challenge. Confounding what is a sociotechnical problem as presenting a simple technological issue therefore risks overlooking the deep-seated origins of bias and the production of solutions which inadvertently sustain or worsen inequality.[48]

The forms of bias considered so far implicate, to varying degrees, the manufacturer of the AI tool. A further form of bias, automation bias, centres on the back end of the process, where the application of the tool to a dataset has resulted in specific conclusions around diagnoses.[49] It therefore arises from the use of these outputs in clinical practice. It is distinct from general user bias in that it does not stem from the user projecting their own social or cognitive biases

---

[45] Ferrara (n 31) 2.

[46] Kate Crawford and Ryan Calo, 'There is a Blind Spot in AI Research' (2016) 538(7625) *Nature* 311, 312–13. As Crawford and Calo note, the performance of medical AI cannot be separated from the institutional and cultural environment in which it operates. For example, hospital guidelines and insurance practices shaped the data that led an AI system to misclassify asthma patients as low risk for pneumonia complications.

[47] Andrew Selbst et al, 'Fairness and Abstraction in Sociotechnical Systems' (Conference Paper, FAT '19: Proceedings of the Conference on Fairness, Accountability, and Transparency, 29 January 2019). The authors argue that fairness and bias are not intrinsic properties of algorithms but of the broader sociotechnical systems, ie social, cultural, and institutional, within which these tools are embedded (at 59–60). The authors also emphasise that technical interventions often reshape existing practices and power dynamics, meaning that user bias is inseparable from the social contexts in which AI systems operate (at 62–5).

[48] Ibid 63–4.

[49] Kate Goddard, Abdul Roudsari and Jeremy Wyatt, 'Automation Bias: A Hidden Issue for Clinical Decision Support System Use' in EM Borycki et al (eds), *International Perspectives in Health Informatics* (IOS Press, 2011) 17, 18.

onto the system, but rather from the user placing excessive trust in AI outputs, often leading to the acceptance of incorrect recommendations generated by the AI clinical tool.[50] Such overreliance may take the form of a failure to act on an error because of a lack of alerts, or of unquestioningly following incorrect system advice.[51] While the absence of alerts or the presence of incorrect recommendations may be attributed to system errors, the failure of clinicians to interrogate AI outputs is indicative of a user placing inordinate trust in the system's guidance. This can be particularly problematic as the output may be incorrect or incomplete.[52] The implications of this form of bias are that, in the event of patient harm associated with the AI tool, the liability lens shifts to the health professional's interaction with the tool. It squarely raises the question of how the health professional's actions and inactions could feature in the determination of civil liability for that harm.

B *Errors in, or of, the AI Clinical Tool*

The identified sources of bias in the design and use of an AI tool need to be conceptually distinguished from errors in the AI tools themselves, such as where the tool contains flawed, outdated or inaccurate source information.[53] While automation bias, where users overly rely on AI outputs, can amplify the impact of these errors, it is not the same as an intrinsic error within the tool itself. This distinction is crucial for understanding and addressing the full range of challenges associated with AI clinical tools. Evans and Snead have identified several types of errors which can manifest in AI clinical tools, noting that both errors and bias can affect the reliability and safety of such tools.[54]

---

[50] Florian Kücking et al, 'Automation Bias in AI-Decision Support: Results from an Empirical Study' (2024) 30(317) *Studies in Health Technology and Informatics* 298, 303.
[51] Mark Sujan et al, 'Human Factors Challenges for the Safe Use of Artificial Intelligence in Patient Care' (2019) 26(1) *BMJ Health and Care Informatics* 1, 2.
[52] Joan Ash et al, 'Some Unintended Consequences of Clinical Decision Support Systems' (Conference Paper, AMIA Annual Symposium Proceedings 2007, 11 October 2007) 26, 27–8; Sarah Jabbour, David Fouhey and Stephanie Shepard, 'Measuring the Impact of AI in the Diagnosis of Hospitalised Patients: A Randomised Clinical Vignette Survey Study' (2023) 330(23) *JAMA Network* 2275, 2279–80; Yaëlle Bellahsen-Harrar et al, 'Exploring the Risks of Over-Reliance on AI in Diagnostic Pathology: What Lessons can be Learned to Support the Training of Young Pathologists?' (2025) 20(8) *PLoS One* 1, 8.
[53] Ash et al (n 52) 26–7.
[54] Harriet Evans and David Snead, 'Why Do Errors Arise in Artificial Intelligence Diagnostic Tools in Histopathology and How Can we Minimise Them?' (2024) 84(2) *Histopathology* 279, 281–3.

These errors are conceptually distinct from the distributional shift, which occurs when the model encounters data during deployment that differs from the data it was trained on.[55] Distributional shift, which may result in misleading outputs, is often associated with changes in population characteristics, clinical practices, or data collection procedures.[56] It can result from poor-quality data, but it is not synonymous with error. For example, low-quality data may be incomplete but still relatively representative, whereas erroneous data, such as wrong labelling of source information, can directly misinform model training and lead to systematic errors.[57]

Separate to those that are intrinsic to the AI tool, errors can also be practical manifestations of biases as described above. Lyell et al explain that errors can arise from the interactions between users and AI clinical tools. These errors are not limited to flaws in the data processed by AI tools but are also a product of how the tools are implemented, understood, and used.[58] For example, inadequate user training or the deployment of AI systems outside their intended scope can lead to erroneous outputs, particularly in high-stakes clinical environments.[59] This type of error can be the manifestation of automation bias. Lyell et al illustrate this with an example of an electronic prescribing system failing to alert a clinician about a potential drug interaction.[60] Because the clinician relied on the system's guidance, the interaction was missed, leading to a prescribing error.[61] As AI technology continues to advance, it is anticipated that new types of errors will emerge. Siontis et al have identified the possibility of AI systems generating outputs that appear plausible but are not grounded in the underlying data or clinical context, a phenomenon known in generative AI as 'hallucination'. This underscores the importance of implementing robust validation processes and ensuring that users are adequately trained to critically evaluate AI-generated outputs.

---

[55] Nicolas Acevedo et al, 'Definition and Detection of Distribution Shift', *arXiv* (Preprint, 23 May 2024) 1, 1 <https://arxiv.org/abs/2405.14186>.
[56] Samuel Finlayson et al, 'The Clinician and Dataset Shift in Artificial Intelligence' (2021) 385(3) *New England Journal of Medicine* 283, 284–6.
[57] Marzyeh Ghassemi, Luke Oakden-Raynerc and Andrew Beam, 'The False Hope of Current Approaches to Explainable Artificial Intelligence in Health Care' (2021) 3(11) *Lancet Digital Health* 745, 748.
[58] David Lyell et al, 'Automation Bias in Electronic Prescribing' (2017) 17(1) *BMC Medical Informatics and Decision Making* 28, 35.
[59] See Yang Gong et al, 'Challenges and Opportunities of Artificial Intelligence in CDSS and Patient Safety' in John Mantas et al (eds), *Digital Health and Informatics Innovations for Sustainable Health Care Systems* (IOS Press, 2024) 1250, 1251.
[60] Lyell et al (n 58) 35.
[61] Ibid 37.

C *Issues with Opacity, Explainability and Interpretability*

While it is well known that bias or error can occur, identifying with any precision the how and when behind these occurrences is a significant challenge. Scholars have identified complexity, opacity (often referred to as the Black Box phenomenon), and unpredictability as the key issues associated with the use of AI operations. These factors can make it difficult to trace errors or verify decision pathways, with resulting complications in establishing accountability for adverse outcomes.[62] The Black Box has been explained as 'a system or program that allows one to see the input data and output results but gives no view of the processes and workings in between'.[63] This description identifies a lack of operational transparency which leaves even developers unable to explain how specific conclusions are drawn.

The lack of accessibility of the decision-making process, from the input of data to the output of conclusions derived from the operation of clinical tools, creates profound challenges for traditional legal liability frameworks. The next part of the article illustrates these challenges by setting out the current WA liability frameworks relevant to the production and use of AI clinical tools, as an essential precursor to addressing the question of allocation of liability for harm associated with these tools. It aims to clarify whether liability could lie with the health professional, the healthcare facility, or the developers or manufacturers of the AI clinical tool. The broad field of therapeutic goods, of which AI clinical tools are a small subset, is a heavily regulated one.[64] As such the analysis will also briefly touch on the potential liability of regulators.

