Workshop report

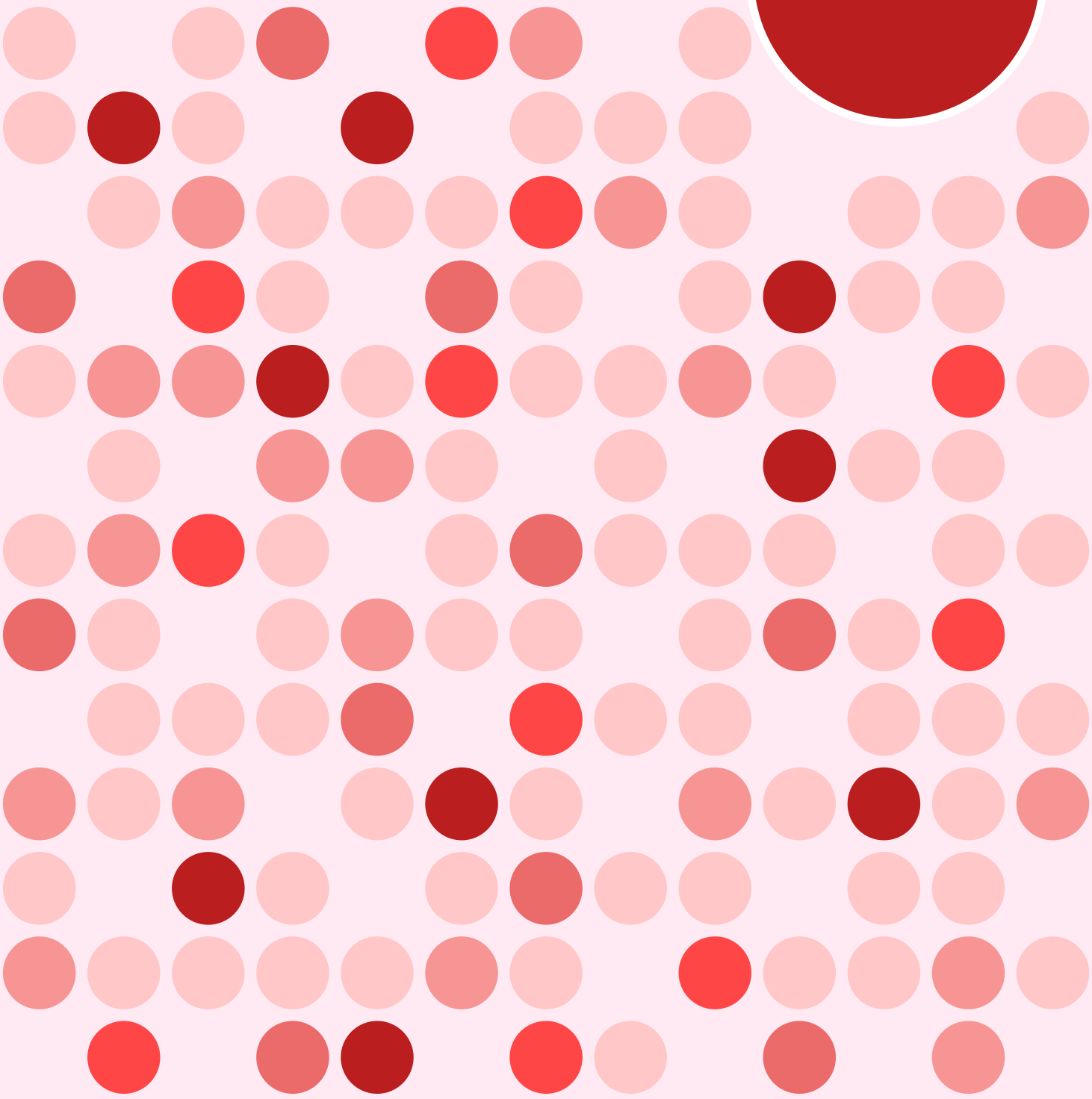
professional standards authority
for Health and Social Care

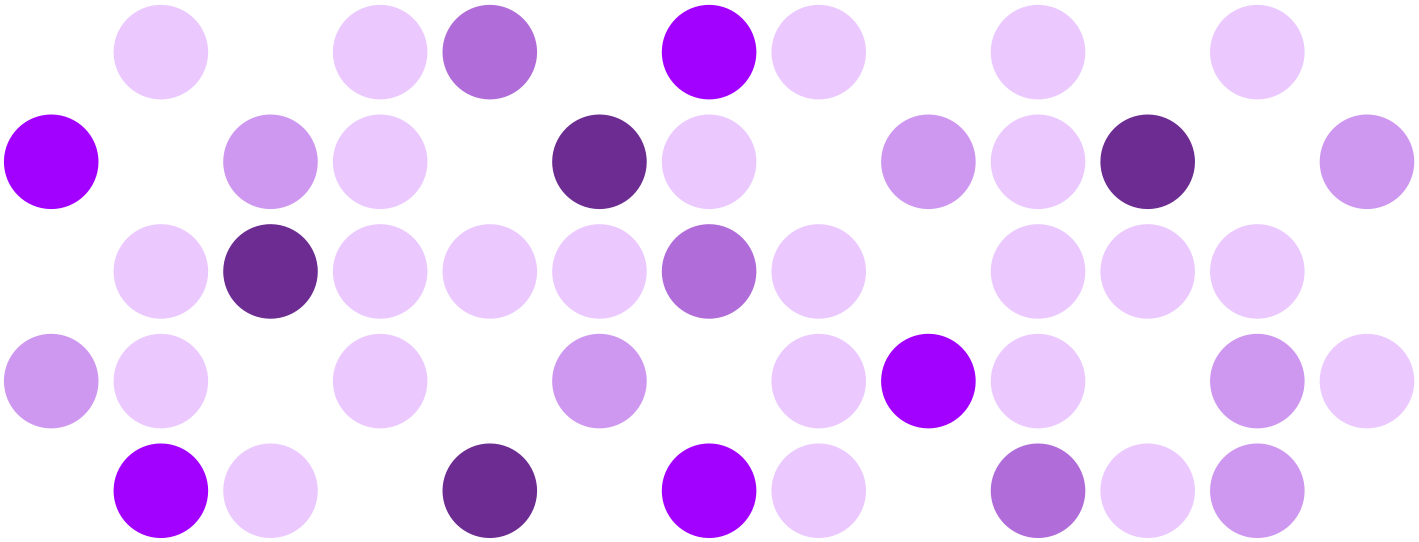
University of BRISTOL
Centre for Ethics in Medicine

Artificial Intelligence

How to guide and regulate for health and social care professionals who use AI

2026





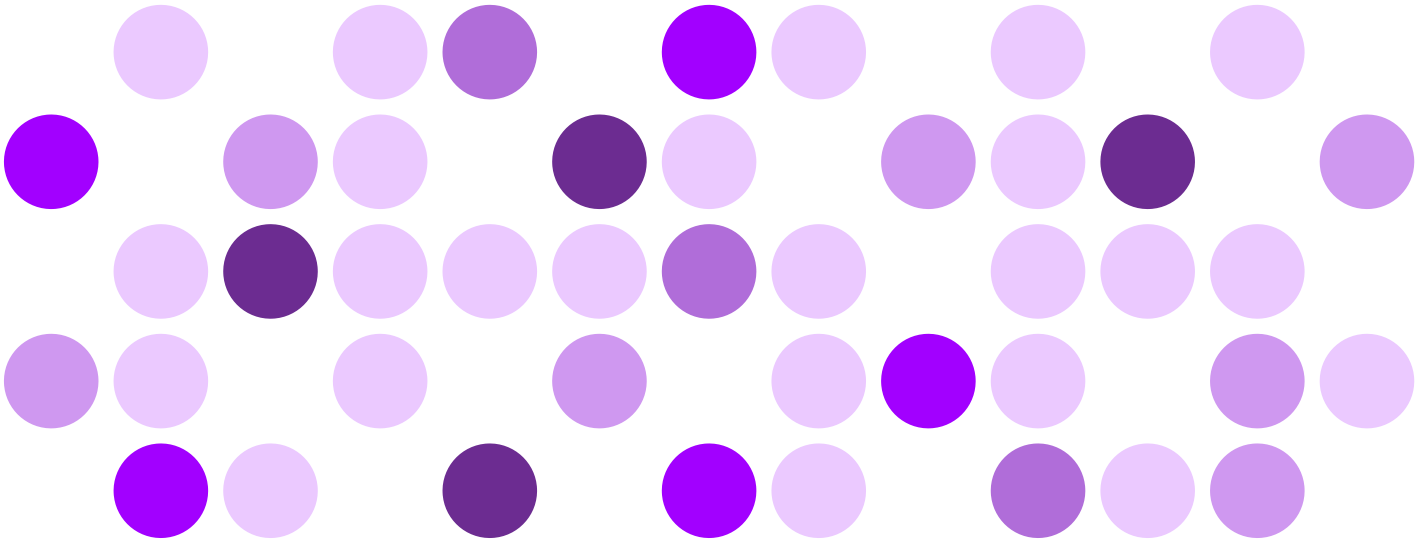
About the Professional Standards Authority

The Professional Standards Authority for Health and Social Care (PSA) is the UK's oversight body for the regulation of people working in health and social care. Our statutory remit, independence and expertise underpin our commitment to the safety of patients and service-users, and to the protection of the public.

