AI and professional liability in health and care:

Opportunities for action

CERSI-AI

ACKNOWLEDGEMENTS

This report was led by Ellen Coughlan, Partnerships and Programmes Manager of the Centre of Excellence for Regulatory Science and Innovation in AI and Digital Health (CERSI-AI) and was published in September 2026.

The authors would like to express thanks to Melanie Venables, Director of Policy and Communications at the Professional Standards Authority for Health and Social Care (PSA), and Mani Hussein, Deputy Director of Provider Assurance at the Medicines and Healthcare products Regulatory Agency (MHRA) for co-producing the workshop series that formed the basis of this work. The MHRA, through work undertaken for the National Commission into the Regulation of AI in Healthcare, and the PSA, have each conducted work on this topic that was foundational to this report.

We are grateful to the patients, members of the public, health and care professionals and experts who took the time to participate in the series of workshops in June 2026. Their thoughtful contributions, candour and expertise were instrumental in identifying the challenges and solutions described in this work.

We would like to thank the wider team at CERSI-AI and the Bright group at the University of Birmingham for their work to arrange and facilitate each of these workshops.

EXECUTIVE SUMMARY

Artificial intelligence (AI) tools are being used by health and care professionals across the NHS and social care system. AI systems are being applied thoughtfully and robustly in many settings, but in others are being used without adequate training, organisational support, or oversight.

Health and care professionals report being anxious about their potential liability when patient harm occurs in the context of AI (whether relating to its use or potentially its non-use). They are particularly concerned at the thought that they may be held accountable for the performance of a tool that they cannot personally verify. Health providers also report that manufacturers of AI tools are securing contracts that minimise any liability on themselves, with the burden largely carried by the deploying organisation (and its health professionals). There are concerns that regulatory guidance has not kept pace, indemnity arrangements contain critical gaps, and patients have no reliable means of knowing when AI was used in their care which hinders the seeking of redress. The risk that clinicians become “liability sinks” has become a prevailing concern.

To explore these concerns and to identify challenges and solutions related to clarity and certainty of AI liability and governance, CERSI-AI convened three workshops in June 2026. The first was a workshop with patients and the public, the insights from which fed into the design of the two following days of workshops. These workshops brought together medical defence organisations, insurance and indemnity providers, professional bodies and unions, law firms, academics, professional and health service regulators, NHS resolution bodies, accredited registers and others. This report describes the findings from each workshop, identifying gaps in professional guidance and regulation and making recommendations related to the use of AI in health and care.

Five themes emerged: a regulatory guidance vacuum for professionals, a legal framework that has failed to keep pace with AI, governance and procurement practices that lack consistent application, an AI education and literacy gap that is already surfacing clinical risk, and equity concerns. There was strong consensus and enthusiasm for cross-regulatory collaboration to address these issues. Of the thirteen actions recommended in this report, participants placed most emphasis on the need for professional regulators to issue a joint statement of intent and to develop high-level principles that could inform their guidance and reduce uncertainty. Following the workshops, the PSA has committed to coordinating this work with the professional regulators, as a key contribution to shaping a new regulatory framework for AI in healthcare that is being led by the National Commission into the Regulation of AI in Healthcare.

THE PROBLEM

The application of artificial intelligence (AI) innovations in health and care presents opportunities and challenges for clinicians, patients, providers, regulators and policy-makers. For example, AI systems used to support imaging and diagnostics have already realised benefits for patients, who are being diagnosed more swiftly, and are met with broad support by clinicians who recognise their potential to improve clinical efficiency. Increasing appetite for AI systems to help resolve system pressures in the NHS is demonstrated in the 10 Year Health Plan for England, which envisages a shift towards a digital NHS where all hospitals are "fully AI enabled". However, support is often conditional on technical, ethical and regulatory terms that ensure that technologies that are carefully governed to deliver safe, equitable and effective care. Regulators and policy-makers are facing the challenge of clarifying the appropriate regulatory approach to AI technologies so that health and care providers and practitioners can realise the benefits while understanding and mitigating the risks.

One challenging domain in need of attention is the regulation of health and care professionals using AI. The subject of liability – including professional liability – in relation to AI harms has been the subject of attention from the UK Jurisdiction Taskforce as well as research raising the risk of clinical, professional and reputational safety in the absence of coordinated professional guidance, and was surfaced as a concern by the AI Commission and the PSA. Their engagement sessions have found that health and care professionals are unclear about their liability when things go wrong, and that defence bodies want regulators to provide further clarity on what they expect from their registrants when using AI technologies. Though there is consensus that whilst clinical negligence legal principles (including the Bolam, Bolitho and Montgomery tests) may apply when AI is being used as a tool to aid decisions, the regulatory gap widens when AI transitions into a role where it makes the decisions, or when the capability to critically operate the AI system as intended is compromised by deskilling, biases or system pressures.