IV RELEVANT LEGAL FRAMEWORKS

The focus of this section is on responses to personal injury, and the analysis therefore centres on the civil law of negligence, which in Australia is organised at a state and territory level. In WA the general law of negligence, as well as medical negligence specifically, stems from a combination of common law and statute. In addition to medical negligence, this section will also consider the role of existing

---

[62] Nicholson Price II, 'Medical Malpractice and Black-Box Medicine' in Glenn Cohen et al (eds), *In Big Data, Health Law, and Bioethics* (Cambridge University Press, 2018) 295, 300.

[63] Frank Pasquale, *The Black Box Society: The Secret Algorithms that Control Money and Information* (Harvard University Press, 2015) 3.

[64] See Eimear Reynolds, 'Machine Learning-Integrated Medical Devices in Australia: Safety Defects and Regulation' (2024) 50(3) *Monash University Law Review* 496, 31–4; Marco Rizzi, Penny Gleeson and Jeannie Marie Paterson, 'Floors and Ceilings: Coordination, Coherence and Consistency in the Relationship Between Therapeutic Goods Regulation and Consumer Protection' (2024) 47(9) *Melbourne University Law Review* 466, 471–6.

consumer protection laws in providing compensation for affected patients and consumers.

A *Civil Liability Act 2002 (WA)*

Proving liability in negligence requires a plaintiff to establish that the defendant owes them a duty of care, that the defendant breached that duty, and that the breach is causative of recognisable harm.[65] In WA, as in other Australian jurisdictions, civil liability legislation governs significant aspects of negligence liability for personal injury. Several parts of the *Civil Liability Act 2002* (WA) ('*CLA*') are relevant to the issues of breach of duty and causation.[66] By way of premise, it is important to note that the issue of establishing the duty of care is largely governed by the common law.[67] There are well-established categories of duty of care between health professionals and patients, as well as between a manufacturer and consumer.[68] Conversely, no established category of duty of care exists between regulatory bodies, such as the Therapeutic Goods Administration ('TGA'), and those claiming to be harmed by their decisions. Indeed, the presence of conflicting statutory duties is often fatal to claims against public authorities.[69] The relevance of this approach for present purposes lies in its implication that the TGA and comparable regulators are unlikely to be held liable in negligence for harm arising from the exercise of their statutory functions. No challenge brought against the TGA — the most recent example being the decision of the Federal Court in the case of *Rose v Secretary of the Department of Health and Aged Care*[70] — has been successful to date.

With regard to the issues under consideration in this article, the use of AI tools in clinical practice is unlikely, in and of itself, to trigger judicial interventions for the purpose of establishing a novel category of duty — although such a development is not impossible to envisage in the future. The sticking points in cases involving the clinical use of AI, as the law of negligence currently stands, are more likely to revolve around breach of already existing categories of duty of care, as well as causation. It is to these elements that we now turn our attention.

---

[65] See *Lochgelly Iron Co v M'Mullan* [1934] AC 1, [25] (Lord Wright).
[66] *CLA 2002* (n 13) pt 1A divs 2–3.
[67] See *Rogers v Whitaker* (1992) 175 CLR 479, 489; *Sullivan v Moody* (2001) 207 CLR 562, [47]–[50].
[68] *Gill v Ethicon Sàrl (No 5)* [2019] FCA 1905, [3624]–[3627] (Katzmann J) ('*Gill Trial*').
[69] See *Minister for the Environment v Sharma* (2022) 291 FCR 311, [192]–[195]. See also *Sullivan v Moody* (n 67) [50].
[70] *Rose v Secretary of the Department of Health and Aged Care* [2025] FCA 339.

1 *Breach of Duty of Care*

Section 5B(1) of the *CLA* outlines the principles for determining whether a duty of care has been breached. It specifies that a person is not negligent in failing to take precautions against a risk unless the risk was foreseeable, the risk was not insignificant, and a reasonable person in the same circumstances would have taken precautions.[71] This risk characterisation, also known as the standard of care, is the yardstick against which a court determines the fault of a defendant. A court will assess whether the conduct of a defendant has fallen below the standard and breached a duty of care by reference to the factors outlined in s 5B(2). These include (amongst other relevant things) weighing the probability of harm occurring, the likely seriousness of the harm, the burden of taking precautions against the risk of harm, and the social utility of the activity that creates the risk.

A key difficulty lies in determining how health professionals can be said to have discharged their duty to take reasonable precautions when they lack adequate information about the training dataset and the functioning of the training algorithm. Although transparency has been widely advocated for, AI developers have in recent years adopted a far more secretive stance toward their training data.[72] Many have shifted from providing detailed explanations of the data used to train a particular model to single sentence descriptions.[73] The inability to access such information makes it difficult to anticipate risks of bias or algorithmic error. In the leading case on breach of duty at common law, *Wyong Shire Council v Shirt*, the High Court held that the assessment of whether reasonable precautions were taken depends on the foreseeability of the risk and the steps available to mitigate it.[74] Similarly, the civil liability legislation requires a risk to be both foreseeable and not insignificant to attract an inquiry into a possible breach of duty.[75] Yet, where critical details are obscured by developers, the capacity of

---

[71] *CLA 2002* (n 13) s 5B(1).

[72] Adam Buick, 'Copyright and AI Training Data: Transparency to the Rescue?' (2025) 20(3) *Journal of Intellectual Property Law & Practice* 182, 184.

[73] Yacine Jernite, 'Training Data Transparency in AI: Tools, Trends, and Policy Recommendations', *Hugging Face* (Blog Post, 5 December 2023) <https://huggingface.co/blog/yjernite/data-transparency>.

[74] *Wyong Shire Council v Shirt* (1980) 146 CLR 40, 47 (Mason J).

[75] *CLA 2002* (n 13) s 5B(1)(b). The statutory requirement that a risk be 'not insignificant' under civil liability legislation reflects only a subtle shift from the former common law test of a 'not far-fetched or fanciful' risk: see *Wyong Shire Council v Shirt* (n 74) 47–8. Courts have observed that while it imposes a slightly more demanding standard, the difference is largely semantic and unlikely to alter outcomes in most cases: see, eg, *Shaw v Thomas* [2010] NSWCA 169, [44] (Macfarlane, Beazley and Tobias JJA).

health professionals to foresee and respond to bias and algorithmic risks is constrained. At its core, this is a transparency problem which challenges the fair application of the negligence standard.

In addition to s 5B of the *CLA*, s 5PB specifically addresses the issues of standard of care and breach of duty in the context of the acts or omissions of health professionals.[76] In essence, the provisions establish that a health professional will not be considered negligent where they have acted 'in accordance with *a practice* that, at the time of the act or omission, is widely accepted by the health professional's peers as *competent* professional practice'.[77] For the provisions of this section to be enlivened, two key ingredients are necessary: there must be 'a widely accepted practice', and that practice must be considered 'competent' by peer professionals at the time of the events.[78]

In the clinical context, *competence* has been referred to as the 'habitual and judicious use of communication, knowledge, technical skills, fundable reasoning, emotions, value, and reflection in daily practice'.[79] Both the wording of s 5PB and established case law allow for the contemporaneous existence of more than one practice regarded as 'competent', accommodating for variation in professional approaches. At first blush, this flexibility appears particularly relevant in the context of rapid advancements in AI-driven healthcare. Given that AI clinical tools are continuously evolving, health institutions are likely to adopt different risk management strategies based on the specific tool and the extent of established medical consensus on their relative risks.