There are 10 organisations that regulate health professionals in the UK and social workers in England by law. We audit their performance and review their decisions on practitioners' fitness to practise. We also accredit and set standards for organisations holding registers of health and care practitioners not regulated by law.

We collaborate with all of these organisations to improve standards. We share good practice, knowledge and our right-touch regulation expertise. We also conduct and promote research on regulation. We monitor policy developments in the UK and internationally, providing guidance to governments and stakeholders. Through our UK and international consultancy, we share our expertise and broaden our regulatory insights.

Our core values of integrity, transparency, respect, fairness, and teamwork, guide our work. We are accountable to the UK Parliament. More information about our activities and approach is available at www.professionalstandards.org.uk.



Contents of this report

Introduction

A report prepared by

- **University of Bristol: Dr Helen Smith and Professor Jonathan Ives**
- **Professional Standards Authority: Dr Patrick Murphy and Melanie Venables**

following a workshop held on 27 February 2026.

The PSA has an interest in using its convening powers to help explore regulatory responses to emerging risks and opportunities such as those presented by Artificial Intelligence (AI). This also helps the PSA to consider how it should set expectations for the regulators and Accredited Registers it oversees in regards to AI use by professionals.

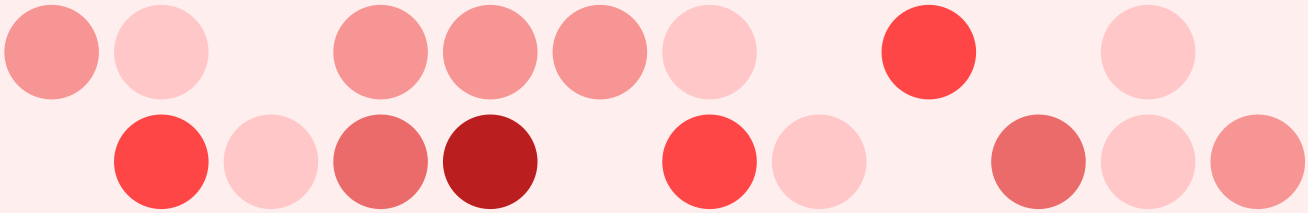
On 27 February 2026 a workshop, commissioned by the PSA, and in collaboration with Dr Helen Smith RN and Professor Jonathan Ives from the University of Bristol, was undertaken to consider professional ethical guidance and regulation for users of AI in health and social care settings in the United Kingdom (UK).

The workshop brought together regulators, Accredited Registers (ARs), service users, service user representatives, and members of the public to discuss the opportunities and challenges of using AI.

Through deliberative group discussions, and with the aid of a series of real-world scenarios, participants explored various themes including AI safety, bias, transparency and accountability.

This report offers an overview of the main points and recommendations from attendees which were raised in oral discussion, documented in written 'Padlet' comments, or typed into the Teams chat-box notes made during the workshop.

We thank each and every attendee for their participation and contribution, which we have endeavoured to faithfully report in this document.



Background

In the context of growing interest in, and anticipated increases in deployment of, AI across health and social care, the workshop was convened to gather perspectives and recommendations on what professional ethical guidance and regulation could do to support safe, appropriate, and fair use of AI in practice, particularly given urgent areas of uncertainty identified in relation to responsibility, oversight, and implementation.

This workshop took place in parallel to the Medicines and Healthcare products Regulatory Agency's (MHRA's) National Commission into the Regulation of Artificial Intelligence (AI) in Healthcare. The purpose of the Commission is to bring together "global AI leaders, clinicians and regulators to advise the MHRA on the development of a new regulatory framework for AI in healthcare", and to "produce recommendations to advise the development of MHRA guidance in the interim, addressing urgent areas of uncertainty for the healthcare system and industry" (MHRA, 2025).

The progressive increase in interest and funding channelled towards the development of AI destined for health and social care has culminated in an aspiration in the NHS 10 Year Plan for England to make the NHS the most AI-enabled health system in the world with AI seamlessly integrated into clinical pathways. There are similar ambitions for AI within the devolved nations. It is hoped that AI will enable staff to be freed up from administrative tasks (using technology such as ambient voice technology and digital triage) to spend more time focusing on patient care (UK Government, 2025).

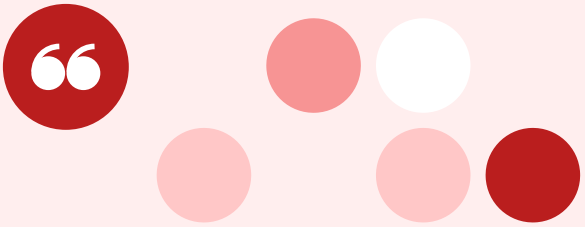
At present, AI cannot operate without human initiation, direction, and oversight. AI entities are not professionally regulated individuals, and their presence and contribution could disrupt the usual decision-making processes of those who use them.

AI deployment is potentially disruptive, provoking a mix of excitement, fear and anxiety, and raises specific professional, ethical, and legal questions for users. For example, to what extent can or should professionals rely on, question, or entirely ignore an AI-generated recommendation? If something were to go wrong, and the AI gives a bad recommendation, then who will be held ethically and legally responsible if that wrong AI recommendation reaches the service user and leads to harm?

At present there is limited guidance available for AI users in health and social care in the UK, and that which is available is directed to a few specific disciplines. Examples include the General Medical Council's (undated) resource on *Artificial intelligence and innovative technologies*, and the updated *Standards of Proficiency for Radiographers* as laid out by the Society of Radiographers (2023).

This compartmentalised approach means that guidance for AI use can be crafted to meet the needs of specific health and social care practices, in line with the requirements of a discipline and profession. However, there are risks associated with a non-uniform approach to AI within wider multidisciplinary health and social care professional groups.

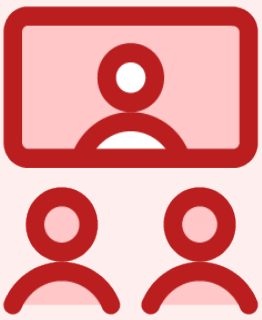
It has been argued that there is scope for some standards to be unified across all clinical groups, no matter which user group is using an AI (Smith, Ives, & Downer, 2024). The aim of a unified approach would be to ensure that all health and social care providers will deliver – and service users will receive – care that is ethically consistent across all services.



"Will we soon be in a time where aspects of AI use are so ubiquitous that we expect and understand that they are in use somewhere through the process?"

The workshop

Workshop participants



17 Service Users

16 Regulators

1 Accredited Registers

4 Other interested Parties

The workshop posed four main questions, identified by the PSA as being of particular interest, these are used to structure the evidence presented in this report:

1. **What should be in place to enable responsible AI adoption?**
2. **Who should be accountable/responsible for any harms caused in the scenarios?**
3. **Who should be responsible for detecting and addressing issues in AI prior to deployment?**
4. **To what extent should the focus be on no-fault learning and correcting of AI users rather than responsibility or blame?**

Aim

The aim of the workshop was to engage with, and gather recommendations from, attendees regarding how to guide and regulate for professionals who use AI in health and social care.

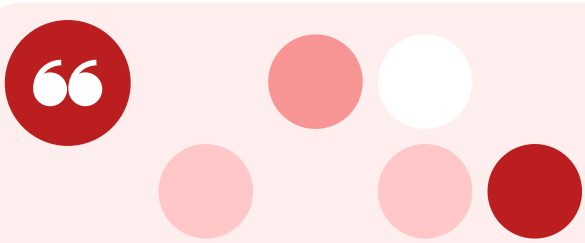
Attendees were invited from a broad range of backgrounds and the workshop was attended by regulators, service users, service user representative organisations, and members of the public. Separate engagement events regarding professional practice were held for AI users (i.e. professional bodies such as the

Royal Colleges) as part of the MHRA Commission's work. Accordingly, AI user professionals were not represented at this workshop to avoid duplication with the same group attending two similar events.

Format

The workshop was arranged online via Microsoft Teams to enable maximal inclusion of attendees and to remove geographical barriers to attendance. The PSA reimbursed public attendees for their time and, where requested, for reasonable expenses, in line with its public involvement policies.

The workshop was attended by service users, regulators and ARs, and otherwise interested parties. In the morning, attendees were allocated into discussion groups according to their affiliation (e.g., regulators, service users). In the afternoon the groups were mixed to enable rich collaborative discussion from different perspectives. Attendees spent the day recursively moving from small group discussions to integrative feedback discussions in the main room. (See Appendix B for the running order for the day.) Online 'Padlets' were used to capture attendees' written thoughts during the small-group work throughout the day. Main-room discussions and written entries in the Teams chat box were recorded. All materials were subsequently analysed and used to produce this report.



"The risk will differ for different AI."

"The agreed AI model needs to be the same across the whole of the UK."

Attendees' names were not associated with written or verbal comments made during the workshop and, as such, names are not connected with quotes in this report. This approach was intended to encourage attendees to share freely throughout the day.

Quotes used in this report have been subjected only to essential editing (for example, to correct spelling and grammar) to aid readability, but have otherwise been left as written or spoken.

