the digital health society
The data arm of ECHAlliance Group

iHD

Creating an environment of Trust and responsible AI adoption and Use

2023 RECOMMENDATIONS BASED ON CALLS TO ACTION ON HEALTH DATA ECOSYSTEMS

Round Table 7

Artificial Intelligence: Creating an environment of Trust and responsible AI adoption and Use.

Round Table findings and recommendations based on Calls to Action on Health Data Ecosystems

This report presents the findings of two multi-stakeholder Round Table consultations that explored what is responsible AI in healthcare and responsible AI adoption and use. They were convened by the Digital Health Society (DHS) and The European Institute for Innovation through Health Data (i~HD).

This report is a consensus of 24 invited expert stakeholders in technology, data science, health informatics, patient representation, regulation, trade and industry representation, academia, and policy-setting. **The Round Tables were held on 18th January and 9th February 2023.** This was followed by a series of deeper dive examinations of key topics that arose during 2023, involving some Round Table members plus further invited experts. This deeper dive collection of papers has just been published. This report should be seen as a companion to that and incorporates some additional recommendations from it.

The two Round Tables were scoped and convened by the DHS and I~HD neutrally and independently from the Round Table Programme collaborators and sponsors Johnson & Johnson, Microsoft, and MSD. This topic is part of a rolling programme of deeper dives drawing on 7 Calls to Action on Health Data Ecosystems that were published in 2020, taking an AI focus in particular on:

Context

AI algorithms, usually embedded within health applications and workflow systems, have huge potential to improve the effectiveness and safety of clinician and patient decision-making, and thereby to improve health outcomes as well as optimising the use of health system resources. However, it is recognised that there are a number of trust and change resistance factors, that are inhibiting the scale up of AI adoption. Healthcare environments (payers, provider organisations, professionals, patients, and carers) may be unclear or uncomfortable about how to engage with AI driven digital health solutions.

The definition of AI system used in the European AI Act is included in the Glossary at page 20 for reference.

There is already a rapid expansion of good practices and rules regarding responsible AI development that safeguard adopting healthcare organisations and patients. These include, at a European level, the ethical principles, and forthcoming AI Regulation/Act. A number of trust-enabling areas of AI R&D, such as the avoidance of bias in training data sets (or, at least, the explicit characterisation of the training data population) and the explainability of AI, are the subject of many European initiatives and current calls for EC proposals in aspects such as multi-modal AI, large language models, accelerated diagnosis in rare diseases, disease stratification, treatment personalisation, image and genomics analysis, personalised disease management tools and closed-loop continuous treatment systems.

Less work is being undertaken to clarify responsible use of AI, such as

a) the extent to which AI-driven diagnostic, treatment, monitoring, or prevention guidance should be complemented by clinical expertise or other evidence,

b) the care pathway scenarios in which AI solutions are gaining their own evidence base of responsible use, and

c) what levels of feedback on the outcomes of AI use are needed to rapidly optimise and improve the AI.

Responsible AI is ideally implemented as a partnership between developers and users (including responsible use cases), so that good practices on both sides can be co-created and learning can be shared.

There are additional areas of emerging good practice, such as which electronic health record interoperability standards-an algorithm can connect with in order to receive patient-specific data on which it will provide its guidance, and how the algorithm and the application(s) and services in which it is embedded can optimally be integrated within clinical and patient workflows. AI reasoning relies upon access to patient data inputs, the availability and quality of which may be critical to its safe and effective functioning. This requires a carefully constructed and trusted handshake between AI components and EHR systems.

There is an urgent need to collate and to distil the emerging learning in all of these areas as a coherent and formalised specification for responsible AI transparency, trustworthiness and use in healthcare.

Examples of responsible use at the health professional level could include ensuring that the AI decision making is:

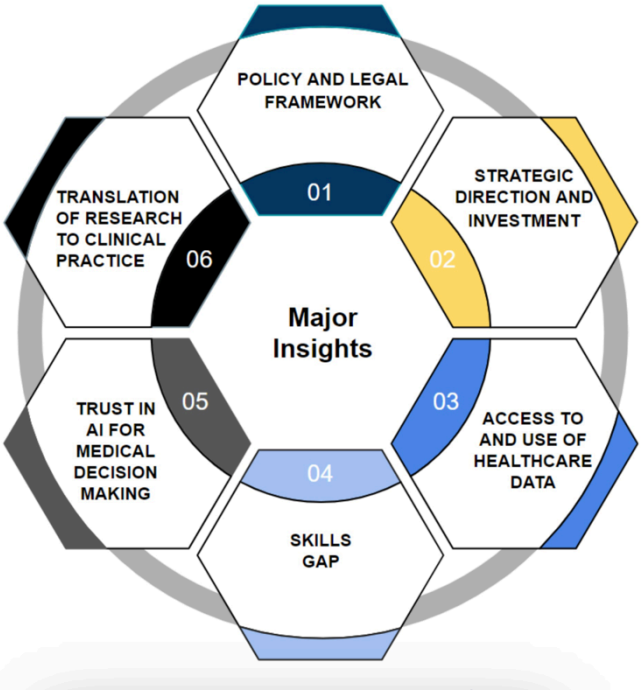
- applied to patients within the age range and disease profile for which it is intended (which is in fact no different to the appropriate use of a medicine)
- utilised at suitable points within a care pathway.
- expected to have based its advice through having access to accurate and complete background clinical information (such as a complete medication list)
- cross-checked for consistency and plausibility against other guidance and clinical practice that would be used if the AI component was not present.