To explore this, CERSI-AI brought together various stakeholders across three workshops to understand what regulatory guidance on AI currently exists for professionals and what gaps remain. This report describes the findings from these workshops and makes recommendations regarding collaboration by regulators and others to address these gaps and ensure that AI technologies are adopted and governed safely, equitably and effectively.

These workshops were delivered in June 2026.

Workshop one: patients and the public discussed their perspectives on the use of AI in health and care by professionals and what should happen when things go wrong.

Workshop two: medical defence organisations, insurance and indemnity providers, professional bodies and unions, law firms, academics and others with ethico-legal expertise related to liability in health and care.

Workshop three: professional regulators, health service regulators, NHS resolution bodies, accredited registers and others with expertise related to standards and regulation of health and care delivery.

Themes of discussion

The workshops surfaced that health and care professionals are already using AI in clinical settings, sometimes opting to use unsanctioned tools like public-facing LLMs rather than tools approved for use in health and care, often without understanding their liability exposure, and without adequate training and organisational support. Participants observed that regulatory guidance for professionals has not kept pace, and were concerned that there may be critical gaps, or at least uncertainties, in the legal framework.

1. **A regulatory guidance vacuum for professionals using AI**

The most consistent findings across both days were that professional regulators are yet to produce guidance on AI that is both specific enough to be actionable and sufficiently broad to cover the range of tools now in clinical use.

Professionals operating today have little to guide them beyond general principles that were not designed with AI in mind, and existing guidance was described as high-level and generic. The General Medical Council (GMC) has recently published a resource regarding how to apply its professional standards to AI and innovative technologies, though this does not constitute formal guidance. Several other major regulators are reviewing their standards but few have produced substantive AI-specific material.

The MHRA yellow card scheme is a safety monitoring system, which facilitates patients, the public and health and care professionals to report concerns about medical devices. However, it was cited as requiring much wider awareness as a route to report quality and safety issues about AI tools. This was supported by the first workshop, which surfaced that people consistently wanted to know how AI was being used in their care and to understand the routes for clinicians and patients to raise concerns.

2. **The legal framework has failed to keep pace with AI**

Workshop attendees considered that clinical negligence principles remain applicable to human-in-the-loop AI use but break down for agentic and autonomous systems. Furthermore, that these principles are undermined in situations where the professional lacks the meaningful ability to evaluate or challenge the AI output owing to deskilling, automation biases or lack of capacity to critically attend to the output. A clear message regarding these risks emerged at the workshops: that they are "foreseeable" and that there is a collective responsibility to mitigate them.

> "*There is a foreseeable atrophy of vigilance when something is proven to be reliable. We know it [an error] is going to happen, therefore we can't place the blame – we can't say 'only use it this way' and then walk off and say 'now it has nothing to do with us'*"
>
> – participant from workshop two

Product liability in the UK is another route to establishing liability when a product is defective and the defect caused harm to a person. However, product liability law only applies to physical goods, so would only apply to AI systems if they were incorporated into a tangible product, rather than systems that operate as pure software. The UK Jurisdiction Taskforce acknowledged this and takes the view that, even without the 'strict' liability' relating to products, the existing civil law in relation to both negligence and contracts provide adequate legal routes for establishing liability and redress.

Beyond liability, participants in workshop two highlighted indemnity arrangements as another area of uncertainty and a potential gap relevant to professional practice and its regulation. For those working in NHS settings, the extent to which NHS Resolution's indemnity schemes cover AI-related harm is being addressed through guidance under development. For professionals working in private settings, standard indemnity policies may exclude IT system failures which could include AI applications. Both of these should be resolved to ensure full coverage.

The “liability sink” problem was a prevailing theme across all workshops. Delegates did not reach consensus on reliable legal mechanisms to hold the “right” party accountable when AI causes harm, and expressed concern that individual clinicians may become "liability sinks" by facing the risk of absorbing consequences they had little meaningful power to prevent. Very few health and care professionals have a significant role in procurement, training data validation or ongoing monitoring of the AI tools that they use, so their assumption of liability was deemed to be unjust. Furthermore, concerns were raised about the speed at which AI tools may be remotely updated, without the professional (or the institution) being aware. This raises questions about protocols for change management and the ability to trace any adverse events to a specific version of the device. Contracts between the product manufacturer and the healthcare provider were described as almost universally pushing liability onto the user, in this case the deploying organisation and the individual professionals using the tool, prompting discussion about fairness and benefits sharing.

> “*How can it be fair that contractual liability for harm might pass from the manufacturer to the organisation, then if the organisation can say “Well, we’ve tried, we’ve made training available and the implementation plan is fully implemented and so it’s the clinician’s fault if it goes wrong.” How can that be?”*
>
> “*Benefits need to be shared. We need to think about a fair outcome and also in terms of developers being properly incentivised to create safe products without overburdening their liability. But, if it’s not the developer, it will be the NHS incurring that liability and that will be the UK taxpayer paying for this [claims], which in certain contexts will not be an optimal outcome. So, it’s about building proper incentives instead of allowing liability shields that push it down in a contract.*”
>
> – participants from workshops two and three

The UK Jurisdiction Taskforce’s legal statement on liability for AI harms advises that that the standard here continues to be what would be expected from a competent member of that profession. Therefore, a professional may be found negligent not only for using AI inappropriately, but also for failing to use AI where a competent member of their profession would have done so. This formalises concerns that workshops surfaced regarding whether professionals would face potential liability, but also highlights a responsibility to engage with the technology where this is becoming a standard of care.