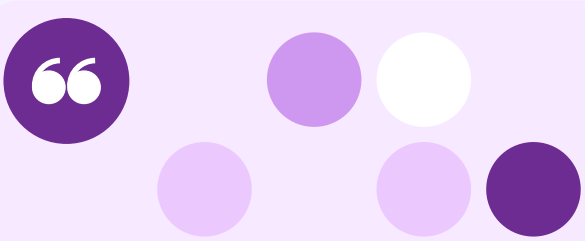
Key findings: a summary

These two pages provide an executive summary to communicate the main findings from the workshop and messages conveyed in this report.

This summary does not comprise recommendations, as further research would be needed on some of the points raised. These points, however, offer a starting point for recommendations about guidance for AI use.

The workshop indicated that there would be further value in exploring:

- **Adopting a risk-based approach**, in line with the PSA's 'right-touch regulation' approach to guidance (Professional Standards Authority, 2025), which distinguishes between different AI technologies and use contexts, balancing proportionality, appropriate specification, professional judgement, and service user safety.Using clear, specific, and consistent terminology in guidance; especially in defining “AI” and “responsibility”.
- There could be **benefit to having cross-profession consistency at a high level**, with specificity at a professional and use-case level, where needed. However, since some professionals are neither subject to statutory regulation, or members of Accredited Registers, it is important that employer requirements and any future legal frameworks for regulating AI use by professionals can ensure protections to the public (no matter how they are accessing health and social care).
- The **starting point for regulatory guidance should be helping registrants apply existing professional standards to AI**, rather than creating wholly new ethical principles.
- **Setting expectations for training as career-long**, with a shared baseline and service-/technology-specific progression, including Continuing Professional Development.
- **Clarify roles** across regulators, professional bodies, and employers in supporting competence for AI use.
- **Strengthening expectations for reporting and learning from AI use**, including clear routes for professionals and service users to raise concerns, and effective collaboration and signposting to relevant post-market surveillance mechanisms.



> “[There needs to be] some level of consistency across how different professions are regulated”

- **Promoting a learning-oriented culture** that supports speaking up, while recognising accountability and associated responsibility may still apply in some cases.
- **Avoiding placing sole responsibility on frontline professionals** for system-level risks; emphasise shared/system accountability to prevent ‘moral crumple zones’ (Elish, 2019).
- **Taking a clear public position that AI used in health and social care should meet appropriate safety and efficacy standards** before deployment, so registrants can reasonably expect this assurance and raise concerns if it is not met.
- **Embracing diversity of input at all points of AI development and adoption**, employing Patient and Public Involvement activities (in addition to professionals, which was taken for granted) to enable detection and avoidance of foreseeable issues arising, and enabling iterative service improvements where AI is used.
- **Maintaining an agile and iterative approach**, with regular review, to enable regulation to keep pace with rapid technological shifts.

Key findings: in-depth

Overall, it was felt ubiquitously that regulation and guidance for AI use is an important area, but presents a complex and multifaceted challenge to construct and implement, especially across multiple different domains.

Discussion acknowledged there is a fine balance to strike between proportionality, specification, allowing professional judgement, and safety.

One problem noted at the start is the lack of a single, clear definition of AI as deployed in health and social care, as different AIs do and will serve different functions in different settings. As such, it may be difficult to consider, design and implement general ethical guidance and standards for AI use that covers all possible uses and settings.

Below, we have laid out each of the four main questions as asked by the PSA, and provided a summary of the discussion, alongside illustrative quotes.

> **“Standards and guidance from regulators needs to complement a clear underpinning legal framework, which is currently missing.”**

What should be in place to enable responsible AI adoption?

Guidance for AI users

The following subsection outlines attendees' input about what guidance should be in place to ensure responsible AI adoption.

Attendees noted that guidance for AI users could be developed in different ways, and that regulators face a complex challenge in finding the right level. Standards may be seen as simultaneously over- or under-specified for registrants working with different technologies and in different settings.

This is compounded by the feeling that 'AI' is not sufficiently well defined, creating a further challenge for regulators in specifying the scope of any guidance. Additionally, it was acknowledged that the pace of AI development poses an additional challenge, requiring regulators to consider both the present and the future.

> "We need to think about the full range of AI tools, from the least to the most complex – but as a result, principles that feel most relevant at the least complex end of the spectrum may not feel equally as relevant or helpful or justifiable for tools at the most advanced end of the spectrum. The pace of technological development is fast and capabilities are growing quickly but adoption at scale could still be further off – how should that impact the approach to guidance in the present?"

> "I think from a regulatory perspective, the term [AI] itself is really unhelpful as it covers multitudes. We need to move to talking about individual technologies and varying the approach according to risk. Trying to write guidance or standards that captures all 'AI' use risks genericising it to the point it stops saying anything."

Some attendees felt that existing high-level ethical guidance for professionals is sufficient to cover the use of AI. However, it was noted that applying these existing standards to AI might be challenging, and that specific guidance may be needed to help professionals operationalise existing ethical principles and guidance in this new context. It was felt that the role of a regulator is to help registrants use AI in line with existing standards, not necessarily to come up with new principles and standards.

> "Feels like this is already covered by existing expectations, do we need to make AI specific expectations here?"

> "As a regulator, we need to focus our support for registrants on using AI in line with our professional standards."

> “It was felt that there is already guidance and accepted professional responsibility. If you are signing off on case notes or producing case notes, they have to be accurate. So it was felt that this was really more about guidance that might help people to ensure they do and can fulfil their professional responsibilities rather than outlining specific or new professional responsibilities in the context of AI. So it's about trying to ensure that people can meet those professional standards. It was noted that... we already use dictaphones, for example... it's fairly standard that people have recorded notes, then give them someone else to write up. The responsibility was still always on the person who dictated notes to check it and [sign] off. So again, there's not really any completely new problems there.” **Facilitator feedback to main room from small group session 2, group 5 (scenario 5)**

It was felt that regulators need streamlined national frameworks to support them in setting any guidance. It was noted, however, that this may be challenging when individual organisations need to work within their own legislative responsibilities.

> “Standards and guidance from regulators needs to complement a clear underpinning legal framework, which is currently missing.”

> “National standards [are needed], which seem to go against the prevailing mood of [NHS] Trusts adopting what standards they want.”

Whilst a high-level approach, with cross-regulatory consistency, was broadly endorsed, it was noted that there would also need to be more specific guidance to accommodate specific AI use cases. It was felt that guidance requirements may need to depend on the risk profile of specific technologies. Attendees also noted that guidance requirements may depend on the level and extent of training provided to professionals.

> “[There needs to be] some level of consistency across how different professions are regulated.”

> “The agreed AI model needs to be the same across the whole of the UK.”

> “But the cost reduction in training and adaptation with be minimised if the same thing is adopted over the whole of the UK (not just England).”

> “The risk will differ for different AI.”

> “Reflexive relationship between guidance and training – what might be reasonable to expect of individual professionals will be impacted in part by what AI training and education looks like for those professionals throughout their careers.”

Whilst this workshop was focused on AI users, there was much discussion about AI itself, and consideration given to how regulators can or should work to ensure AI technologies in the health and social care space are responsibly developed prior to it being adopted (for example, tackling the risk of bias, preventing the exacerbation of inequalities, giving consideration to those in vulnerable and/or disempowered positions).

> "We already know there are huge existing inequalities in Health + Care – How do we regulate AI use so that as a minimum it doesn't widen inequalities + at best it reduces those existing inequalities? If the AI research data that feeds the policies and regulations for practice does not hear the voices that really need to be heard (that experience the greatest healthcare inequalities) that's where harm can come in and inequality gaps in healthcare provision can widen. This workshop is about regulation and practice. I'm very concerned about the voices that are unheard when it comes to AI research that feeds into all of this."

> "And then you have your hierarchies of academic evidence too – often the research that represents the voices of those who are not often heard is qualitative community research (storytelling etc) and way down that hierarchy of evidence."

> "Just to add to group 3 as well, we also discussed the huge complexities and concerns raised when it comes to people admitted to mental health wards/ detained under the Mental Health Act. People are already not being asked consent for many "small" decisions (who decides this?), how do we safeguard against AI being added to this?"

The risk of creating or exacerbating inequalities through AI in health and social care arise both from use in practice and from its development, and requires special consideration.

Whilst there may be limited opportunity for professional regulators to have an impact on development, it may be feasible for them to take an explicit position on the need for responsible development, and to offer specific guidance around how it is used in the context of caring for vulnerable populations (for example, service users detained under the Mental Health Act 1983).

> "Should consider adding guidance in the Mental Health Act code of practice handbook which is regularly used as a reference point in [mental health] services."

Training for AI users

The following subsection lays out attendees' deliberations surrounding the training needs for AI users.

Attendees felt strongly that professionals need to have training for AI use that spans their entire career, with a basic shared level of understanding at qualification, progressing to higher-level training specific to the service that they provide.

> "Should be trained and competent for whatever they do."

> "High level of AI literacy required to support all AI use. Therefore training along the lines of 'evidence-based medicine' needs to be built into undergraduate and postgraduate training/CPD."

> "Need to see training as career-long, rather than static. need ways to check people are keeping up to date, such as through mandated CPD."

> "Potential for revalidation to be a tool to address issues in AI competence"

> "And protected time to enable them to actually attend and complete the training, in work time, not outside work time!!"