Notably, the WA Court of Appeal in *Child and Adolescent Health Service v Mabior* interpreted the existence of a practice widely accepted as competent at the time of the events very narrowly.[80] This case involved 16-month-old Sunday John Mabior who suffered severe scald burns. It was alleged that hospital doctors negligently failed to recognise and treat his sepsis and thus caused Sunday to suffer harm in the form of Systemic Inflammatory Response Syndrome ('SIRS')

---

[76] Note that a section specifically limited to the standard of care of health professionals is unique to the civil liability legislation of WA. Most other jurisdictions have a general provision on professional liability, typically in the form of a defence. For a discussion of the topic, see Marco Rizzi, 'Health Professionals' Standard of Care and Breach of Duty in Western Australia: A Requiem for the 'Peer Professional' Test at a Time of Uncertainty' (2020) 26(2) *Torts Law Journal* 179.
[77] *CLA 2002* (n 13) s 5PB(1) (emphasis added).
[78] See *Child and Adolescent Health Service v Mabior* (2019) 55 WAR 208, [313]–[327].
[79] Ronald Epstein and Edward Hundert, 'Defining and Assessing Professional Competence' (2002) 287(2) *JAMA* 226, 226.
[80] *Child and Adolescent Health Service v Mabior* (n 78). In this case, the doctors failed to suspect, recognise and treat sepsis in the toddler, and did not administer antibiotics in a timely manner, which would have been a reasonable and practicable precaution.

and Acute Respiratory Distress Syndrome ('ARDS'), ultimately leading to cardiac arrest and multiple-organ failure.[81] The trial judge found the doctors negligent, concluding their actions were not in accordance with a widely accepted practice, which in this case would have required treatment for sepsis.[82] The Court of Appeal upheld this decision, providing a strict textual interpretation of s 5PB of the *CLA*. They emphasised that for s 5PB to apply, there must be an *established course of conduct* that was widely accepted as an *actual fact* at the time of the incident, rather than merely an opinion.[83]

When applied to an AI clinical tool, this requirement, for a health professional to demonstrate the existence of a widely accepted practice considered as competent by the peer professional group at the time of the incident, represents a significant challenge. The Black Box nature of algorithms, limited disclosure about training data, and the risk of automation bias create evidential obstacles in negligence claims. These factors complicate the determination of whether reliance or non-reliance on an AI clinical tool constitutes a widely accepted practice as competent by a peer professional group at the time of the incident. In traditional medical procedures, competence can be validated through peer-reviewed literature, clinical trials, and professional guidelines. However, when considering whether to accept or reject an AI clinical tool's recommendation, the assessment becomes less straightforward. If a health professional chooses not to follow the suggestion of an AI clinical tool, would that omission be regarded as a failure to follow a widely accepted practice? If the tool is not used, would that omission be regarded as incompetent practice at the relevant time, given that departmental protocol required its prior use? These questions illustrate the uncertainty in relation to how courts will assess liability in cases where AI-driven decisions either deviate from conventional practices, or where no established and widely accepted professional courses of conduct yet exist.[84]

---

[81] Ibid [2].
[82] Ibid [82] referring to the primary reasons of the trial judge: see *Mabior v Child and Adolescent Health Service* [2018] WADC 12 [714].
[83] Rizzi (n 76) 8.
[84] Ibid. Note that s 5PB of the *CLA 2002* (n 13) allows for diverging views to co-exist. A widely accepted practice does not have to be the only accepted practice. It does not even have to be the dominant one — it simply must be accepted by a significant proportion of the health professional's peers and be deemed reasonable by the court.

The Australian medical community has taken steps to develop guidance documents and policies on AI clinical tools.[85] Even though these documents are non-binding, the WA and New South Wales ('NSW') courts in *Ellis v East Metropolitan Health Service*[86] and *South Western Sydney Local Health District v Gould*[87] confirmed that appropriate adherence to college, hospital, or other accepted professional guidelines constitutes persuasive evidence that the defendant health professionals were practising in accordance with widely held peer professional opinion.[88] In WA, this is likely to be relevant regardless of the applicability of s 5PB of the *CLA*. Even where the special standard of care is not enlivened by the circumstances, the normal assessment of standard of care and breach, conducted under ss 5B(1)(c) and (2), allows taking into consideration all relevant factors, which certainly can include professional guidelines.[89] Below are examples of such guidelines with an assessment as to their potential relevance in the WA civil liability context.

In 2024, the Australian Health Practitioner Regulation Agency ('AHPRA') published guidance outlining how health practitioners can meet their existing professional obligations when using AI.[90] The guidance makes it clear that individual health professionals remain ultimately responsible for the outcomes of using AI tools in their clinical practice, requiring them to critically evaluate and apply professional judgement to the outputs of such tools rather than relying solely on the generated results.[91] The guidance also outlines key professional

---

[85] See, eg, Australian Medical Association ('AMA'), 'Artificial Intelligence in Healthcare' (Position Statement, 8 August 2023) <https://www.ama.com.au/sites/default/files/2023-08/Artificial%20Intelligence%20in%20Healthcare%20-%20AMA.pdf>; 'Meeting Your Professional Obligations When Using Artificial Intelligence in Healthcare', *Australian Health Practitioner Regulation Agency* ('*AHPRA*'), (Web Page, 22 August 2024) <https://www.ahpra.gov.au/Resources/Artificial-Intelligence-in-healthcare.aspx>; Australian Alliance for Artificial Intelligence in Healthcare ('AAAIH'), 'A National Policy Roadmap for Artificial Intelligence in Healthcare' (Guidelines, 2023) <https://aihealthalliance.org/wp-content/uploads/2023/11/AAAiH_NationalPolicyRoadmap_FINAL.pdf>.

[86] *Ellis v East Metropolitan Health Service* [2018] WADC 91.

[87] *South Western Sydney Local Health District v Gould* (2018) 97 NSWLR 513.

[88] David Pakchung, Morag Smith and Catherine Hughes, 'The Role of Clinical Guidelines in Establishing Competent Professional Practice' (2019) 48(1–2) *Australian Journal of General Practice* 22, 23. The authors refer to the judgment in *Ellis v East Metropolitan Health Service* (n 86) and *South Western Sydney Local Health District v Gould* (n 87), illustrating how courts use clinical guidelines as evidence of competent practice, while also recognising that guidelines are not determinative of the legal standard of care.

[89] Rizzi (n 76) 188.

[90] *AHPRA* (n 85).

[91] Ibid.

obligations including the requirements to ensure human oversight, to understand how the AI clinical tools are trained, and to recognise their inherent biases.[92]

Similarly, the Australian Medical Association emphasizes the importance of transparency and explainability in AI use within healthcare, highlighting the need for all stakeholders to understand its role in clinical decision-making, including addressing inherent biases.[93] It also includes a requirement that patients should be informed when diagnoses or treatment recommendations are determined by AI programs.[94]

The Australian Medical Council ('AMC') is the national standards body responsible for accrediting medical education and training programs, with the aim of ensuring quality and consistency across the healthcare sector in Australia.[95] It oversees 25 specialist medical colleges, each of which plays a critical role in guiding clinical practice and professional standards within their respective specialities.[96] Given the growing integration of AI in healthcare, these colleges are expected to provide leadership and guidance on its ethical and clinical implications.[97] However, among the 25 specialist medical colleges under the AMC, only four have published position statements on AI in healthcare.[98] The position statements from three colleges (Australasian College of Dermatologists,[99] Royal Australian College of

---

[92] Ibid.
[93] AMA (n 85).
[94] Ibid.
[95] See 'About the Australian Medical Council', *Australian Medical Council* (Web Page) <https://www.amc.org.au/about-the-amc/>.
[96] 'List of Specialist Medical Colleges', *Australian Medical Council* (Web Page, 4 February 2025) <https://www.amc.org.au/accredited-organisations/assessment-and-accreditation-of-specialist-medical-programs/specialist-medical-colleges/>.
[97] AMA, 'AMA Calls for Expert Clinical Oversight of the Use of AI in Healthcare'(Media Release, 12 August 2025) <https://www.ama.com.au/media/ama-calls-expert-clinical-oversight-use-ai-healthcare>.
[98] See Australasian College of Dermatologists, 'Position Statement: Use of Artificial Intelligence in Dermatology in Australia' (2022); Royal Australian College of General Practitioners ('RACGP'), 'Artificial Intelligence in Primary Care' (Position Statement, 17 April 2025) ; Royal Australian and New Zealand College of Radiologists ('RANZCR'), 'Ethical Principles for AI in Medicine' (Position Statement, 15 September 2023) <https://www.ranzcr.com/whats-on/news-media/ethical-principles-for-artificial-intelligence-for-medicine-version-2>; Royal Australasian College of Medical Administrators, 'RACMA Position Statement on AI in Healthcare' (June 2025) <https://racma.edu.au/app/uploads/2025/06/RACMA-Position-Statement-on-AI-in-Healthcare-Final.pdf>.
[99] Australasian College of Dermatologists, 'Use of Artificial Intelligence in Dermatology in Australia' (Position Statement, July 2025) <https://www.dermcoll.edu.au/wp-content/uploads/2025/07/ACD-Position-Statement-Use-of-Al-in-Dermatology-in-Australia-July-2025.pdf>.