3. **Governance and procurement practices lack consistent application.**

Participants in both workshops two and three expressed concern that governance frameworks for AI procurement were seldom applied consistently or effectively. For example, standard clauses for NHS IT procurement were cited as unfit for AI systems. Without specific clauses being consistently applied in contracts for AI systems, providers will lack adequate clarity and contractual protections around change-control, algorithmic drift, post-market surveillance and monitoring, and liability. These matter in cases where AI systems have contributed to harm as a result of remote updates or performance drift.

Procurement of AI systems was described by attendees as contributing to system risk. They expressed concern that procurement decisions were too often driven by commercial pressure and poorly substantiated productivity claims rather than a robust assessment of need and benefit. Several participants noted that published research on AI productivity gains is based predominantly on self-reporting by AI companies, lacks rigorous evaluation, and frequently fails to account for implementation costs and the time clinicians spend checking AI outputs. Better procurement processes that enable potential purchasers of AI systems to predict the performance of a given tool in situ would resolve this market asymmetry and better assure the quality and efficacy of systems once implemented.

A specific call was made for NHS procurement standards to be strengthened so that AI tools cannot be bought by NHS organisations unless they meet defined safety criteria, including mandatory training requirements, communication protocols for updates and safety alerts, and contractual terms that “fairly” establish liability. The point was made that this approach empowers the NHS as a consumer rather than relying on individual professionals to be the last line of defence against a tool they did not choose and cannot evaluate.

Clinical risk management standards DCB0129 and DCB0160 were mentioned across both days as useful frameworks but were reported to be seldom implemented with full compliance and poorly enforced even where mandated. The DCB0160 compliance gap was seen as an active risk because tools are being deployed without the clinical risk management assessments that are supposed to govern them.

DCB0129 and DCB0160: what are they and which organisations do they apply to?

DCB0129 and DCB0160 are both digital clinical standards, developed by NHS England. They are mandated in England and adopted as best practice in Wales. Scotland and Northern Ireland operate their own governance frameworks.

DCB0129 applies to manufacturers of health IT software and sets forth requirements to promote and ensure the application of clinical risk management with regard to the development and maintenance of health IT systems.

DCB0160 applies to health and care organisations and providers responsible for deployment, use, maintenance and decommissioning of health IT systems. The standard supports organisations to assure the clinical safety of the health IT software that they use.

Small- and medium-sized providers, care homes, domiciliary care providers and sole-practitioner registrants were cited as having little or no infrastructure to implement governance frameworks even where these exist. Infrastructure to support appropriate governance and procurement of these tools may include named AI or digital governance leads, clinical safety officers fulfilling DCB0129 and DCB0160 obligations, procurement teams with sufficient legal and technical expertise to interrogate AI contracts, AI procurement policies, training requirements, incident reporting mechanisms and real-world monitoring capabilities.

Some large providers may be adequately equipped to implement these functions and governance frameworks. However, participants suggested that these organisations are the minority, and that many more providers lack the capability to consistently implement and maintain this infrastructure, creating wide variations between providers and undermining the responsible, robust use of AI across the health and care system.

"*As a patient, I don't care whether the mistake was made by a human or an algorithm. What matters is whether the harm could have been prevented. The focus should therefore be on designing systems where AI and clinicians act as safety nets for one another, rather than creating new opportunities for errors to be missed.*"
– participant from workshop one

4. **The AI education and literacy gap is already surfacing clinical risk**

A consistent and urgent finding across both days was that health and care practitioners are actively using AI tools in clinical settings without adequate training, understanding, or organisational support. This includes the inappropriate use of publicly available LLMs. Examples cited include doctors using consumer applications to interpret diagnostic images, a multidisciplinary team uploading case notes to a consumer version of an AI tool rather than an enterprise version, professionals using browser versions of ambient scribes without knowing where the data is stored, and pharmacists being asked to use tools that challenge the professional duty to check every prescription.

Attendees raised concerns about what they termed the "expertise paradox". This is when an AI system is deployed in a clinical setting because it is superior to a human in certain tasks, yet the human professional remains the arbiter who must critically review the output and make the decision to accept or reject it. This paradox is exacerbated by what attendees considered to be a foreseeable atrophy of vigilance, skill and courage to challenge outputs over time as professionals become habituated to reliable AI technologies. Therefore, though existing clinical negligence principles remain applicable to human-in-the-loop AI use in theory, there is a foreseeable risk that this will break down in situations where the clinician lacks the meaningful ability to challenge the AI output owing to a lack of capacity or attenuated capability. One attendee described the risk that clinicians become "moral crumple zones" and assume liability for harm associated with risks across the system, including scarce capacity and poor implementation.