It was felt that training needs to be mandatory, informative, engaging, and broad, for example covering the limitations of AI, and using case studies to show how general principles and standards can be applied to specific uses of AI. It was noted that training needs to build confidence in students, and be accessible to all.

> "Online and mandatory, on bias, simple usage, examples good and bad. At present, NHS staff in trust I train have just been allowed to look at, but not use, copilot...needs more than this but not aim to make them experts. "

> "Case studies and practical examples which is profession-specific will be helpful in addition as AI expansion happens and we have generations which are very unused to it as well as those who are."

> "Should be a part of mandatory training in any service where AI is used (ideally for all people working in health/care settings) to make sure we all have a basic shared level of understanding. And then potentially higher-level training on specific service use/tech use?"

> "The biggest challenge as regulators is how we ensure that training is providing practitioners the confidence and competence to use AI appropriately and safely (and to understand the warning signs that it is not). That underpins anything else that we can put in place."

Attendees identified three broad levels of capability that training needs to address. First, the standards of proficiency informed by the AI user's regulator who oversees the general required standard of practice; second, the AI user's professional body who oversees their specific area of practice; and third, the AI user's employer.

> "Good to have standards of proficiency for every profession for their area of expertise – what do we expect the professional to know about the newest emerging technology related to their field, should this be a requirement or an outcome based approach, to keep up with it relevant to your professional, or very specific requirement or specific to technology. [Regulator redacted] requirements are non prescriptive, standards not designed to work in that way, but professional bodies have more specific more prescriptive standards, and then down to the employer may have very specific standards. We are using these AI technologies and you must be proficient in using them"

> "Also – should be incorporated in pre-registration training programmes (e.g. The [regulator redacted] produces guidelines on what [healthcare professional] programmes must include – maybe AI should be included in this going forward)."

However, attendees also raised concerns that additional training would add to existing workloads and completing this extra professional development would be burdensome:

> "Aren't we just adding to [healthcare professional's] existing workloads expecting them to do this/receive additional training in it?"

Reporting systems for AI use:

This subsection describes attendees' discussion about factors that need to be considered, and elements needed, when developing reporting mechanisms for AI use.

Attendees recognised that both humans and AI make errors, at both use and design stages, which highlights the importance of reporting systems and joined up post-market surveillance to catch and address issues as they arise.

> "Humans make errors and info input may have flaws."

> "But some clinicians may be over confident in their ability to mitigate against the bias in the AI. This could then end up leading to harm."

> "Challenges around post market surveillance: Because if an AI, if let's say there's a discrepancy between an AI and a healthcare professional, in theory that discrepancy needs to be documented. It needs to be written down. The manufacturer needs to be notified so that they over time can monitor whether that's a one off or whether there's sort of whether there's a pattern occurring."

> "So where problems with a system like this [are] identified that they can be flagged and that kind of other clinics or other practices that are using similar systems can be aware that that such issues have been identified. So, the developer needs to be aware, but also places that have those in use [that] might be experiencing some similar problems might need to respond accordingly." **Facilitator feedback to main room from small group session 2, group 4 (scenario 4)**

This supports the idea that decision-making power must remain with people, and emphasises that people need to remain in the loop to both ensure that AI remains a tool that enhances care without replacing humans, and to ensure there is human oversight that can recognise and report errors.

> "AI tools should be a tool used for guidance by [healthcare professionals]. That it requires oversight and thorough checking by [healthcare professionals] at all stages."

> "Patients and [healthcare professionals] remain the decision-makers."

However, it was acknowledged that this assumes that the AI user will be in a position to know that an error has occurred. The fact that an AI is being overseen by a trained human user does not guarantee the human will spot a flaw, either in the AI or in their own practice. When fallible AI users work with fallible AI systems, errors can be magnified and hard to spot.

> "How will [healthcare professionals] know if something is wrong?"

> "Users may not recognise that the AI is incorrect."

> "Need to know what is wrong, to be able to report appropriately."

> "People shouldn't use AI until they are an expert – you need to be better than the AI to know whether the AI is useful/correct."

This risk further supports the need for training that prepares professionals for appropriate AI use, and equips them to spot and report errors.

Further discussion of this issue emphasised that identification and reporting of errors should not just be the purview of professionals, and that service users ought to be sufficiently informed, equipped, and supported to report issues arising from AI use as well.

> "Needs to capture patient and public experience."

> "Patient and public involvement at every point."

> "Needs to be easy to report and needs to also act as an early warning system to others across health and care. Surveillance."

> "Speed may be of the essence. Something like the MHRA Yellow Card system?"

> "Yellow card AI reporting for patients, families and carers."

Attendees also noted that recording and reporting systems need to collect different types of data about similar and dissimilar AI systems, and that systems not only capture where things are going wrong, but also what is going right.

> "Concerns should be shared across similar services across a trust/region/nationally if the same tech is used."

> "Specific AI codes on patients' health care records – useful for safety and research."

> "Needs a Safety 1 and Safety 2 perspective. Capture AI failures as well as where AI is working well, needs to look at organisational culture and capacity to learn from when things go wrong. Needs mix of: 1. Concern reporting 2. Automatic ongoing performance monitoring 3. Outcome tracking 4. Easy reporting systems that can be linked and analysed nationally or even internationally."

It was additionally noted that reporting and monitoring is just one part, and that there needs to be a commitment to act on the information collected.

> "Facility and responsibility to escalate concerns and enable overarching action to correct/ remove AI that is malfunctioning?"

Alongside the creation of monitoring and reporting systems, it was emphasised that a culture where users feel safe to report errors must be developed.

> "...there was an appreciation that whilst [the AI user's actions when using the AI] is the professional's responsibility, that doesn't necessarily mean they should automatically be sanctioned... there were sometimes system pressures that can mean that whilst the professional is still responsible, they may not be blameworthy in a way that requires or demands sanction. So, we want to think about a developing a culture where there is freedom and a safe way to report problems so that we can learn from them without necessarily being scared about sanction." **Facilitator feedback to main room from small group session 3, group 5 (scenario 5)**

> "Also need to have systems in place to protect those who raise concerns/whistleblowers."

It was also noted that there is a disconnect between independent and National Health Services (NHS), and that reporting mechanisms for AI must cover both NHS and non-NHS services. AI users practising in independent services might not have the same level of support available as those in the NHS.

> "There's quite different considerations in terms of responsibility and accountability for those working in private practice versus those working in a much more structured kind of NHS setting where there's more layers, more supports available. Whereas potentially someone might just be virtually on their own, they might not have kind of sort of options in terms of people to carry out the different roles. They might not be able to access so much support." **Facilitator feedback to main room from small group session 3, group 4 (scenario 4)**

Who should be accountable / responsible for any harms caused in the scenarios?

This section addresses the second of the workshop's four main questions, which examines who should be accountable / responsible for any harms caused in the scenarios.

This question was directly discussed in a small group session in the afternoon, but the subject of responsibility was frequently raised over the course of the day.

Firstly, and importantly, attendees highlighted the need for a clear definition of responsibility, as the term could denote multiple things, ranging from an obligation to log a concern, through to liability for harm. This highlights the need for regulators to use clear and specific language.

> "...need to define 'responsibility' for the purposes of guidance – may not mean taking full responsibility but logging issues, raising concerns, interacting with oversight mechanisms." **Facilitator feedback to main room from small group session 1, group 4**

It was acknowledged that there are a vast array of actors who could potentially hold some form of responsibility and accountability for AI use, including:

- AI users
- Supervisors of student AI users
- Managers
- Employers
- Organisations, e.g., hospitals, dental practices
- Regulators
- Developers
- Commissioners.

It was widely and strongly felt that regulators had a role in helping to clarify lines of responsibility and accountability around AI use, and that it was very important to do so, and yet the discussion here did not generate any clear sense of how responsibility could be determined.

Whilst clarification on responsibility and accountability was viewed as essential, it was at the same time recognised that allocating responsibility and accountability for AI use (viewed in terms of liability for harms) to specific actors is nuanced and likely needs to be determined on a case-by-case basis (dependent on the unique details of any given AI use scenario), and it was also recognised that this makes prospective guidance challenging to write.

> "...hard to imagine situation where there could be blanket guidance on who is always responsible or strict lines on if clinicians, or AI developers. so different and so different in different professions. when harm arises, need to look at case by case situation." **Facilitator note from small group session 3, group 1 (scenario 1)**

At times, during discussion of specific scenarios, attendees felt that lines of responsibility and accountability were very clear, and could be allocated to individuals or organisations.

> "The AI Kit should be developed with this safeguard built in... send responsibility back to the developer." **Attendee comment towards scenario 2**

> "...professional is responsible for accuracy of record." **Attendee comment towards scenario 5**

> "The hospital – the provider is responsible for providing care that meets fundamental standards." **Attendee comment towards scenario 3**

> "Developers continuing to retain responsibility after deployment. So it's not as though the system has just been handed over and then can, you know, create harm or they are absolved with any responsibility within settings." **Facilitator feedback to main room from small group session 3, group 1 (scenario 1)**

> "The majority perhaps being on the product developer. If they hadn't released a product that they could assure as being safe or more likely to be safe than not." **Facilitator feedback to main room from small group session 3, group 3 (scenario 3)**

It was also felt that in some cases, responsibility and accountability could be shared across multiple different actors.