General Practitioners,[100] and Royal Australian and New Zealand College of Radiologists ('RANZCR')[101]) emphasise the importance of transparency and explainability and human oversight. The RANZCR has developed standards of practice for the use of AI in radiation oncology, providing a risk-management framework for safe and effective implementation.[102] The document outlines the standards required, among others, for training in ML and AI, consent and privacy, algorithm selection, clinical oversight, and decision-making.[103] In relation to clinical decision-making, the RANZCR has determined that ultimate responsibility rests with the radiation oncologist.[104] A full discussion of this standard is beyond the scope of this article, but it is worth noting that this kind of non-legal regulatory framework does engage with the issue of allocating responsibility.

While all professional guidelines can gain relevance in the context of a medical negligence claim through the broad reference in s 5B(2) to 'other relevant things', they have the potential to become 'a widely accepted practice', thereby determining the standard of care. Arguably, this is particularly the case for speciality-specific standards and guidelines, especially those issued by RANZCR, which may serve as a guidance for other specialist medical colleges.

2 *Causation*

Section 5C of the *CLA* addresses the requirement of causation in stating that a breach of duty must be a necessary condition for the harm suffered by the plaintiff (factual causation) and that it is appropriate for the scope of liability to extend to the harm (scope of liability).[105] The former is exclusively factual and a restatement of the well-known 'but for test'.[106] The latter is normative, and its purpose is to set limits around a defendant's liability based on policy considerations, the foreseeability of the harm, or the nature of the chain of causation.[107] Proof of causation has long been established as a challenge for

---

[100] 'Artificial Intelligence in Primary Care', *RACGP* (Web Page, 17 April 2025) <https://www.racgp.org.au/advocacy/position-statements/view-all-position-statements/clinical-and-practice-management/artificial-intelligence-in-primary-care>.
[101] RANZCR, *Standards of Practice for Artificial Intelligence in Radiation Oncology* (30 October 2024).
[102] Ibid.
[103] Ibid 11–18.
[104] Ibid 11.
[105] *CLA 2002* (n 13) ss 5C(1)(a), 5C(1)(b), 5C(4).
[106] See *Adeels Palace v Moubarak* (2009) 239 CLR 420, [41]–[44]; *Strong v Woolworths Ltd* (2012) 246 CLR 182, [18].
[107] *Wallace v Kam* (2013) 250 CLR 375, [14], [24].

plaintiffs seeking compensation for harm associated with clinical settings.[108] The patient's physical condition and variables associated with predictions of the impact of a proper diagnosis or treatment have traditionally complicated the plaintiff's requirement of proving both factual causation on the balance of probabilities, and the normative allocation of causal liability.[109] In *Tabet v Gett*, the High Court ruled that loss of a chance of a better medical outcome is not a recognised form of compensable harm in negligence claims.[110] As such, a plaintiff must establish, on the balance of probability, that the defendant's negligence was a necessary condition of the actual personal injury suffered.

3 *The 'State of Negligence' with Respect to AI Clinical Tools.*

The provisions discussed so far form the basis for evaluating whether liability can be established in negligence in situations involving the use of AI clinical tools. They set out the criteria for proving that the defendant breached its duty of care, and that the breach caused the patient's harm. The relevant sections come into play by providing a framework for assessing the actions of health professionals who incorporate AI into their clinical practice. Under s 5B of the *CLA*, determining whether there has been a breach of the duty of care involves evaluating whether the medical professional appropriately used the AI clinical tool based on the foreseeable risks associated with its limitations and the accuracy of its outputs.

Making this determination requires examining if the health professional took reasonable precautions to verify the AI-generated advice, considering factors such as the likelihood of the AI's errors and the severity of potential harm to the patient. As specific courses of action with respect to the use of AI in clinical practice become well established, the alternative standard of care for health professionals under s 5PB is likely to gain relevance. Causation under s 5C is particularly challenging in AI-related cases, as the provision requires a plaintiff to establish that a health professional's reliance on the AI's recommendation was a necessary condition for the harm, distinguishing this from other causal influences such as the AI's inherent flaws, user misuse, or a combination of factors.

In situations in which the use of an AI clinical tool by a health professional is considered negligent, and harm to a plaintiff ensues because of the AI's inherent flaws rather than their misuse, interesting questions arise with respect to the allocation of liability. Under the current WA legal regime, health professionals

---

[108] See, eg, ibid; *Tabet v Gett* (2010) 240 CLR 537; *Waller v James* [2013] NSWSC 497.
[109] *Tabet v Gett* (n 108) [54]–[62].
[110] Ibid [21]–[25]. See Meredith Blake, 'Loss of a Chance in Medical Negligence: A Lost Cause' (2009) 17(3) *Tort Law Review* 123, 125.

retain responsibility for clinical decision-making and cannot be fully exonerated of liability. However, where the use of an AI tool is involved, liability could arguably be jointly attributed to the AI developer, the institution adopting the AI, and potentially other stakeholders. The Black Box nature of AI discussed above, specifically its lack of transparency and predictability, complicates the fundamental questions around attribution of liability, particularly regarding the standard of care and causation. There are clear challenges for both health professionals and adjudicators in evaluating the reasoning behind AI-generated recommendations. It therefore remains unclear whether liability should rest solely with clinicians or extend to developers and institutions.[111]

This issue is of pressing concern as the risks identified in the previous part of this article are particularly alive in the clinical context. For example, the deployment of clinical decision support systems ('CDSS') is linked to increasing forms of automation bias among clinicians.[112] This is of concern at a time when health services across the world are actively integrating these systems into routine care delivery, increasing their scale of use, frequency of application, and clinician reliance.[113] For example, in the United Kingdom, the National Health Service has already deployed AI in high-volume settings such as radiology, pathology, and diagnostic decision support, where automated tools assist in detecting breast cancer, analysing chest radiographs, and prioritising high-risk patients.[114] These overseas developments have particular relevance for the WA context, where provisions specifically dedicated to the standard of care and liability of health professionals will be put to the test while 'widely accepted' practices emerge. It is worth noting that the particular legal landscape of WA creates a lack of clarity

---

[111] Smith and Fotheringham (n 5) 136–7; Nicholson Price II, 'Artificial Intelligence in Health Care: Applications and Legal Issues' (2017) 14(1) *SciTech Lawyer* 10, 11.

[112] Susanne Gaube et al, 'Do as AI Say: Susceptibility in Deployment of Clinical Decision-Aids' (2021) 4(31) *NPJ Digital Medicine* 1, 4.

[113] Priya Jeyaraj, 'Role of Artificial Intelligence in Enhancing Healthcare Delivery' (2023) 11(12) *International Journal of Innovative Science and Modern Engineering* 1, 10–11.

[114] Jordan Smith et al, 'Real-World Performance of Autonomously Reporting Normal Chest Radiographs in NHS Trusts Using a Deep-Learning Algorithm on the GP Pathway', *arXiv* (Preprint, 28 June 2023) <https://arxiv.org/abs/2306.16115>; Sam Blanchard, 'AI SAY "Landmark" Moment as NHS Clinics Use AI to "Detect Breast Cancer Cases Earlier and Faster"', *The Scottish Sun* (online, 4 February 2025) <https://www.thescottishsun.co.uk/health/14275629/artificial-intelligence-ai-nhs-trial-breast-cancer-screening/>; Andrew Gregory, 'Algorithm could help prevent thousands of strokes in UK each year', *The Guardian* (online, 28 December 2024) <https://www.theguardian.com/society/2024/dec/28/algorithm-could-help-prevent-thousands-of-strokes-in-uk-each-year>.

around the standard of care expected of health professionals, which can cause uncertainty among health professionals and encourage defensive practices.[115]

4 *Expanding the Picture: Liability of the Developer?*

As previously discussed, there are two major sources of the risks of harm associated with clinical AI tools: the four types of bias identified, and errors of the AI clinical tool. Where harm has resulted to a patient because of bias or error, it is arguable that the developer has breached their duty of care, which is likely to fall within the established category owed by manufacturers to consumers.[116] This is the case notwithstanding a finding that the health professional is liable where the use of AI clinical tools is found to be in breach of their specific duty of care.