Deskilling and automation bias emerged as key concerns across each workshop, both of which attendees framed as organisational concerns. Health and care practitioners must remain able and equipped with the right skills to make decisions about the patient in their care in the absence of the AI tool such that they can effectively maintain their role in the loop. However, the passive role of reviewing versus actively reasoning and analysing was raised as a risk of deskilling.

> "*You scan rather than actually reading, and that might be where we're kept busy, because you have to be mentally engaged, whereas if you're actually recording, you're thinking about your plan, you're documenting your plan.*"
>
> – participant from workshop three, talking about the use of ambient scribes

Analogies to aviation were presented as instructive by attendees; safety agencies recommend incorporating manual flying into training programmes, and simulator scenarios are designed to test a pilot's situational awareness. These were suggested as models for what professional training in an AI-enabled environment may look like in the future in order that practitioners are able to maintain their skills and the critical ability to challenge a confident, plausible, yet incorrect output.

5. **Equity concerns**

AI tools may perform differently across different patient populations owing to biases and lack of representativeness in training data. The Whitehead Review of Equity in Medical Devices identified the "real danger" that innovations in AI as a Medical Device (AIaMD) will exacerbate inequity, and urged coordinated action across the device lifecycle to identify and eliminate bias. This is an area of active research, including a number of projects being funded by CERSI-AI. For example researchers have developed a methodology to stress test AI systems and identify where they may perform differently between subgroups of the patient population. These types of approaches are not yet common-place. Professionals may therefore be ill-equipped to know whether the AI tool they are about to use will perform appropriately for the specific patient in front of them, since performance data will often be based on 'average' performance without consideration of potential underperformance in certain groups, for example relating to age, sex or ethnicity. Questions may therefore emerge about whether health and care practitioners have enough information to meet their existing consent obligations, including whether the materiality principles established in the Montgomery test of consent extend to disclosure of medical device performance.

Patients have no reliable way to know when AI has been used in their care. The professional duty of candour requires openness and honesty with patients when something goes wrong. However, the absence of specific transparency obligations around AI use and the lack of consistent auditing of AI systems within organisations means that patients may never know that an AI tool contributed to a safety incident in the first place, nor have a meaningful route to seek redress. Patient experiences of, and response to, harm in healthcare already show patterns of inequity, emphasising the need to take an equity-centred approach to governing liability and empowering patients.

"*As a professional using an AI tool, are you able to tell your patient in front of you – 'Okay, so these tools have been found to perform well for your ethnic group' or 'your ethnic group was represented in 15% of training data, so it might not necessarily be the best in terms of outcomes for you.' Regulators need to be writing high level statements that address whether professionals need to have this information at their fingertips when deploying care. That's a gap - when you tell people, 'Yeah, you have the liability for using AI', you need to create critical professionals who also think around bias when it goes to asking questions*."

– participant from workshop two

OPPORTUNITIES FOR COLLABORATION AND COORDINATION

The workshops revealed appetite for joint working on AI across the regulatory community, and action to collaborate has been taken by the PSA as well as independently, though this has mostly been informal. The absence of formal, systematic collaboration to this point, partly stemming from challenges presented by different legislative frameworks, reveals missed opportunities to optimise resources and effort, and to collectively respond to issues encountered by professionals who are often working in concert. Precedents, including the joint commitment on the Duty of Candour, the joint regulators statement on Using AI in Health and Care Professional Education and the Digital Regulation Cooperation Forum, provide workable models.

There was a strong, cross-day consensus that collaboration between professional regulators offers specific benefits including:

- A collective voice for developers and policy-makers to hear and respond to that no individual regulator could produce alone
- Consistency of standards for professionals working in multidisciplinary teams (MDTs) who are currently subject to different expectations depending on their registration
- Avoidance of duplication of effort across multiple regulators consulting separately on similar issues
- A clearer signal to patients and the public that regulators are acting coherently, which may better engender trust in the system as a whole
- Leverage for the insurance and indemnity sector. A shared ethical framework makes it easier to define coverage conditions and reduces the risk of professionals unknowingly operating outside their cover.

The PSA described their efforts to conceptualise and address the challenges relating to fragmentation in regulation across professions. Their objective is to support a cohesive and dynamic approach to regulating health and care professionals for AI use, while balancing the role of organisational autonomy and complementing existing frameworks of standards and guidance set by regulators and Accredited Registers (ARs).

The PSA set out four possible options for collaboration to the attendees of workshops two and three, to be pursued either individually or in parallel. The table below presents these options in order of most supported to least supported by workshop attendees and includes a description alongside the advantages and disadvantages surfaced in discussions with workshop attendees.