> "I think people were fairly clear that that was laid out largely with the developer where this sort of original fault had occurred in terms of the system not being able to understand all accents. But also some responsibility for the dental practice who should have carried out kind of due diligence in, in the way that both in their procurement and in their deployment of this particular tech." **Facilitator feedback to main room from small group session 3, group 4 (scenario 4)**

> "There would be a team responsible for making a decision you would hope – even with an AI output – so joint accountability rather than individual blame."

The potential for proactively sharing responsibility was described, including across service users when a decision to use AI is shared.

> "Has to be a partnership, team work, all the way through. Tension between no fault/blame, and outcomes for patients."

> "Where there are high levels of uncertainty, there needs to be shared decision-making with the patient"

> "I would suggest that shared decision-making is critical here and that the consent should be linked to this."

Taking a step back from individual actors and looking more broadly, some attendees noted that responsibility and accountability should be carried across the system.

> "That it wasn't likely to sit with the practitioner unless they failed to follow the guidance and the existing shared decision-making practice in place. So, with the AI or otherwise, they should already have safeguards in place to mitigate the risk of an individual decision. More likely it should go to the hospital with the wider providers that put them in the situation. The majority perhaps being on the product developer. If they hadn't released a product that they could assure as being safe or more likely to be safe than not. Providers who are using AI, they should be already having their own risk assessments in place. They should understand the risks and mitigate them as much as possible before they allow harms to potentially happen. In doing so, we're when deploying it, the developer should be able to explain the technology's limitations, trustworthiness like the percentages and errors that are likely to happen and what those margins are. And this should be unclear at all stages of the process from the hospital down to the practitioner." **Facilitator feedback to main room from small group session 3, group 3 (scenario 3)**

> "Ultimate accountability for patient harm should sit at system level rather than solely with the frontline clinician."

This kind of system-wide approach seems appealing in light of concerns expressed from some attendees about shifting all responsibility onto individuals, exemplified by one attendee pointing out that it was unfair for AI users to understand/factor for the technical issue of 'drift'*.

> "Does not feel fair for the individual professional to be responsible for this. This is part of clinical governance org level."

> "If we are putting responsibility onto the clinicians to report and identify problems, are we shifting responsibility away from the developers?" **Facilitator feedback to main room from small group session 1, group 5**

One facilitator's feedback to the main room noted that their group felt AI developers cannot wash their hands of the effects of an AI device once it is in the hands of a user.

> "Developers continuing to retain responsibility after deployment. So it's not as though the system has just been handed over and then can, you know, create harm or they are absolved with any responsibility within settings." **Facilitator feedback to main room from small group session 3, group 1 (scenario 1)**

***Drift can be understood as data and populations changing – or drifting – over time. For example, if an AI voice scribe had not been updated in a while, it might not be able to capture the latest terminology (for example a new drug or diagnosis) when the clinical user dictated to it. Drift captures the idea that the world has changed since the AI was trained, and the AI has not been updated to reflect those changes.**

Whilst some attendees were firm in the view that the professional user bears ultimate responsibility for any consequences of their own AI use, it was also acknowledged that in the presence of institutional/structural pressure to use AI (for example, to maintain or improve efficiency), the notion of responsibility may become blurred.

It was felt that employers would have a significant impact on the way AI is used – for both good and ill. A concern was raised that professionals may be put under pressure to use AI to make efficiencies and not be given the time to undertake the necessary checking and assurance.

Consequently, it was felt that employers will need to ensure there is sufficient time for professionals to undertake necessary assurance and be respectful of the professionals' risk analysis and decision-making. It was suggested that, at a local level, to prospectively plan for responsible AI adoption in an organisation, an AI officer post could be created to champion local level guidance for AI use to support AI deployment.

> "It's really important to make sure that we to give people time to do the assurance checking that is needed. So it was noted that even if guidance here is not needed to highlight what the professional responsibility is, guidance might be useful for the professional to use to give to their employer or to give to the institution that may be putting pressure on them to use these tools to make efficiencies, to say no, you have to give me time to do the appropriate assurances." **Facilitator feedback to main room from small group session 2, group 5 (scenario 5)**

> Regarding discussion about taking AI out of service if risk of harm/injury: "Will their providers / employers allow this?"

> "One thing to look at could be local AI guidance through the development of an AI Officer post at maybe the Integrated Care System level or local NHS trust level or primary care network level. They can help the organisations in their patch develop safety policies, complaints policies and PALS guidance. They could also support implementation and training. They would also need some patient involvement skills to co-develop these systems."

Overall, it was felt that there needs to be clear regulation to cover post-market monitoring and reporting, and that collaboration would be key to this.

> "We had a conversation about pre and post market regulation and pre and post market surveillance. So it's not just before the things go out into the world, but it's important to reflect on who continues to be responsible and potential regulatory gaps and the role of regular continuing role of regulation within both health and social care." **Facilitator feedback to main room from small group session 3, group 1 (scenario 1)**

> "One thing that I feel came through in the conversation has come through all, all day is the importance of collaboration and the need for continuing collaboration and conversation, whether that's across sectors or across the different actors who are part of this the system as it's being deployed and adopted." **Facilitator feedback to main room from small group session 3, group 1 (scenario 1)**

One attendee highlighted the Responsible Handover Framework (Sense about Science, 2024), which they felt could serve as an example of how involved parties could effectively communicate and share information.

Attendees acknowledged that uncertainty presents a regulatory challenge and that, in the absence of either a clear legal framework, and until there is a sufficient body of experience, regulators will need to be comfortable dealing with uncertainty.

> "Over time law practice and dealing with incidents will help us develop our understanding and the important ethics and how we keep pace with that and how we change as we need to develop a more informed understanding... Regulators like certainty and this is everything everywhere all at once right now. Have a plan for dealing with uncertainty, a better, more coherent approach to AI right now."

Broadly, and notwithstanding the recognised challenges, attendees did feel that some guidance from regulators around responsibility is desirable.

> "Would be helpful to have guidance on clinician responsibility and basic questions to assess when harm is caused, and whether they had appropriate competence, if they applied it right, and for employer whether the applied training and the use of the technology, if there is appropriate supervision. Different checks for the AI researchers and developers, and owners. do not think there should be very strict release of liability regardless of who they are."
> **Facilitator note from small group session 3, group 1 (scenario 1)**

Who should be responsible for detecting and addressing issues in AI prior to deployment?

This section addresses the third of the workshop's four main questions, which explored who should be responsible for detecting and addressing issues in AI prior to deployment.

Attendees again noted a variety of potential actors when responding to this question:

- Product developers
- MHRA and NICE
- Organisations, eg, hospitals, dental practices
- AI users.

The overarching response to this question was that individual AI users cannot and should not be responsible for detecting issues with AI prior to deployment, but that there is scope for shared responsibility across other actors. Notably, relying on developers being left alone to identify and address issues was viewed as problematic, suggesting a need for independent oversight and verification.

> "This technology is not fixed. it will continually develop and change over time. Therefore increasing need to consider moving towards no fault learning and some form of shared liability across all agencies."

> "The user and developer are most responsible but they also need to allow service users to feedback issues with using the device."

> "Should the developer be testing their own product?"

> "No device manufacturer should be self-regulating. As well as the developer itself, obviously you need the regulators, the MHRA and NICE." **Facilitator feedback to main room from small group session 3, group 2 (scenario 2)**

One attendee illustrated the idea of shared responsibility using scenario 4, noting that the scenario reflected a failure on the part of both the developer and the practice to undertake reasonable pre-deployment checks or consider foreseeable issues.

> "Both the developer of the software and the dental practice for not procuring correctly and not doing proper risk assessments of the deployment of the software. Ultimately from a regulator point of view, it's unlikely we would be able to pick this up with the developer of the software (in most cases) and therefore we would focus on the organisation/individuals using/deploying."

The quote below (left) also captures the concern that health and social care regulators are not currently in a position have any influence over the development process. This was picked up elsewhere, where concerns were voiced about the lack of joined up processes pre- and post- deployment, with the suggestion that regulators could play a role in remedying this.

> "Detecting prior to deployment, not the responsibility of professional regulation to do that and can contribute but there needs to be an infrastructure that does not exist in social care at the moment, no NHS commissioning body or NICE, or MHRA, nothing like that that could be used to regulate devices before they hit market. that is problematic, it creates a vacuum where devices are being tested and they are evidencing and evaluating their own product and there is no one checking the outputs. it absolutely should exist, it should be the responsibility of a centralised regulator of an AI product and the produce or the technology." **Facilitator note from small group session 3, group 1 (scenario 1)**

> "I don't know the answer to this question, but I think that there ought to be some sort of oversight and feedback for learning from mistakes/improper use/ inappropriate suggestions from the AI – I think that the developers ought to be responsible for their products but I imagine it would also take the establishment of some sort of regulatory body to organise this."

Other attendees highlighted the importance of Patient and Public Involvement (PPI) during testing, development, and deployment, although this is not the same as holding public and service users responsible (even in shared sense).

> "The user and developer are most responsible but they also need to allow service users to feedback issues with using the device."

> "Is there a support group for this AI? (Common in IT systems) Form a test team in the Trust? Include PPI and users to test practicality."

To what extent should the focus be on no-fault learning and correcting of AI users rather than responsibility or blame?