Examples of responsible use at the healthcare provider level could be to:

- appropriately connect the AI component into an eHealth infrastructure
- ensure technical and semantic interoperability with relevant EHR data.
- ensure that guidance generated by the AI component and presented to professionals or patients is captured with the EHR of that patient *
- ensure that system and personnel interactions with the AI component are captured in an audit trail. *
- Both of the examples * above will assist to clarify how liability issues could be both transparent and appropriately apportioned between clinicians, providers and AI developers.

Rationale for this Round Table topic

This topic aligns with Challenge 5 to the wider adoption of AI in healthcare (actions addressing culture issues and building trust in the use of AI in the healthcare sector) of the Study on eHealth, Interoperability of Health Data and Artificial Intelligence for Health and Care in the European Union published in October 2021 (https://digital-strategy.ec.europa.eu/en/library/artificial-intelligence-healthcare-report). The diagram below shows the six challenges in that report.



A literature survey published in December 2023 (http://dx.doi.org/10.1055/s-0042-1742516),that highlights the rich availability of ethical principles governing AI development, but the paucity of specific and operational guidance for good adoption.

In 2023 AI has attracted considerable media attention in 2023 and even after our two Round Tables the pace of change has been astonishing. The introduction and uptake of ChatGPT and subsequent generative AI models and systems have led to broad discussions around its societal benefits as well as its potential risks, including its impact on productivity and jobs, in particular in the creative sector. As a consequence, this has led to a long list of governance and regulatory initiatives, such as the US Executive Order on AI Standards for Safety and Security, the G7 code of Conduct and others, as well as amendments to the AI Act.

Healthcare has been quietly using AI to assist healthcare professionals, lighten cognitive load, discover new treatments and drugs, and saves lives. This Report is stiving to present a better balance to discussions around the use of AI in healthcare.

Recommendations

The summary of the Working Group discussions in this report contains more detailed supporting information explaining the recommendations below.

1. Formalise and invest further in measures and evidence to ensure that AI is trustworthy

- The EC should promote the standardised use of terms, definitions and a common language for developers, HCP, policy makers, regulators patients and citizens including for example transparency, explainability and value.
- The EC and Member States should support further research in, and facilitate European alignment on, risk stratifications and mitigation strategies relating to data protection, information security, research governance, the quality and safety of medical devices and AI, and better alignment between these components.
- The EC and Member States should co-develop and contribute to
    - a global evidence-based evaluation framework for risk and opportunity for use cases that allows for local sensitivities such as demographics, healthcare priorities and personalisation, including ethics.
    - a global standard to define and implement explainability that meets the needs of all stakeholders including the citizen, professionals, and leaders.The standard should address explainability requirements for materially different use cases rather than one standard for every use case.
    - a global data transparency framework ensuring AI training data is fit for purpose and can be benchmarked globally, including dataset characterisation, from the start, and continues to be fit for purpose as it learns from deployment data including continuous monitoring of large language models for bias and drift.
- The EC and Member States should promote and support communities of interest and user groups to foster the development of guidance on responsible AI development and use.

2. Engage across stakeholders to communicate the basis on which AI should be trusted and adopted

- Effectively communicate the benefits (risks and rewards) of AI for the different stakeholders.
    - Develop patient & clinician informed use cases for AI, which must be capable of scaling.
    - Invest in expert use case and benefits “AI brokers” who understand and can connect developers, decision makers and potential users of AI across healthcare and research.

- Grow stronger partnerships and co-operative developments between AI developers (including computer scientists, engineers, and data scientists), clinicians, and patients.
- Effectively communicate the trust factors, and counter lack of trust that is only based on fears.
- Launch communication campaigns to citizens and HCPs from trusted brokers (like Data Saves Lives)

3. Invest in the digital, data and AI skills needed for robust and responsible AI development, well-informed decision making and responsible usage

- Invest in education skills & training for HCPs, patient organisations, policy makers, regulators, and healthcare managers.
- Develop evidence-based competency frameworks for the skills required by the health and care workforce, foster Europe-wide alignment to education interventions needed to achieve them.
- Develop an exemplar global health literacy skills framework for the citizen to empower them to engage with use of AI in their health system and other related services.
- Investigate how AI could augment the knowledge and skills of the workforce, including but not limited to HCPs and patients.