The rationale for integrating the developer of the AI clinical tool within the scope of one claim for medical negligence (as opposed to treating it as a separate instance of liability) can be summarised as follows. The development of the algorithm and the data training that informs the AI clinical tool lie within the developer's purview. While they may not always have direct control over the data source, particularly if it is provided by third parties or acquired through external datasets, they oversee critical aspects such as the selection, preprocessing, and curation of the data used for training.[117] Consequently, developers should reasonably foresee the potential for user bias in algorithms designed to continuously learn from user-provided data, as these systems inherently reflect and amplify patterns in the input data.[118] Since user-generated data can be influenced by individual behaviours, cultural norms, or systemic biases, algorithms may inadvertently adopt and reinforce these biases over time.[119] This issue is well-documented in ML literature, emphasising the need for proactive bias mitigation strategies during the algorithm's design and deployment stages to ensure fair and reliable outputs.[120]

---

[115] Tom Lawton et al, 'Clinicians Risk Becoming "Liability Sinks" for Artificial Intelligence' (2024) 11(1) *Future Healthcare Journal* 100007 1, 4.

[116] *Gill Trial* (n 68) [3755]–[3761], [3797]–[3807], [3828]–[3834] (Katzmann J). The Court determined Ethicon owed and breached a duty of care to patients by failing to adequately test device safety before release, appropriately monitor devices post-approval, or provide proper warnings about risks and probabilities.

[117] See generally El Kindi Rezig, Michael Cafarella and Vijay Gadepally, 'Technical Report on Data Integration and Preparation', *arXiv* (Preprint, 2 March 2021) <https://arxiv.org/abs/2103.01986>.

[118] See Sribala Chinta et al, 'AI-Driven Healthcare: A Review on Ensuring Fairness and Mitigating Bias' (2025) 4(5) *PLOS Digital Health* 1, 14–15.

[119] Ibid 8.

[120] Kirsten Martin, 'Ethical Implications and Accountability of Algorithms' (2018) 160(4) *Journal Of Business Ethics* 835, 845–6.

5 *Refining the Picture: Limitations of the Human-in-the-Loop*

As noted earlier, automation bias occurs when health professionals develop an over-reliance on AI clinical tools, such as CDSS, which can lead to potentially harmful outcomes when these systems provide inaccurate or incomplete advice.[121] This bias is produced by the perception that AI systems are highly reliable, which may discourage clinicians from questioning or verifying their outputs.[122] The main contributors to this reliance include clinicians' excessive trust in the system and the assumption that the tools are safe and effective simply because they are adopted and provided by the hospital.[123] Additionally, practical constraints such as limited time and high workloads may reduce clinicians' ability to critically assess AI-generated outputs, inadvertently reinforcing reliance on these tools.[124] This raises questions about the legal responsibilities of the human-in-the-loop ('HITL').

The HITL approach, where clinicians are tasked with verifying or correcting errors flagged by the AI system before finalising decisions, is intended to enhance oversight and reduce errors caused by blind reliance on automation.[125] However, questions remain about the effectiveness of this strategy. Specifically, concerns arise regarding whether institutions are adequately providing the rigorous training necessary to equip clinicians with the skills to critically engage with AI outputs, rather than passively accepting them.[126] Furthermore, there is ongoing debate about whether regulatory bodies should impose stricter oversight to ensure the safe integration of AI clinical tools into healthcare workflows, to minimise the risk that the use of these tools may undermine patient safety.[127]

---

[121] Kate Goddard, Abdul Roudsari and Jeremy Wyatt, 'Automation Bias: A Systematic Review of Frequency, Effect Mediators, and Mitigators' (2011) 19(1) *Journal of American Medical Informatics Association* 121, 123.

[122] Kun-Hsing Yu and Isaac Kohane, 'Framing the Challenges of Artificial Intelligence in Medicine' (2019) 28(3) *BMJ Safety & Quality* 238, 239.

[123] Lyell et al (n 58) 8.

[124] Emely Rosbach et al, 'Automation Bias in AI-Assisted Medical Decision-Making under Time Pressure in Computational Pathology' (Conference Paper, Bildverarbeitung für die Medizin, 2 March 2025) 4.

[125] Ujué Agudo et al, 'The Impact of AI Errors in a Human-in-the-Loop Process' (2024) 9(1) *Cognitive Research: Principles and Implications* 1, 2.

[126] Lotte Schuitmaker et al, 'Physicians' Required Competencies in AI-assisted Clinical Settings: A Systematic Review' (2025) 153(1) *British Medical Bulletin* 1, 9.

[127] Haider Warraich, Troy Tazbaz and Robert Califf, 'FDA Perspective on the Regulation of Artificial Intelligence in Health Care and Biomedicine' (2025) 333(3) *JAMA Network* 241, 244–5.

Further research will be necessary to fully address the challenges of the HITL approach, particularly in light of recent global developments that have established it as a regulatory requirement.[128] However, for present purposes, the HITL approach raises the question of what constitutes a competent professional practice in relation to the deployment and use of AI clinical tools. In the absence of case law, we suggest that the proposition that reliance on AI clinical tools is widely accepted as competent professional practice would be relevant in the context of establishing the standard of care and assessing breach (whether under s 5B or 5PB of the *CLA*). Notwithstanding this, it is unlikely that complete, unquestioning reliance would be regarded as exonerating, even where the use of a specific AI clinical tool rises to the level of a widely accepted practice per s 5PB.

A separate but related issue is that arising in connection with the health professional's duty to inform patients of the risks associated with a process of diagnosis or treatment that relies on AI tools. The Black Box nature of algorithmic processes makes it difficult to anticipate risks of harm that might constitute 'material risks' for the purposes of the legal test for disclosure of information. As established in *Roger v Whitaker* and subsequently reaffirmed in *Wallace v Kam*,[129] a risk is material if a reasonable person in the patient's position is likely to attach significance to it, even if the probability of the risk of materialising is low or altogether uncertain.[130] In this sense, the intrinsic risks associated with AI, discussed above, add a further layer of complexity to the health professional's duty to disclose risks, making it more challenging to determine what must be communicated, especially in light of the uncertainty around the general risks associated with AI clinical tools.

B *Australian Consumer Law*

The Australian Consumer Law ('ACL') is a statutory regime that operates alongside the tort of negligence.[131] Its relevance in the context of AI clinical tools lies in the nature of such tools as products, the design and application of which are subject to the protections offered by this regime. The ACL provides a comprehensive framework designed to offer remedies to consumers who are harmed through the conduct of corporate bodies or commercial entities. Liability under the ACL can be established through multiple legal bases, including misleading and deceptive conduct, breach of consumer guarantees, and liability

---

[128] *Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 on Artificial Intelligence (Artificial Intelligence Act)* [2024] OJ L 12.7.2024, art 14.
[129] *Rogers v Whitaker* (n 67) 490; *Wallace v Kam* (n 107) [8].
[130] *Rogers v Whitaker* (n 67) 490.
[131] *Competition and Consumer Act 2010* (Cth) sch 2 ('Australian Consumer Law').

for defective products.[132] A key feature of the ACL is its compensatory no-fault regime, which provides strict liability under pts 3–5, ensuring that injured parties can seek remedies without needing to prove fault for harm caused by an unsafe product.[133] This is defined in s 9 as a product that does not have the level of safety 'persons generally are entitled to expect'.[134] Section 138 outlines the circumstances in which this strict liability regime is enlivened. Importantly, the protections offered by this provision extend to patients injured by medical devices,[135] which would include AI-powered clinical tools.[136]

The ACL and negligence were both invoked as bases for compensation in the successful challenge by consumers injured by defective pelvic mesh implants in *Ethicon Sàrl v Gill*.[137] The case underscores the complementary roles of ACL and negligence in medical product litigation. The ACL's strict liability provisions under s 138 ensured manufacturer accountability for safety defects,[138] while the negligence claims addressed the failure to take reasonable precautions in product design and risk disclosure.[139]

The findings in the *Gill* class action are directly relevant to AI clinical tools, as both involve medical products impacting patient safety. AI-integrated medical devices such as CDSS could also be subject to ACL claims if they were to generate erroneous recommendations, cause or contribute to misdiagnosis, or be otherwise associated with patient harm. Section 138 appears applicable to an AI tool's algorithmic flaws, biases, or system failures resulting in unsafe outcomes for patients.