The final of the workshop's four main questions probed attendees' views of no-fault learning and correcting of AI users, as opposed to responsibility or blame.

There was, amongst attendees, an almost ubiquitous endorsement of the need to encourage learning from mistakes, which would enable issues with AI to be observed, reported and tackled, rather than a culture of blame. For some, this was about developing a culture where people are supported to report problems, and for others it was important to develop a culture of no-fault learning.

> "Learning and improvement is very, very important." **Facilitator feedback to main room from small group session 3, group 4 (scenario 4)**

> "Must be no-fault learning to avoid stress and over caution."

> "want to encourage and support people to report problems."

> "All understand the risks and could be expecting there will be learning from mistakes, there should be an open dialogue between the developer and the clinicians, always two-way street where we are developing constantly, even though the tech is already deployed and used, should always be reflection and open dialogue to learn from those mistakes. Post Market, needs to be an iterative approach to improving device on market, should go without saying." **Facilitator note from small group session 3, group 1 (scenario 1)**

> "Do need to monitor and address multiple mistakes."

> "It should be a learning experience partly based on the feedback of people who use it."

> "Unless you speak to the users it might not always be obvious about the context or when to step in."

> "Problem is that there is a bit of a blame culture in the media or when patients are harmed."

Concerns were also raised, however, with idea of no-fault culture, because it may encourage a passive and uncritical acceptance of AI alongside a lack of vigilance, and also because sometimes there may in fact be fault. Rather than an assumption of no fault, cases need to be determined on their own merits.

> "Need to encourage critical thinking and challenge rather than passive acceptance that AI knows best."

> "It would depend on the impact of the harm and the track record of the organisation we are regulating. Is this a one-off issue and that learning is taking place and putting in place proper risk assessment."

Others felt that 'no-fault' did not mean that individuals could never be held to account. Rather, a culture is required where people can speak up or admit mistakes without fear of punishment, supported by an understanding of the complexities and challenges of deploying AI, while still ensuring individual accountability and responsibility where appropriate.

> "There was no harm in this scenario [scenario 2] so absolutely can be accepted as a no-fault learning. May need to be considered differently if there was harm as a direct consequence."

> "Monitor performance to appropriately attribute responsibility in the instance where [an] individual persists in errors."

> "I think learning and improvement should be the focus to an extent and appropriate for this situation – however when there is deliberate use of a tool by a professional in the care of a patient there should be an element of professional responsibility – just like how practitioners choose medication, dressings, care pathways etc."

There was, however, some evident scepticism about this. Some attendees suggested that regardless of original intent, there will always be the risk of litigation and individuals being unfairly held responsible for system issues.

> "There's also the thread of litigation weaving its way through the end to end process, so while learning and transparency approach is important, the reality of 'formal' blame via litigation will be a factor?"

> "Who should be accountable, responsible, as regulators have to hold professionals responsible on some level and that has to be risk-based, and that should be part of a system and that is shared. At the moment it feels like it is filtered downstream. Feels like responsibilising [the] practitioner which seems unfair." **Facilitator note from small group session 3, group 1 (scenario 1)**

Other pertinent points raised by attendees

A number of other key themes were identified throughout the discussions. These did not relate specifically to any of the key questions from the workshop, but are nonetheless highly relevant to consider in relation to thinking about developing guidance and regulation in this area.

First, attendees emphasised the importance of equity, inclusivity and representation in all areas of activity, across development, use, and reporting. Concerns included the (lack of) representativeness in training data, access to opportunities to provide feedback and report, the potential to exacerbate health inequalities, and the risk of minority voices being overlooked, with lived experiences getting filtered out or ignored.

> "Preventing unintentional or unrecognised harm and widening healthcare inequalities:
> - unintentional harm comes from the gaps in evidence when not everyone's voices have been heard or not the right voices have been heard
> - those impacted by the greatest health inequalities are also those that face barriers to engaging to being heard
> - If the AI research data that feeds the policies and regulations for practice does not hear these voices that's where harm can come in and inequality gaps in healthcare provision can widen
> - we already know there are huge existing inequalities in health and care – how do we regulate AI use so that as a minimum it doesn't widen inequalities and at best it reduces those existing inequalities?
> - We already know that data fed into AI systems leads to discriminatory actions – e.g. the racial bias within AI systems fed on predominantly White datasets
> - how do we ensure that for example that existing inequalities in detecting jaundice in White vs Black or Brown babies is not worsened
> - diverse and underserved patient groups must be involved in decisions about how AI is deployed
> - There is absolutely no one size fits all."

This historical lack of inclusive research design gave rise to concerns about the presence of bias in AI, and the negative impact this can have. As noted above, the discussion included questions about how to regulate AI use so that it does not widen inequalities and acts to reduce existing inequalities. Responsibility for mitigating bias was generally viewed to lie with all actors, from developers, through regulators, to AI users.

Going forward, regulators may have a role in ensuring that all work in AI, ranging from the gathering and selection of data with which to

train the AI through to procurement and use, is taken in partnership with all groups who will be affected by it. This was expressed through the repeated attention called to the importance of meaningful service user and public involvement and engagement during AI development and deployment.

> "Iterative and ongoing."

> "Co designed with strong PPI."

> "Absolutely taking PPI input into account wherever relevant."

> "Form a test team in the Trust? Include Patient and public involvement and users to test practicality."

Attendees raised concerns that AI users may not trust or believe what is told to them by service users, which will affect the information that will be provided to the AI (which it then learns from), resulting in further risk of erroneous outputs. When the people training or interacting with the AI have a bias against the people providing the raw data, it creates a specific kind of technical failure.

> "Concern that clinicians do not believe patients, and the inaccuracy may be put into AI - errors get magnified."

Second, it was noted that regulatory bodies are not the arbiters of all things, and do not always have the expertise or skill to advise registrants on every area of practice. Sometimes, all a regulator can do is refer enquiries to other relevant bodies.

It was felt that whilst regulators may have released guidance about AI which links to their practice standards, they do not always have the expertise or skill to fully advise their registrants, for example in areas such as use of judgement in when to use AI tools to support clinical thinking.

> "Particularly with small professions where lots of people are independent practitioners working in sole practice, they will often look to the regulator to be the arbiter of all things and effectively take the place of management. And so, we set standards on complying with data protection laws and regulations and health and safety laws and things like that, but if they come to us with questions about 'actually how do I do this?' we would refer them to other authorities like the Information Commissioner's Office or whatever. We have published guidance on the use of AI in practice for [our registrants] linked to [our] practice standards, but we don't always have the expertise or the skill to say 'this is what that means' – that's the difficulty. These technologies are out of our expertise and so it's about how they exercise their judgement in determining where something is appropriate to use and how they might use that." **Verbal quote in main room from attendee**

Third, it was felt strongly that service users should be aware of AI use, and its role, in any and all aspects of their care, and this calls for transparency. This transparency is a prerequisite for shared decision-making, ensuring that consent is not a mere formality but is linked to a collaborative process where the patient's role is central.

However, the informed element remains a challenge, as it is uncertain if everyone can sufficiently understand what the AI is doing, including the risk of error and the potential for harms that may take time to detect.

> "I would suggest that shared decision making is critical here and that the consent should be linked to this."

> "I think there are issues with the 'informed' element of consent to AI use. will everyone always be able to understand precisely/ sufficiently what the AI is doing?"

> "...patients need to understand the risk of error, especially if they have an input on decision made by made by the AI. And there's a risk that some harms caused by AI might take months or years to detect." **Facilitator feedback to main room from small group session 3, group 3 (scenario 3)**

There were, however, mixed views about to what extent AI use must be part of the informed consent process, with some suggesting that special consideration might not be needed. As AI is adopted more widely and becomes part of day-to-day practice in health and social care, the need to treat AI as a distinct issue requiring special consent may no longer be so relevant. Furthermore, there is a risk in treating AI differently than other medical technologies or interventions; instead, an approach may be preferable where patients are informed and provided with choice, rather than requiring a formal step of explicit consent for every use of AI.

> "Yes but surely fairly soon it just will be most of the time?"

> "Until it becomes adopted as 'usual' practice and therefore does not warrant special consideration"

> "However....if AI has become 'usual care' the 'AI' as an issue is no longer relevant."

> "It will be hard to know all the parts of the chain of care where AI might be involved."

> "Risk here of treating AI differently to other medical tech or interventions; I would take a slightly softer approach that patients should be informed and have choice, rather than needing explicit consent."

One especially salient comment called attention to the evolving nature of AI, professional clinical practice, and societal values, which could force AI users to walk an ever-moving tightrope between giving too little and too much information about AI use.

> "Will we soon be in a time where aspects of AI use are so ubiquitous that we expect and understand that they are in use somewhere through the process?"

Summary discussion

Discussion over the course of the workshop identified a number of concerns about the risks of AI use in health and social care. Participants felt strongly that regulators have a role to play in supporting professionals navigate these risks, though they also acknowledged the significant challenges regulators face in doing so.

Risks

Both risks to service user safety and to professionals were surfaced in the discussions and formed a consistent focus of concern.

Service user safety is at the heart of health and social care, and the overarching concern that all AI used in this space should be safe and fit for purpose was visible throughout all the discussions. Service user safety starts with responsible development involving those which the AI would affect, and it is taken for granted that risks arising from biased training data and insufficient PPI need to be addressed at the development stage, and mitigated prior to deployment.