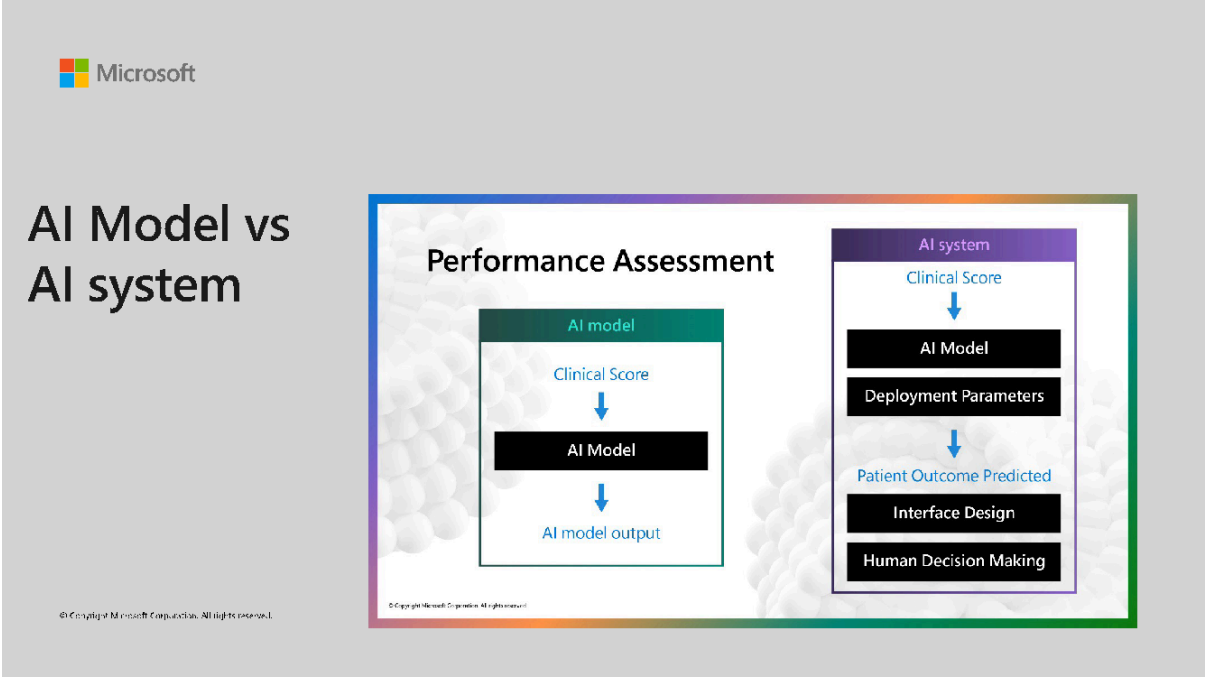
4. Promote the importance of European and Member State policy that stimulates responsible AI adoption by appropriately targeted and evidenced uptake of assured AI solutions within healthcare and clinical research

- The EC should facilitate European alignment of health systems adoption and approvals for reimbursement of AI.
- The EC should publish an annual catalogue of responsible AI evidenced use cases or maintain an online catalogue of such use evidenced use cases.
- The EC and Member States should provide greater legal clarity on AI liability to ensure this does not inhibit the development of AI healthcare products and services and provides balanced protection for developers, AI systems providers, HCP, patients and health systems.
- The EC and Member States should co-develop and contribute to a global framework for standardised but flexible evidence-based impact metrics and process evaluation for AI in health that also allows global comparison and continuous monitoring of AI performance in real world deployment.
- The EC and Member States should promote and invest more in formulative evaluation of the introduction of AI into clinical practice, in particular impact on patients and outcomes, and in the capacity of the system to deliver services, noting that:
    - funding of AI development and adoption should both include a budget for evaluation.
    - rigorous evaluation is a risk mitigation for AI.

Other key points that emerged from the Round Table

Below are important points we have extracted from the summary of the Working Group discussions.