Furthermore, the *Gill* class action emphasised the importance of manufacturer responsibility in disclosing risks and ensuring adequate testing, a principle that extends to AI developers and software vendors operating in the clinical and

---

[132] Ibid. Sections 18 and 29 specifically address misrepresentations, making it unlawful for businesses to engage in such practices. Parts 3–5 of the Australian Consumer Law impose strict liability on manufacturers and importers for goods with safety defects.
[133] See Jeannie Marie Paterson, *Corones' Australian Consumer Law* (Thomson Reuters, 4th ed, 2023) 504.
[134] For an in-depth analysis of the operation of the Australian Consumer Law with respect to therapeutic goods, and its complex interaction with the regulatory regime that governs the approval of therapeutic goods, see Rizzi, Gleeson and Paterson (n 64).
[135] Julia Symons and Marco Rizzi, 'Consumers or Patients? Medical Device Recipients under Australian Law Straddle Two Worlds' (2023) 30(3) *Journal of Law and Medicine* 572, 573.
[136] Reynolds (n 64) 502.
[137] *Ethicon Sàrl v Gill* (2021) 288 FCR 338, [4] ('*Gill Appeal*'), on appeal from *Gill Trial* (n 68) — together the cases are referred to as the '*Gill* class action'.
[138] *Gill Appeal* (n 137) [266]–[271].
[139] Ibid [715]–[723].

healthcare context.[140] The Federal Court made clear that the sufficiency of warnings does not depend on the manufacturer's subjective view of what medical professionals ought to know.[141] Instead, manufacturers must actively provide sufficient information, advice, and warnings about potential risks to ensure that clinicians and consumers can make informed decisions.[142] This reasoning is particularly relevant to AI clinical tools, where algorithmic failure modes, bias, and data limitations can significantly impact patient safety. If an AI clinical tool is deployed without sufficient validation, lacks transparency in decision-making, or fails to account for known limitations, manufacturers could face liability under the ACL for safety defects.

This would be particularly relevant where an AI model fails to generalise across diverse populations, leading to inaccurate diagnoses or treatment recommendations. As Reynolds' in-depth analysis of AI-powered medical devices notes, AI models often underperform when deployed on populations that differ from the training data, raising serious concerns about generalisability, bias, and health disparities — in other words, they can be prone to data bias, as described above.[143] AI manufacturers must manage the expectations of both patients and clinicians,[144] and provide explicit guidance on the intended use of their tools. If an AI clinical tool has been trained primarily on homogeneous patient datasets but is then deployed for broader use without appropriate disclaimers or safeguards, the risk of systemic errors increases. In the *Gill* class action, the Federal Court indicated that a medical product would not necessarily be defective if clear warnings and limitations were provided to clinicians, allowing them to make balanced, cautious, and informed judgements.[145] By analogy, AI developers must clearly communicate the limitations of their models, ensuring that users are aware of potential risks and constraints.

It is worth noting at this point that the European Union has recently adopted a revised version of its Product Liability Directive ('Directive').[146] To overcome the significant hurdles that individuals harmed by AI-powered products may encounter in demonstrating the safety defectiveness of a product, the new

---

[140] See, eg, Australian Government, Department of Health, Disability and Ageing, *Demonstrating Evidence to Comply with the Essential Principles* (Guidelines, 25 August 2022) 1, 3.
[141] *Gill Appeal* (n 137) [607].
[142] Symons and Rizzi (n 135) 580.
[143] Reynolds (n 64) 516.
[144] Symons and Rizzi (n 135) 579.
[145] *Gill Appeal* (n 137) [608].
[146] *Directive (EU) 2024/2853 of the European Parliament and of the Council of 23 October 2024 on Liability for Defective Products and Repealing Council Directive 85/374/EEC* [2024] OJ L 18.11.2024.

Directive establishes a series of presumptions that significantly ease the plaintiff's burden.[147] This is certainly an interesting development at the international level, and it will be interesting to follow whether the current 'Review of AI and the Australian Consumer Law', led by the federal Australian Treasury, will reach similar conclusions.[148] The process is however still very much ongoing, and no reform is expected in the immediate future.

V EXPLORING ALTERNATIVES TO ESTABLISHED LIABILITY MODELS

A *Underpinning Normative Challenges*

The discussion so far has outlined the challenges that AI clinical tools can raise for established liability frameworks in the face of patient harm associated with their use in the clinical setting. As this article aims to explore equitable ways to account and compensate for patient harm, it is important to expose the limitations of existing frameworks in order to present viable alternative mechanisms to address this goal. The (at least partial) mismatch between negligence and consumer law frameworks is no surprise. As outlined above, negligence is a fault-based liability regime and, as such, it focuses on proving that the conduct of an identifiable actor, such as a doctor or hospital, has fallen below the expected standard of care and caused an identifiable harm. Product liability on the other hand addresses the safety defect in particular and identifiable products. It responds to the separate logic of protecting consumer safety by removing the fault requirement and shifting the focus away from the conduct of the defendant, relocating that focus on the product and the legitimate safety expectations of the public.

Both the negligence and product liability frameworks can therefore struggle to address complex allocations of responsibility when harm arises from the use of AI clinical tools, as this arguably requires a twofold focus on both product and conduct, which in turn requires consideration of the contributions of different parties. In this sense, both negligence and product liability are founded on normative premises that are not necessarily suited to recognising the collective and interdependent nature of AI-driven decision-making, where liability may be diffused across multiple stakeholders, including developers, manufacturers, healthcare providers, and regulatory bodies. While complexity in allocation of

---

[147] See Miriam Buiten, 'Product Liability for Defective AI' (2024) 57(1–2) *European Journal of Law and Economics* 239, 267.

[148] Australian Government Treasury, *Review of AI and the Australian Consumer Law* (Discussion Paper, 15 October 2024) 3 <https://treasury.gov.au/sites/default/files/2024-10/c2024-584560-dp.pdf>.

liability (particularly with respect to the use of technology in clinical care) is certainly not a novel situation, the types of risks and harms involved with the use of AI clinical tools as described in Part III provide us with a unique opportunity to reflect on the viability of our current models. Indeed, the mismatch between those risks and the characteristics of the legal frameworks analysed in Part IV can lead to accountability gaps, or to liability being unfairly concentrated on frontline medical professionals who may have limited control over the AI's outputs. So, as AI clinical tools become more sophisticated and embedded in decision-making, including in the WA healthcare system, the time is ripe to explore alternative approaches to redressing harm.

B *Considering Alternatives: Collective Responsibility and No-Fault Compensation, or Risk Pooling for AI Clinical Tools*

Alternative approaches to centring redress on proof of an individual actor's fault or an identifiable product defect embrace the notion that loss distribution mechanisms should reflect the distinct normative underpinning of societal responsibility for individuals' harm. The topic is vast, and a full review of these approaches is beyond the scope of this article. However, it is important to note that the idea of shifting from liability models and refocusing on alternative methods that centre on ensuring recovery for harms suffered may be well-suited to the present context, given the social utility associated with the use of AI as a way to improve accessibility to high quality healthcare treatments.[149]

Separate from, but connected to, issues of fair and adequate compensation is the equally pressing question of identifying appropriate models for the allocation of liability. Clara Cestonaro et al, for example, identified six broad approaches to liability arising in connection with the use of AI in clinical settings in their systematic review.[150] These are medical malpractice, product liability, learned intermediary rule, common enterprise liability, vicarious liability, and AI personhood.[151] The first two represent the traditional regimes that are currently available in WA and have been unpacked in the previous part. The other approaches provide useful elements worth reflecting upon, particularly the idea of common enterprise liability.