> **“Humans make errors and info input may have flaws”**

Focusing on the user end, risks tended to be articulated in terms of users either having insufficient training in the proper use of AI or being in a position where they had to use AI but were not in a position to undertake the necessary assurance. One specific and illustrative concern was that employers may put pressure on professionals to use AI in order to make efficiencies and save time, but that the time saved through the use of AI would then have to be used to see more service users, leaving insufficient time for quality assurance.

The risk here is that AI will eventually become a *de facto* replacement for some human roles, where humans retain responsibility but are not in a position to undertake necessary due diligence. This potentially puts at risk both service users and professionals.

The importance of training is linked to the AI user's ability to report issues, as reporting is entirely dependent on their being able to recognise a fault or error in the first place. Reporting need not, however, be the sole responsibility of professional AI users, and could be shared more widely, e.g. service users and their significant others can feedback using the MHRA's Yellow Card scheme. PPI work was suggested, not just during testing and development (as mentioned above), but also throughout deployment, enabling additional opportunities for recognition and reporting of AI issues as they arise.

Risk is not fixed, nor solely determined by the AI. Different AIs, used by different professionals who have different levels of skill, would present different levels of risk. Such fluctuations underline the difficulty of attempting any kind of universal adoption guidance for any given AI within organisations or systems. However, common issues, such as the adoption of biased AI or exacerbating inequalities through AI use, were foreseen by attendees; guidance could have a role here in informing AI users to identify and navigate anticipated risks which are universally present.

The forward planning needed to determine who should bear responsibility for the risks associated with using an AI system was recognised as inherently challenging. While individual actors will inevitably remain responsible for their own decisions and actions, participants emphasised that there is scope for risk to be distributed across the wider system. Moreover, service users need to understand the risks they may be subjected to when AI is involved in their care, and that harm caused by using an AI might not be detected for some time. Employers play a crucial role in creating the conditions for responsible and safe AI use (as noted above) in particular the need to provide professionals with sufficient time and support to undertake the necessary assurance and checking activities. This was seen as a key component of creating an environment that enables responsible adoption of AI while mitigating foreseeable risks. Ideally, risk management strategies such as assurance and checking would be streamlined across all healthcare services, yet there are significant differences between the national and independent sectors – the regulatory implications of determining and carrying these responsibilities are further discussed below.

Challenges

It was fully appreciated, and referenced multiple times during the day's discussions, that there are significant regulatory challenges associated with AI.

Regulators face the challenge of finding the right level of specificity in any guidance they issue or standards they articulate. This challenge is not unique to AI, but it is made complex by the multifaceted nature of the technology and a very wide range of possible uses across the health and social care professions.

The challenge above is linked to the challenge presented by the need for consistency across professions. Different health and social care professions will have different challenges arising from different clinical contexts and differing technologies. Regulators will need to

find a way to articulate consistent high-level standards, which can then be articulated and implemented in different ways in different professional contexts.

> “This technology is not fixed. it will continually develop and change over time. Therefore increasing need to consider moving towards no fault learning and some form of shared liability across all agencies.”

AI health technologies remain new, which means explicit legal consideration is relatively scarce, with few examples of the legal ramifications of the use of AI tools in health and social care available and a corresponding absence of case law in this area, including for liability and accountability. Regulators are very unlikely to be able to prospectively determine lines of responsibility, and this will only become with the development of legislation or case law. This is a particular challenge, given that a primary concern raised during the discussions was a lack of clarity and uncertainty around the circumstances in which a professional would be responsible for any harms arising from AI use, making it impossible to draft definitive guidance in this area.

A further challenge is that case law arising from decisions made by statutory professional regulators does not necessarily apply in the same way to Accredited Registers. It is also important to recognise that some professionals using AI are neither subject to statutory regulation nor members of an Accredited Register. In these circumstances, employer requirements – and for self-employed practitioners, the wider legal frameworks that apply to all businesses – take on greater significance. This fragmented nature of the regulatory landscape means that professional regulators and Accredited Registers can only go so far in providing consistent, high-level ethical guidance. This may, in turn, fall short of, or even conflict with, the expectations of patients and service users. It is therefore important that employer requirements, and any future legal frameworks for regulating AI use by professionals, can make sure there are protections to the public no matter how they are accessing health and social care.

A significant challenge to regulation in this area is presented by the pace of technological development. This suggests that any attempt to provide guidance will need to be agile and iterative, with regular review to keep pace with rapid technological shifts and so that guidance remains effective as AI technologies advance. This would also need to take into account the anticipated rapid advancement of AI technologies, including potential implementation and scaling of AI in health and social care.

Regulatory support

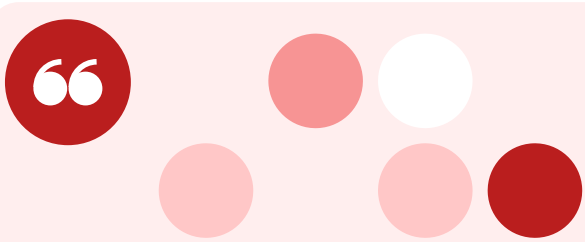
Whilst it was recognised that professional regulators cannot control or regulate all aspects of AI, it was felt that regulators can and should set standards and support good

practice in areas where they do have influence. Even if guidance has to be high level, there is the opportunity for professional regulators to set the tone around AI use, which will help health and social care professionals feel confident that they know the standards expected of them when they use AI, and to be empowered to reasonably reject or challenge the use of AI.

The major concern that drives the need for some guidance in this space is the risk to professionals created if they are put in a situation where (1) AI use is either mandated or encouraged, (2) where they are not sufficiently empowered to challenge its use or outputs, and (3) where they retain all personal responsibility as decision makers. The combination of these three factors risks a 'perfect storm' where professionals become moral crumple zones (Elish, 2019) that soak up risk.

It is important that any guidance is informed by parties that it would affect – that means inclusion of the views of communities such as service users, those who have historically been under-represented in health and social care research and policy-making, those who would use AI, and those who would regulate AI users. Meaningful service user and community involvement is essential to undertake.

It needs to be acknowledged that whilst there are already standards, and ethical guidance and principles, relevant to the use of any technology or instrument in health and social care, AI is a disruptive technology that is likely to challenge the *status quo*. As such, whilst entirely new principles and standards are unlikely to be needed, it is important for regulators to anticipate the disruptive impact of AI in practice, and provide support – where feasible – to professionals to help them navigate the challenges it presents.



"As a regulator, we need to focus our support for registrants on using AI in line with our professional standards."

Although professional regulators likely have very little ability to influence the development pathway, and the testing and approval of these technologies sits outside of their remit, they may nonetheless be able to take a public position on the need to ensure high safety and efficacy standards of AI technologies before they are deployed. This would provide a demonstrable standard to support a) AI users' judgements when deciding whether to adopt AI into their practices, making clear that registrants can rightly expect those standards, and speak out if they are not met, b) service users, who can refer to the standard when discussing and deciding their care with the teams caring for them, and c) AI developers, who will have a clear understanding of what is required of them by end users.

Regulators are unlikely to be in a position to determine exactly where lines for responsibility lie (this is for law to determine), but they are in a position to set the tone and lay out the conditions under which a professional might reasonably use or reject an AI in their practice, and provide clarification on what would comprise adequate due diligence around these technologies at the point of use. It might be specified, for example, that a professional is required to

satisfy themselves that the technology is approved for the use to which it is being put, that a professional is required to follow guidance on usage, scrutinise the output for face validity, and has an obligation to report errors and concerns. Specifying such conditions would also serve to make clear organisational level responsibility. For example, if a professional has an obligation to report errors and concerns, it follows that there is organisational responsibility to ensure there are adequate reporting mechanisms. By setting out basic standards and guidance for the use of AI in clinical spaces for professionals, regulators can help empower individuals and clarify the structures required for them to operate safely and responsibly.

Regulators are in a position to determine training requirements and standards, and the importance of doing this ought not be underestimated. Sufficient understanding of what AI is, how it functions in a clinical space, what its limitations are, and what it can and cannot be used for, is foundational to safe and effective use. Whilst it is unnecessary for all end users to have full technical understanding of AI, being a responsible and critical AI user does require a basic level of understanding of the technology (just as to be safe car driver one needs to have a basic understanding the car). When life and wellbeing is at stake, this becomes more than critical. Different regulators will have different requirements, but all ought to take a position on the requirements for training at both pre- and post- qualification stages, and work with educators and employers to ensure standards are met.

Related to the above point about training, regulators are in a position to support training by co-developing AI use case scenarios that illustrate how existing standards might be operationalised. A clear example, arising from the workshop, is informed consent. There are already very clear standards and expectations around the need for informed consent, but the extent to which AI use can be, and needs to be, explicitly discussed, is unclear and will depend on the particular service user care context, and on evolving social norms. Managing complexity around consent is not a new problem, and AI use does not require new standards or principles in this area, but professionals may need support in thinking through consent when AI is needed. Regulators cannot prescribe how this can and should be done in every possible case, but they could clearly articulate the need and support the development of training. Whilst educators are already grappling with these challenges at a pre-qualification level, there may be a more urgent need for regulators to intervene to ensure that these kinds of cases are included in continuing professional development.