Option	Description	Advantages	Disadvantages
High-level principles for regulators and ARs.	Owned and coordinated by the PSA, working with MHRA and professional regulators. A shared set of principles applicable across all regulated professions and ARs, serving as a foundation to inform guidance from each regulator. This would draw on research such as by Smith & Ives.	Allows regulators to retain autonomy and profession-specific expertise. The abstraction of the principles would secure a longevity that more prescriptive rules would lack. The PSA's oversight role would provide a structural mechanism to require regulators to demonstrate alignment.	There is a risk of abstraction without practical action; principles set at a level that is too high to be meaningful would not be useful. Attendees suggested that this option must include worked examples to have utility. This would not address the immediate guidance vacuum.
Common ethical guidance for professionals across regulators and ARs.	Owned jointly by all professional regulators and ARs, but coordinated through the PSA, this would provide ethical guidance applicable to all professionals regardless of registration.	This would provide consistency to professionals working in MDTs who currently follow different regulations, and to the patients receiving care. Strongly supported in principle at each workshop and was cited as a powerful instrument to signal collectively to innovators and government on minimum standards.	This would be constitutionally complex because individual regulators are legally accountable for their own standards. The diversity of regulated professions would result in trade-offs between specificity and relevance. The time and investment from different organisations required to produce this would be costly.
Stand-alone AI specific guidance from individual regulators and ARs.	This would involve each individual regulator or AR developing their own AI-specific guidance, position or standards tailored for their own registered professionals. This approach has already been initiated by several regulators.	Already underway, and allows for specificity and context that high-level principles could not capture, which signal intent and reassurance to professionals in the near-term. A useful way for regulators and ARs to translate higher-level principles into profession-specific practice.	If pursued in isolation, this approach could exacerbate fragmentation and inconsistency for professionals working in MDTs. This could duplicate effort across multiple regulators who may consult separately on the same questions.
Iterative updates to existing standards and guidance.	Owned by each individual regulator as part of their standard review cycles, existing professional guidance would be updated over time to incorporate references to AI.	This process is familiar and will likely happen within existing regulatory frameworks as case law, complaints and consultations shape the broader context over time. This would ideally be additionally supported by higher-level principles.	This was considered to be inadequate alone because relying on regulatory development cycles over long time periods would lack the proactivity and timeliness that the risks and opportunities of AI demands. Without a shared framework, these iterative updates could worsen inconsistency.

RECOMMENDATIONS
SHORT TERM

The following are actions identified through the workshops that could be taken quickly, within existing powers:

- **Regulators should publicly communicate what they are currently doing on AI**. This is the single most universally requested action across all workshops. A published statement could provide meaningful reassurance to professionals and indemnity providers, and would represent an act of regulatory transparency. The current lack of coordinated communication was argued to be itself harmful, potentially fuelling a perception that regulators are not engaged. This communication could acknowledge uncertainty while demonstrating intent, rather than being definitive.

- **A joint statement from professional regulators.** The PSA was identified as a natural convener for this joint statement, which needn't be a set of principles but could instead be a signal of collective intent and a commitment to cross-regulatory alignment. The Duty of Candour joint commitment was cited across multiple tables as a useful model to follow.

- **Clarity for professionals about their current liability exposure when using unsanctioned AI tools**. Professionals using direct-to-consumer AI tools with patient data need to be warned clearly, specifically, and through their professional bodies and defence organisations that they are operating outside their indemnity cover and at professional risk.

- **Minimum AI governance standards embedded in health service inspection frameworks**. Attendees suggested that providers should be assessed on whether they have AI policies, training requirements, and monitoring arrangements in place. This creates a universal floor of governance accountability regardless of provider size or type.

- **Tool-specific training as a condition of use.** Any professional using an AI system as part of a patient's care should be required to complete training for that specific tool, with documentation, before using it in clinical practice. The model of data protection training was proposed as the template. This does not require new legislation; it could be incorporated into existing workplace governance.

- **An AI statement for each provider, including a transparency register of AI tools**. There was broad consensus that NHS trusts and independent providers of health and social care should publish which AI tools they are deploying and in which clinical settings. This statement could also communicate to patients what they can expect when AI is used in their care, including transparency regarding the use of AI, what information practitioners should be able to provide, how performance is monitored and what happens if things go wrong. This suggestion is strengthened by public polling finding that between 75% and 79% of patients believe that they should be told in advance about AI being used to support diagnostics and triage. Further discussion should focus on the type of application that would be relevant; for instance, AIaMD and AI tools used in consultations with patients were deemed to be an appropriate subject of such a register, but AI tools used to support administrative functions including HR and training were suggested to be less relevant. This action could build patient awareness and trust, could help professionals understand what is regulated and how, and may support a more open approach to market navigation and post-market monitoring. This would mirror a similar transparency obligation placed on public sector bodies in the UK.