---

[149] Reynolds (n 64) 526.
[150] Clara Cestonaro et al, 'Defining Medical Liability when Artificial Intelligence is Applied on Diagnostic Algorithms: A Systematic Review' (2023) 10 *Frontiers in Medicine* 1, 8.
[151] Ibid 5–7.

A central challenge in clinical AI regulation is addressing the responsibility gap, ensuring that liability for patient harm is fairly distributed. This challenge is compounded by the fact that the use of AI clinical tools involves multiple stakeholders, each with distinct roles and responsibilities. Collective responsibility refers to the shared accountability of relevant stakeholders for the outcomes of AI-driven decisions.[152] This approach recognises that AI systems are the product of collaborative efforts involving clinicians, developers, institutions, and regulators, and that each stakeholder plays a role in ensuring the safety and efficacy of AI in clinical settings. The case for collective responsibility is rooted in the recognition that harms associated with the use of AI clinical tools are often the result of systemic failures rather than individual negligence or the defectiveness of a particular product.[153]

By distributing responsibility among stakeholders, collective responsibility models can ensure that each party is held accountable for its role in the AI system's development and deployment. Within this framework, health professionals are ultimately responsible for patient care and must exercise professional judgement when using AI clinical tools. However, the model holds that their responsibility should be limited to their actual influence over the decision-making process.[154] This is complemented by developers holding responsibility for ensuring that AI algorithms are accurate, transparent, and free from biases. Developers must also provide adequate documentation and support to enable clinicians to use the tools effectively.[155] In parallel, healthcare institutions are responsible for procuring and deploying AI systems, as well as for establishing protocols for their use, ensuring that clinicians are adequately trained to interpret and act on AI-driven recommendations. Regulators play a crucial role in this model by establishing standards for AI development and deployment, as well as overseeing the implementation of safety and efficacy

[152] Isaac Taylor, 'Collective Responsibility and Artifical Intelligence' (2024) 37 *Philosophy and Technology* 1, 3.
[153] Bart Custers, Henning Lahmann and Benjamyn Scott, 'From Liability Gaps to Liability Overlaps: Shared Responsibilities and Fiduciary Duties in AI and Other Complex Technologies' (2025) 40(5) *AI and Society* 4035, 4037.
[154] Changhyun Lee, Hun-Yeong Kwon and Kyung Jin, 'Sharing Accountability of Versatile AI Systems: The Role of Developers and Practitioners' (2024) 37(1) *The International FLAIRS Conference Proceedings* 1, 2. The authors introduced the concepts of controllability and openness, where practitioners often lack the techniques to modify the algorithm model.
[155] Antian Chen, Chenyu Wang and Xinqing Zhang, 'Reflection on the Equitable Attribution of Responsibility for Artificial Intelligence-Assisted Diagnosis and Treatment Decisions' (2022) 3(2) *Intelligent Medicine* 139, 143.

protocols.[156] The clear definition of these roles is essential in distributing responsibility across the AI ecosystem.

In light of these challenges, the question of how collective responsibility should be implemented arises. Based on the literature analysed in this article, we suggest the following two primary mechanisms for implementing collective responsibility: (1) a two-pronged approach based on the establishment of a no-fault compensation fund supplemented by enterprise liability; and (2) the adoption of risk pooling frameworks within the enterprise liability model.

1 *No-Fault Compensation and Enterprise Liability*

Søren Holm et al argue that AI systems, due to their opacity and complex causal chains, render traditional fault-based liability both impractical and unjust, particularly where clinicians are held accountable for outcomes influenced by AI clinical tools they do not fully control.[157] They propose that no-fault compensation schemes, designed to circumvent the burden of proving negligence, can provide more equitable outcomes for patients while supporting transparency and error reporting within healthcare systems.[158] Nynke Vellinga advances a policy-oriented argument for an AI-specific compensation fund for high-risk systems.[159] This fund, to be financed by AI developers, distributors, and institutional users, would aim to ensure full compensation for personal injuries, including those involving mental harm, where the precise source of failure is indeterminate.[160] This is in line with the call from the World Health Organisation for the introduction of a 'Faultless Responsibility Model' under which all agents involved in the development and deployment of AI are required to promote integrity and minimise harm.[161] A key issue for any no-fault compensation regime, as identified by Genevieve Grant and Harold Luntz, is to ensure that it is tailored to effect 'a more just and comprehensive response to the burden of injury

[156] Argyri Panezi, 'Liability Rules For AI-Facilitated Wrongs: An Ecosystem Approach To Manage Risk And Uncertainty' in Pablo García Mexía and Francisco Pérez Bes (eds), *Artificial Intelligence and the Law* (La Ley, 2021) 1, 15–16.

[157] Søren Holm, Catherine Stanton and Benjamin Bartlett, 'A New Argument for No-Fault Compensation in Health Care: The Introduction of Artifcial Intelligence Systems' (2021) 29 *Health Care Analysis* 171, 173–4.

[158] Ibid.

[159] Nynke E Vellinga, 'Rethinking Compensation in Light of the Development of AI' (2024) 38(3) *International Review of Law, Computers & Technology* 391, 408.

[160] Ibid 406.

[161] World Health Organisation, *Ethics and Governance of Artificial Intelligence for Health* (Guidance, 28 June 2021) 1, 28 <https://www.who.int/publications/i/item/9789240029200>.

and disease on the basis of disability and need, rather than cause'.[162] The transformative potential of the technology in clinical care arguably calls for a model that offers secure safety nets to patients when the risks identified in Part III materialise. The complexity of AI clinical tools may otherwise put injured patients at a serious structural disadvantage if required to individually pursue liability actions (whether in negligence or product liability).[163]

In parallel to no-fault compensation, which addresses the challenge of equitably responding to harm, is the issue of actual attribution of liability. One suggested model is that of enterprise liability, in which institutional actors, including hospitals and AI developers, collectively bear legal responsibility.[164] This model acknowledges that accidental harm is often a by-product of organised activities rather than isolated individual actions. It reflects the reality that harm frequently results from systemic interactions involving technology, institutional processes, and clinical practice.[165] It is also grounded in the principle that those who benefit from an activity should also bear its burdens, thereby promoting fairness and accountability.[166] It may involve holding developers, manufacturers, or deployers of AI systems jointly liable for harm associated with the use of AI clinical tools. By attributing liability to organisations as part of the broader enterprise engaged in the design, deployment, and governance of AI clinical tools, enterprise liability can provide a structurally responsive framework to the multilayered nature of the challenges posed by AI clinical tools.

While theoretically sensible, this two-pronged approach can be politically challenging as shifting away from established liability models is not necessarily practical. As Mark Forwood highlights, Australian jurisdictions have historically shown limited appetite for broad no-fault compensation schemes, with the Whitlam Government's Woodhouse proposals and later initiatives after the 2002 insurance crisis failing to secure national adoption.[167] Similarly, a 'No Fault Compensation' paper prepared for the NSW Parliament emphasised that political

---

[162] Genevieve Grant and Harold Luntz, 'The Accident Preference, "Unrigorous Thinking" and Injury Compensation Schemes' in Kylie Burns et al (eds), *Torts on Three Continents: Honouring Jane Stapleton* (Oxford University Press, 2024) 473, 473.
[163] Luciano Floridi et al, 'AI4People: An Ethical Framework for a Good AI Society: Opportunities, Risks, Principles, and Recommendations' (2018) 28 *Minds & Machines* 689, 694.
[164] Benny Chan, 'Applying a Common Enterprise Theory of Liability to Clinical AI Systems' (2021) 47(4) *American Journal of Law & Medicine* 351, 352, 375–6.
[165] Gregory Keating, *Reasonableness and Risk: Right and Responsibility in the Law of Torts* (Oxford University Press, 2024) 336, 359.
[166] Ibid 341–2.
[167] Mark R Forwood, 'Whither No-Fault Schemes in Australia: Have we Closed the Care and Compensation Gap?' (2018) 43(3) *Alternative Law Journal* 166, 166.

caution, cost concerns, and resistance from insurers and litigation lawyers have constrained reform, even as technological and systemic changes have exposed gaps in fault-based systems.[168] These findings are consistent with the ongoing resistance of both federal and state governments to deploy general no-fault compensation schemes for vaccine injuries, despite their common nature in comparable countries.[169] However, existing no-fault schemes, such as WA's Catastrophic Injuries Support scheme, are confined to clearly identified types of harm or contextual settings.[170] When emerging risks, such as catastrophic motor injuries, have revealed intrinsic weaknesses of traditional liability models, calls for no-fault approaches have re-emerged. Arguably, AI in clinical settings presents a comparable inflection point. In August 2011, the Productivity Commission ('PC') of the federal Treasury recommended the establishment of a National Injury Insurance Scheme ('NIIS').[171] The PC recommended that a NIIS be developed for catastrophic injuries caused by four types of accidents: motor vehicle accidents, workplace accidents, medical accidents, and general accidents (occurring in the home or community).[172] Subsequently, the Treasury published a 'draft without prejudice' medical treatment injury discussion paper to examine the issues around possible minimum benchmarks for compensation for medical injury within a NIIS.[173] Revisiting no-fault models, whether through a tailored extension of the WA scheme or the NIIS, or through the introduction of a bespoke one, could provide timely redress for AI-related harms.