Limitations of the workshop

There were some technical difficulties throughout the day that affected some attendees' ability to fully engage at certain points, which included difficulty for some people moving in and out of the small group sessions, and difficulty editing Padlets. Whilst these issues were addressed as quickly as possible, we acknowledge the impact this had.

Every effort was made to meet the needs of attendees with disability access requirements, through early access to the workshop materials and small group facilitators offering to scribe to the Padlets wherever a need was identified. However, we recognise that this process was imperfect, especially in small group session one. We are grateful for attendees' patience with us.

Recognising the impact that these challenges had, all attendees were invited to contact us separately via email with any additional comments that they would have liked to have made over the course of the day. Padlets were also kept open after the close of the workshop (and for the rest of the day) to give people time to contribute anything that had been missed.

Quick links/find out more

- Find out more about our work around Artificial Intelligence
- Find out more about our February 2026 workshop
- Read our blog Artificial Intelligence in healthcare: who holds power?
- Read our Response to the call for evidence by the National Commission into the Regulation of AI in Healthcare's call for evidence

References

Elish, MC (2019). *Moral Crumple Zones: Cautionary Tales in Human-Robot Interaction*. Engaging Science, Technology, and Society. https://doi.org/10.17351/ests2019.260

Department of Health and Social Care (2015). *Code of practice: Mental Health Act 1983*. https://www.gov.uk/government/publications/code-of-practice-mental-health-act-1983

General Medical Council (undated). *Artificial intelligence and innovative technologies*. https://www.gmc-uk.org/professional-standards/learning-materials/artificial-intelligence-and-innovative-technologies

Medicines and Healthcare products Regulatory Agency (MHRA) (2025). National Commission into the Regulation of AI in Healthcare https://www.gov.uk/government/groups/national-commission-into-the-regulation-of-ai-in-healthcare

Professional Standards Authority for Health and Social Care (2025). *Right-touch regulation 2025*. https://www.professionalstandards.org.uk/publications/right-touch-regulation-2025

Sense about Science (2024). *Responsible Handover: How to share and implement data analytics and AI tools*. https://senseaboutscience.org/wp-content/uploads/2024/11/Responsible-Handover-implementation-guide.pdf

Smith, H., Ives, J. (2025). *Developing professional ethical guidance for healthcare AI use* (PEG-AI): an attitudinal survey pilot. AI & Soc 40, 5439–5455 (2025). https://doi.org/10.1007/s00146-025-02276-z

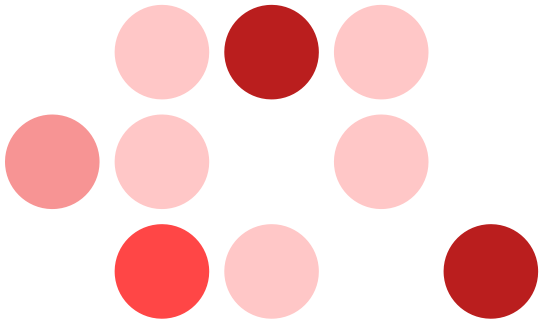
Smith, H., Downer, J., Ives, J. (2024). *Clinicians and AI use: where is the professional guidance?* Journal of Medical Ethics 50:437-441. https://doi.org/10.1136/jme-2022-108831

Society of Radiographers (2023). *Updated Standards of Proficiency for radiographers from 1 September 2023*. https://www.sor.org/news/professionalism/new-standards-of-proficiency-for-radiographers-tak

General Chiropractic Council, General Dental Council, General Medical Council, General Optical Council, General Pharmaceutical Council Nursing and Midwifery Council and the Pharmaceutical Society of Northern Ireland. *The professional duty of candour: joint statement*, (2014). https://www.osteopathy.org.uk/news-and-resources/document-library/our-work/joint-statement-on-the-professional-duty-of-candour/

UK Government (2025). Fit for the Future: 10 Year Health Plan for England. https://www.gov.uk/government/publications/10-year-health-plan-for-england-fit-for-the-future

Appendix A: Scenarios



Scenario 1:

A GP practice implements a new AI-based skin-screening tool to assist with early detection of dermatological conditions. Maya, who is a patient with darker skin, submits an image of a new hyperpigmented patch on her arm. The AI system classified the patch as low concern and advised no further follow-up. A clinician reviews the image on their computer and follows the AI's recommendation. Over the weeks that follow, the patch continues to change. And after requesting an in-person appointment, Maya is reviewed by a clinician who identified features suggestive of melanoma. They make an urgent referral to dermatology, who confirms that the lesion is an early-stage melanoma. An audit later reveals that the AI model had been trained predominantly on images of lighter-skinned individuals, reducing its accuracy for patients with darker skin tones. And Maya's case highlights to us how reliance on an insufficiently diverse training dataset can contribute to diagnostic delay and potential harm.

Scenario 2:

Four-year-old Sky is brought into the Emergency Department with a fainting episode at daycare. A student nurse applies an AI-based physical monitoring alert system to Sky, which was designed for adults, not children. The AI picked up that the heart rate was normal for an adult, without alerting that the heart rate would have been low for a four-year-old child. The student nurse did not recognise that the AI would interpret paediatric data differently, and the wider multidisciplinary team did not check whether the system was set to adult or paediatric mode. As a result, the patient's abnormal heart rate was not flagged promptly, causing a delay in assessment, diagnosis, and directing the child to definitive care. Although the low heart rate was eventually identified and the child received appropriate treatment, the incident caused reputational harm to both the AI system and the professionals who had relied on it as it had been used without verifying the correct settings.

Scenario 3:

Lewis is a diagnostic radiographer whose hospital uses an AI which can read CT scan images for abnormalities. The AI picks out an area of tissue and flags it as a tumour but is unable to explain itself any further. The area that has been flagged by the AI looks entirely clear to Lewis. The difficulty is that the AI is so advanced that it might be able to identify changes before a human eye can. Hospital policy says that the AI's outputs can be

followed, and the patient wants medication to stop the potential tumour from growing, but Lewis is not sure that there is even a tumour there. This leaves the multidisciplinary team to choose between two risks: do they treat for something that they cannot independently confirm with the risk of the treatment causing side effects? Or do they risk a tumour growing?

Scenario 4:

To streamline administrative tasks such as appointment bookings and bill payments, a dental clinic operates all of its enquiries, bookings, and pre-check assessments with AI-driven chatbots and apps. Whilst a human reviews all contacts with the practice, and it does allow patients to contact the clinic 24/7, it does reduce interactions with the dental clinic's team to the absolute minimum – i.e. only the examination and treatment in the dentist's chair. Patients are left feeling dehumanised from the loss of direct communication and empathy. Angus called to make an enquiry about an appointment, but the AI couldn't understand his accent. Unable to access routine dental care, Angus attended A&E, where staff determined his issue was non-urgent and dental in nature. This resulted in a prolonged wait, use of emergency resources for a problem that should have been managed in primary care, and frustration for both Angus and A&E staff. A subsequent review found similar issues had been experienced by patients with strong regional accents and those for whom English was not a first language, many of whom abandoned attempts to contact the clinic altogether.

Scenario 5:

A social worker has used the AI transcription tool to create a summary for a service user, Kai. The AI tool incorrectly indicates in the summary that Kai was experiencing suicidal ideation, but at no point in the encounter did Kai actually talk about suicidal ideation. The social worker – who is pressed for time thanks to high caseloads – takes a couple of minutes to briefly scan the transcript and then pastes it to submit the AI-generated summary as the official record. The summary is uploaded to Kai's shared care record, where the recorded risk triggers an urgent mental health assessment. Although professionals act appropriately based on the information available to them, the assessment is unnecessary and distressing for Kai, who feels misrepresented and alarmed by the process.

Appendix B: Workshop running order

Time	Activity
10.00-10.10	Welcome
10.10-10.35	**Introductory talk** Instruction session on how to use a Padlet (an online tool where people captured their thoughts throughout the day)
10.35-10.40	**Short break**
10.40-11.05	**Small group discussion 1:** To look examples of possible professional ethical guidance (from Smith & Ives's (2025) pilot research) & consider what guidance is & is not needed.
11.05-11.30	Feedback from small groups to Main Room
11.30-12.00	**Main room discussion:** **What should be in place to enable responsible AI adoption?** E.g. • **Guidance for AI users** • **Training for AI users** • **Reporting systems for AI use**
12.00-13.00	**Lunch**
13.00-13.10	Welcome back Summary of the morning
13.10-13.30	**Small group discussion 2:** **Each group was given a scenario and asked to consider what guidance would be useful to help health and social care professionals navigate the scenario.**
13.30-13.50	Feedback to Main Room
13.50-14.00	**Short Break**
14.00- 14.20	**Small group discussion 3:** **Scenarios were used to discuss accountability and learning in AI use, asking the following questions:** • **Who should be accountable/responsible in this scenario?** • **Who should be responsible for detecting and addressing issues in AI prior to deployment?** • **To what extent should the focus be on no-fault learning and correcting of AI users rather than responsibility or blame?**
14.20-14.40	**Feedback discussion Main Room**
14.40- 15.00	**Summary, final comments, thanks, and close**

professional standards authority
for Health and Social Care

University of BRISTOL
Centre for Ethics in Medicine

www.professionalstandards.org.uk