> "*Patients would want reassurance that the technology was properly tested and continuously monitored before being used in real clinical settings.*"
>
> "*Patients need clear accountability pathways and routes to challenge decisions*."
>
> – participants from workshop one

- **Model contract clauses**. The use of standard contact clauses was understood to be unsuitable for the procurement of AI systems in health and care, and attendees expressed enthusiasm for model contractual clauses to be developed and shared for any organisation to adopt freely. This would mirror the approach that the European Commission has taken to produce AI Model Contractual Clauses for public organisations, which have been published for high-risk and non-high-risk AI systems.

RECOMMENDATIONS

LONG TERM

The workshops also surfaced longer-term actions that could support the resolution of the issues raised in this report:

- **A coordinated professional regulatory AI framework setting out high-level principles**. This was the primary ambition of the workshop series, to be owned and coordinated by the PSA. Participants described this framework as ideally being: applicable across all regulated professions without requiring each regulator to start from scratch; future-proof through abstraction rather than prescription; accompanied by worked examples and case studies that ground the principles in practice; reviewed regularly with a published cycle; and jointly endorsed by the PSA, MHRA and professional regulators.

- **A national AI literacy training programme for health and care professionals**. Unlike the immediate quick win of tool-specific training, this long-term ambition involves embedding AI literacy in education across all health and care professions. This should include critical engagement with AI outputs, understanding of bias and training data, and the ability to identify when an AI output may be unreliable or incorrect. The Medical Licensing Assessment was proposed as one vehicle for setting a minimum standard in medicine; equivalent mechanisms exist or should be developed across other professions.

- **Reform of UK product liability law to encompass AI software**. AI software that is not incorporated into a tangible product does not currently meet the definition of “product” and therefore falls outside the scope of product liability law in the UK. Workshop attendees raised the need for greater integration across the legislative framework, and proposed that the UK follow the EU’s approach of bringing software & AI in scope. The Law Commission's current consultation on the Consumer Protection Act may result in legislative change that achieves this.

- **Mandatory safety cases for AI tools deployed in health and care**. Drawing on practice from nuclear, defence and aviation, health and care organisations that deploy AIaMD and other AI tools used in consultations should be required to develop and publish structured safety cases. It is noted that these are required within DCB0129 and DCB0160, although as noted earlier these are not universally implemented. These safety cases would make explicit the argument, evidence and assumptions underpinning the claim that a tool is safe for its intended use in its intended population. This was identified as both a transparency mechanism and a learning tool.

- **Mandatory co-design of AI tools used in health and care**. AI tools should not be deployed in NHS settings without documented co-design with clinical professionals and patients. This addresses both the adoption gap (clinicians do not use tools they had no role in designing) and the equity gap (tools trained on non-representative data that are then used on populations they were not validated for). This point was raised across both days and echoes an earlier recommendation by the Whitehead Review.

- **Resolved indemnity arrangements for AI use**. The insurance and indemnity sector, NHS Resolution, defence organisations, professional regulators and others need to reach a shared understanding of how accountability, liability and indemnity apply when AI is involved in clinical decision-making.

AREAS OF NON-CONSENSUS

The workshops also revealed areas where no consensus exists and where further work is needed before conclusions can be drawn:

- **The severity and urgency of current liability risk**. Workshop delegates agreed that clinicians face disproportionate liability exposure when AI causes harm, but differed on their perception of how great this risk is. Some workshop two participants, drawn from defence and insurance, expressed alarm at professionals operating in what they considered to be a liability vacuum with inadequate cover. They suggested that the window for proactive action is narrowing; professionals are being exposed to liability for which little support infrastructure exists, and regulation is in a reactive position to harm that is occurring now. Participants in workshop three, drawn from regulators, expressed less immediate concerns in their discussions, noting that existing standards provide a foundation and that regulatory frameworks are more resilient than is sometimes claimed. Though aware of the need to act swiftly, they acknowledged that developing principles through proper consultation takes time and that premature guidance can create as many problems as it solves. This difference may partly reflect that the participants in workshop two are likely to see the downstream consequences of things going wrong, and feel personally and imminently exposed, whereas participants from workshop three were focused on upstream standard-setting and may have a longer-term, more macro view.

- **The right for professionals to opt out of AI tools.** Some participants saw the right to opt out on grounds of competence, conscience, or professional obligation as vital. Others argued that if AI represents evidence-based best practice, then refusing to use it becomes professionally unsound in the same way that refusing to follow clinical guidelines would be, and that this position will become untenable as AI tools become the clinical norm and continue to advance in performance. This latter view appears to be supported by the findings of the UKJT.

- **Patient consent to AI use.** Discussion in the first and second workshops identified two levels to this: consent to AI being used in the care pathway, and transparency about AI use without requiring explicit consent. There was no consensus on where the line should be drawn, or whether AI in clinical settings requires a different approach to consent than the use of AI tools to support administrative functions. The materiality principle from established consent law was referenced in relation to this tension: whether AI use crosses the materiality threshold for disclosure will depend on context, risk, and patient population, but working out what this means in practice was not resolved.