---

[168] Talina Drabsch, NSW Parliamentary Library Research Service, *No Fault Compensation* (Briefing Paper No 6/05, May 2005) 11.

[169] Duncan Fairgrieve et al, 'Comparing No-Fault Compensation Systems for Vaccine Injury' (2023) 31(1) *Tulane Journal of International and Comparative Law* 75, 78, 82–4; Australian Government, Department of the Prime Minister and Cabinet, *Commonwealth Government COVID-19 Response Inquiry Report* (Inquiry Report, October 2024) 270–1; Anastasia Tsirtsakis, 'No-Fault Indemnity Should Apply to all Vaccines', *News GP* (online, 10 August 2021) <https://www1.racgp.org.au/newsgp/professional/no-fault-indemnity-should-apply-to-all-vaccines-ra>.

[170] 'Catastrophic Injuries Support', *Insurance Commission of Western Australia* (Web Page) <https://www.icwa.wa.gov.au/motor-injury-insurance/catastrophic-injuries-support>. The Catastrophic Injuries Support ('CIS') Scheme is a no-fault scheme which provides eligible people with necessary and reasonable treatment, care, and support, in some cases for their lifetime. Anyone who experiences serious personal injuries as a result of a motor vehicle accident in WA on or after 1 July 2016, or a workplace accident on or after 1 July 2024, may be eligible to receive support from the CIS Scheme.

[171] 'National Injury Insurance Scheme', *The Treasury* (Web Page) <https://treasury.gov.au/programs-initiatives-consumers-community/niis>.

[172] Ibid.

[173] Australian Government Treasury, *Medical Treatment Injury Discussion Paper* (Draft Discussion Paper, 2019) 1, 2–4 <https://treasury.gov.au/sites/default/files/2019-03/Medical_treatment_injury_discussion_paper.pdf>.

2 *Risk Pooling*

Another proposed model to address the challenges of current liability frameworks is risk pooling, which involves the spreading of liability among stakeholders, typically through insurance mechanisms.[174] It provides a structure for compensating patients harmed through the use of AI clinical tools by distributing risk across developers, healthcare providers, insurers, and institutions. Related to an enterprise liability framework, risk pooling is based on the principle that those who benefit from a risky activity should share responsibility for any resulting harm.[175] This model invites comparison with the United Kingdom's *Automated and Electric Vehicles Act 2018* ('*AEVA 2018*'), a joint insurance scheme where compensation is awarded without fault attribution and insurers can later seek recovery if appropriate.[176] The *AEVA 2018* was enacted to establish an insurance and liability framework for Connected and Autonomous Vehicles ('CAVs'), mandating that insurers compensate third parties in accidents caused by the vehicle's autonomous operation, with a right of recovery against the manufacturer or responsible party.[177] This legislation is regarded as proactive, and therefore suitable for addressing future challenges, and as providing certainty for accident victims and stakeholders, ensuring compensation even in complex liability scenarios.[178] However, the *AEVA 2018* faces practical barriers due to some broad and imprecise definitions associated with, for example, 'automated vehicle' and 'accident', and has therefore been criticised for creating uncertainty for insurer recovery processes and software liability.[179] Such uncertainty has raised concerns in relation to regulatory disconnection and the hindering of innovation.[180] Despite the criticism, the principle of risk pooling can still provide useful guidance to the WA healthcare setting, where the current liability regimes lack clear mechanisms to allocate liability when harm arises from health practitioner reliance on AI outputs. This model also aligns with growing calls at an international level for law reform to

---

[174] Smith and Fotheringham (n 5) 146.
[175] Ibid 147–8.
[176] Ibid 150.
[177] *Automated and Electric Vehicles Act 2018* (UK) ss 2(1), 5(1).
[178] Matthew Channon, 'Automated and Electric Vehicles Act 2018: An Evaluation in Light of Proactive Law and Regulatory Disconnect' (2019) 10(2) *European Journal of Law and Technology* 1, 2.
[179] Ibid 6–7.
[180] Ibid.

consider the use of risk pooling, alongside detailed professional guidance for the use of AI in healthcare spaces.[181]

VI CONCLUSION

This article has examined the current legal frameworks in WA which potentially apply in a scenario where patient harm is associated with the use of AI clinical tools in clinical practice. Rather than assessing whether the current system is fully adequate, this article has presented an overview of the key legal principles that may apply in such situations, particularly those found in negligence law and the ACL. These frameworks provide established mechanisms for addressing liability in healthcare settings, but they were developed in the context of human decision-making and traditional medical products, raising questions about their adequacy in cases involving AI-driven technologies and the specific risks these technologies present.

The discussion highlights the dependence of negligence actions on proof of fault, which can be prohibitive in cases where AI plays a role in clinical decision-making. Similarly, the ACL's product liability provisions are designed for static products rather than evolving AI systems that adapt over time. While these legal mechanisms continue to offer viable avenues for redress, we question their ability to adequately address AI-related harm in their current form.

Regulatory discussions on AI governance in Australia are ongoing, with recent proposals indicating a shift toward a risk-based approach to AI regulation. In September 2024, the federal Department of Industry, Science, and Resources proposed mandatory guardrails for high-risk AI applications,[182] including clinical AI tools, emphasising transparency, accountability, and risk management across the AI lifecycle.[183] Additionally, the federal Department of Health, Disability and Ageing has conducted a legislative review focusing on high-risk AI systems in healthcare.[184] However, at this stage, action is yet to be taken on these proposals,

---

[181] Kit Fotheringham and Helen Smith, 'Accidental Injustice: Healthcare AI Legal Responsibility Must be Prospectively Planned Prior to its Adoption' (2024) 11(3) *Future Healthcare Journal* 1, 2.
[182] Australian Government, Department of Industry Science and Resources, *Safe and Responsible AI in Australia: Proposals Paper for Introducing Mandatory Guardrails for AI in High-Risk Settings* (September 2024).
[183] Ibid 30–42.
[184] Australian Government, Department of Health, Disability and Ageing, *Safe and Responsible Artificial Intelligence in Health Care: Legislation and Regulation Review* (Final Report, March 2025) 1, 3.

and in the absence of legal enforcement mechanisms, they can only serve as policy guidance.

As AI continues to integrate into clinical practice, it is important to consider whether WA's current liability framework provides sufficient clarity for both patients and health professionals when harm occurs in association with the use of AI clinical tools. This article has outlined liability mechanisms currently available in WA and canvassed potential alternatives. In particular, the combination of a no-fault compensation system with enterprise liability or a form of risk pooling could provide a more appropriate alternative to managing AI-related harm. While conceptually viable, these options will of course raise questions around political feasibility, financial implications, and specific details of regulatory reform involved, and further research is needed to fully unpack these aspects.

Finally, further research is also needed to understand how healthcare professionals navigate liability risks under the current system, especially the challenges associated with real-world implications of the HITL approach discussed in this article. Studying how health practitioners respond to AI-related legal challenges in practice will provide critical evidence to understand what reforms are necessary to ensure accountability and confidence in AI-driven healthcare.