- **The liability position for agentic AI**. The legal framework for agentic AI that makes decisions without a human in-the-loop is uncertain and will possibly require both legislative intervention and test case law to resolve. One view expressed and debated was that this distinction has been somewhat overworked in the professional regulatory context, and that existing professional standards already cover professionals' use of AI tools at any level of autonomy. Others argued that agentic and autonomous AI presents new accountability challenges that existing standards cannot address without modification.

What is agentic AI?
Agentic AI refers to AI systems that are able to autonomously plan, decide and execute actions to achieve a goal. This is sometimes referred to as “human-on-the-loop” which reflects the human’s role as a supervisor, rather than “human-in-the-loop” which describes an AI system that provides an output for a human to interpret and act upon.

- **Whether AI should be treated as exceptional or technology-neutral.** A fundamental disagreement was whether AI requires a qualitatively different regulatory approach, or whether existing frameworks are sufficient if properly applied. Those who saw AI as qualitatively different emphasised its role as an "epistemic agent" that disrupts professional judgement and creates paradoxes of accountability. Those who saw it as technology-neutral argued that the professional obligation to use tools competently and within their scope applies regardless of what the tool is.

NEXT STEPS

This work intended to explore the consensus, non-consensus and areas of uncertainty of AI liability across different stakeholder groups, identifying where regulatory guidance on AI for professionals requires further clarity, and how collaboration could resolve those gaps to ensure that AI technologies are adopted and governed safely, equitably and effectively.

Following the workshops, commitments have been made by key organisations and initiatives are underway that respond to the risks and issues surfaced:

- The Professional Standards Authority are coordinating professional regulators and ARs to issue a joint statement of intent
- The Professional Standards Authority are coordinating professional regulators and ARs to develop a high level principles framework as endorsed by the workshop attendees
- CERSI-AI will work with partners to produce draft model clauses to share with providers to support greater clarity regarding ownership of risks relating to AI health technology deployment
- The National Commission into the Regulation of AI in Healthcare will publish its recommendations on 10th September 2026, with the response from government expected soon afterwards

Glossary

Professional regulators are responsible for ensuring that the professionals working in UK health and social care settings who must be regulated by law are appropriately trained and qualified. This includes doctors, nurses, paramedics, dentists, pharmacists, physiotherapists, chiropractors, and osteopaths.

Accredited Registers enable other practitioners working in health and care roles that are not regulated by law, including practitioners that deliver talking therapies and non-surgical cosmetic procedures, to voluntarily register for accreditation and commit to maintaining standards of good practice.

The Professional Standards Authority oversees the 10 professional regulators and the Accredited Registers programme.

AIaMD describes an application of AI in healthcare that is used for a medical purpose, including the diagnosis, prevention or treatment of injury, the replacement or modification of anatomy of physiological process, or the control of contraception. All medical devices, including AIaMD, are regulated by the MHRA.

The Bolam Test is a legal standard used in the UK to decide if a doctor or professional was negligent, established in case law by *Bolam v Friern Hospital Management Committee [1957]*. It provides that a professional is not guilty of negligence if they act in accordance with a responsible body of medical opinion.

The Bolitho Test clarified Bolam by stipulating that a professional's action is not negligent if that responsible medical opinion is found to have a logical basis by the court. This was established by *Bolitho v City and Hackney HA [1998]*.

The Montgomery Test established in *Montgomery v Lanarkshire Health Board [2015]* refines negligence law further, placing a duty on the practitioner to take reasonable care to disclose information to the patient on the material risks involved in any recommended treatment and of any reasonable alternative or variant treatments. Reflecting a shift toward patient-centred care, the judgment requires a test of materiality where a reasonable person in the patient's position would be likely to attach significance to the risk. With regard to AI, the UK Jurisdiction Taskforce suggested that in future negligence cases, attention may be drawn to the reasonable use and operation of AI as well as transparency regarding the risks of using or not using AI for that patient.

REFERENCES

‘AI to speed up cancer diagnosis for millions of NHS patients’ (Department of Health and Social Care, June 2026)

‘Artificial intelligence and innovative technologies’ (General Medical Council)

Binesmael A and others, ‘Attitudes to technology and AI in health care’ (The Health Foundation, March 2026)

Campbell D, ‘Doctors and NHS could be sued for mistakes made by AI tools, report warns’ (The Guardian, June 2026)

Crocker H and others, ‘*Patient-reported harm from NHS treatment or care, or the lack of access to care: a cross-sectional survey of general population prevalence, impact and responses*’ [2026] 35(1) BMJ Quality & Safety

‘Fit for the future: 10 Year Health Plan for England’ (NHS England, July 2025)

‘Joint statement from statutory regulators of health and care professionals: using AI in health and care professional education’ (Health and Care Professions Council, February 2026)

‘Joint statement from the Chief Executives of statutory regulators of healthcare professionals’ (PSNI, 2024)

Lawton T and others, ‘Clinicians risk becoming “liability sinks” for artificial intelligence’ [2024] 11(1) Future Healthcare Journal

‘Legal statement on Liability for AI Harms under the private law of England and Wales’ (UK Jurisdiction Taskforce, July 2026)

‘Model stress testing using causal generative AI’ (CERSI-AI)

‘Report on the research and engagement programme for the National Commission into the Regulation of AI in Healthcare’ (Medicines and Healthcare products Regulatory Agency, July 2026)

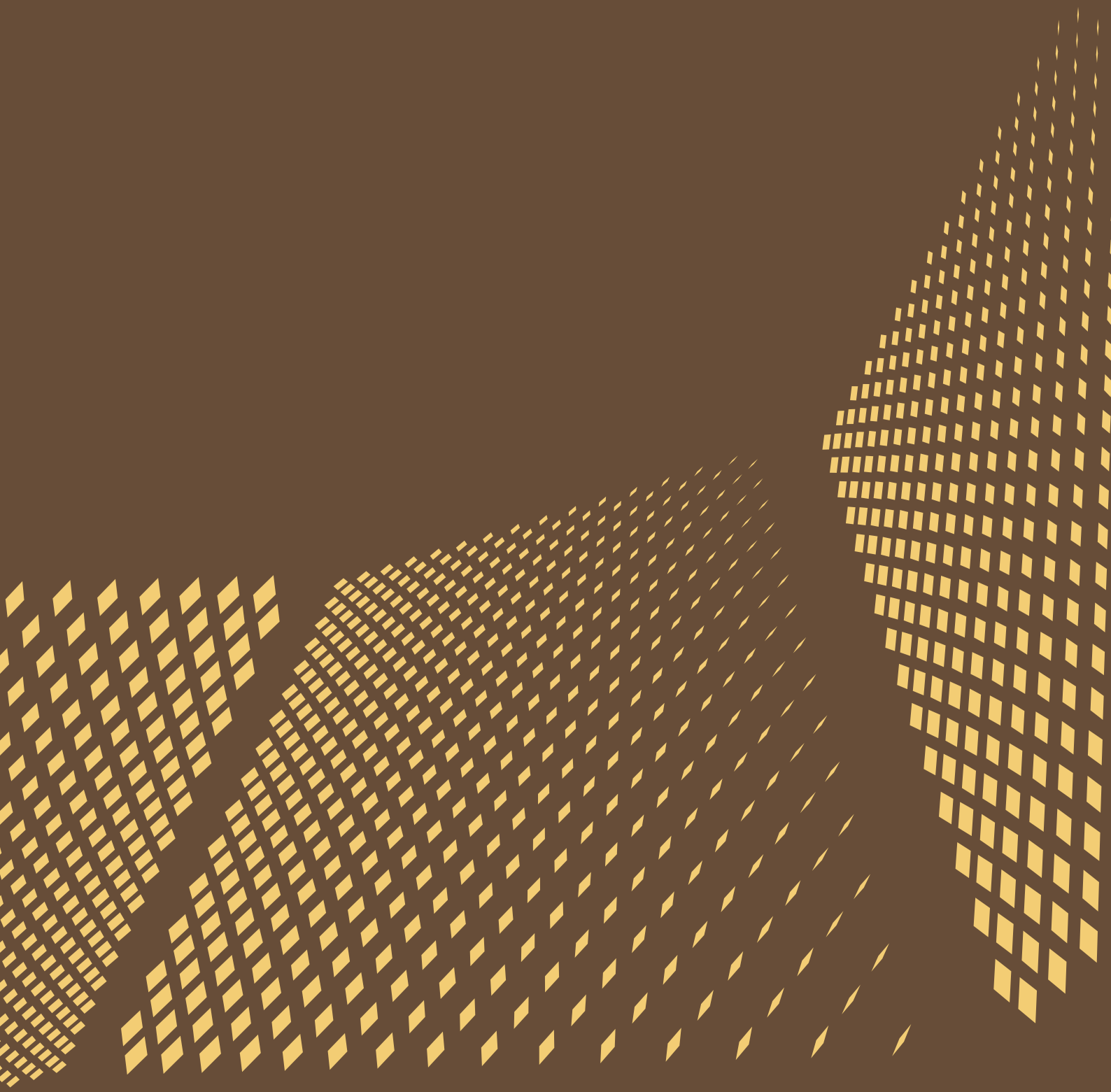
Smith H, Downer J and Ives J, ‘Clinicians and AI use: where is the professional guidance?’ [2024] 50(4) J Med Ethics

‘The use of AI in UK healthcare: public perceptions and healthcare priorities’ (King’s College London, May 2026)

Toro-Tobon D and others ‘Clinician in the loop: a flawed solution for AI oversight’ [2026] 393 BMJ

‘Updated EU AI model contractual clauses’ (European Commission, March 2025)

Whitehead M, ‘Equity in medical devices: independent review’ (Department for Health and Social Care, March 2024)



CERSI AI