- Quality operating management systems, procedures and practices and the strategic management of these are important trust factors.
- Clinicians need to have enough understanding of AI models and AI systems to have confidence in using them and to provide the necessary reassurance to patients ("This works I trust it") A component of this trust will be generated by operating systems, processes, and practices.
- Trust must be collectively secured involving the providers of data sets (healthcare systems or others), developers, clinicians, and AI system providers. We need a circle or Ecosystem of trust where trust is collectively built and maintained.
- No current methodology for how to assess self-learning medical devices in Europe (assessed pre and post market)
- Notified bodies are going to have skills gaps and Europe will need to plan to address these.
- Adoption of AI involves organisational change/change management and much learning around this topic exists as well as an AI case study with AZ (Mökander & Floridi Oxford Internet Institute). This involves conversations with HCP for engagement, feedback, useability as well as listening. The main conclusion of the AZ case study is the introduction of AI is essentially no different to other change management programmes, so success factors as known (e.g., leadership, clear scoping, communication, and responsibilities)
- AI systems need to be auditable, and a discussion is required with providers, industry, and regulators about the technical means to facilitate auditability.
- AI trust needs to involve a standardised risk assessment system that travels through the life cycle from ideation to use which allows the management of both risk and liability. This life cycle management identifies who manages what risk and who is responsible for controlling those risks.
- The development of standard metrics is occurring through the HTA network at the European level, but not presently through industry driven cooperation. This might be an area to explore further for potential alignment opportunities.
- AI Models V AI Systems



Summary of Key Working Group Discussions

The Round Table compromised two virtual meetings with an opening keynote from Dr Cecily Morrison from Microsoft Research Cambridge on AI Models v AI Systems. Attendees were broken into two Working Groups for each meeting followed by a plenary for the two Working Groups to feed back.

The summary discussions have been organised below under a series of headings and bullet points are in no specific order.

Potential AI Benefits for Healthcare

- Investments in AI must focus on ensuring it delivers good health systems value, especially on health outcomes.
- AI powered disease management should help improve health systems resilience, by reducing the exclusive dependence on direct patient-clinician interactions.
- Better adherence by clinicians when appropriate to published Guidelines with feedback loops (data collection) from clinicians to improve both the AI and the Guidance which may lead to improve care or AI models or new models.
- Better adherence by patients to treatment plans (reminders etc)
- Ability to quickly find latest Guidelines and other sources of clinical information.

- Standardisation of processes and care services.
- Increased quality of health services/better patient outcomes.
- Using the substantial data generated by AI to analyse workflow processes, sample flows and patient journeys to improve them.
- Efficiency gains (service and financial)
- Greater accountability for treatment decisions taken (audit trial)
- Facilitating more automated working reducing human cognitive/administrative tasks and workload of HCPs including remote working; Enabling change management in clinical working practices.
- More advanced clinical decision support systems reducing lost clinical time and reliability risks.
- Improved access to care with self-diagnosis
- Language translations for clinicians and patients
- Upskilling junior clinicians with more advanced clinical support systems
- Facilitating more personalised services with personalised AI systems (for senior clinicians in specific settings)
- With the shortage of HCP worldwide the use of AI provides an important ability to mitigate these shortages for the benefit of HCP, patients, and healthcare systems.

Trust Factors & Establishing trust.

- Transparency about the use and limitations of the AI being used (especially for more complex AI uses including what factors influence the AI decisions (including if something goes wrong what could happen so the clinician or patient is able to decide does this matter to me) and is data relating to my condition included and so relevant to me (which is more of a data issue the AI is working on)
- Explainability for patients (which should be a fundamental principle for patients) and clinicians and understanding the decision made by AI. However, this will not always require detailed information about how the black box works.
- The quality of communications about the use of AI (both transparency and explainability) are important whether to patients or clinicians or by clinicians.
- The exact information each patient will want will vary by patient and treatment type. HCPs will communicate in their own words with AI providing the necessary levels of detail and information. Alternatively, is a type of new role required to explain AI – like the experts who explain genetic testing to patients and families in the NHS?

- Clear messaging that AI is not replacing clinician but providing clinical support tools.
- Quality operating management systems, procedures and practices and the strategic management of these.
- Lack of trust leads to fear and apprehension by both clinician and Patients, so the provision of enough and clear information is a balance of what is practical in the specific circumstances (what services/treatments and understanding of the clinician/Patient)
- Clinicians need to have enough understanding of AI models and AI systems to have confidence in using them and to provide the necessary reassurance to patients ("This works I trust it") A component of this trust will be generated by operating systems, processes, and practices.
- Another component of this trust is understanding how models work and the inherent risks they have and an acceptance that mistakes or errors will be discovered and corrected. So, this is another aspect of balanced risk/reward for patients, clinicians, and health care systems. Indeed, these corrections and the risks associated with them should be viewed in the same way as drug side effects and the complications of certain treatment as explained risks and risk and reward choices.
- Trust in machine learning /AI) requires good practices in data collection, curation, calibration, model and AI system development and deployment, clinical use and associated change management of service pathways or treatments.
- Trust must be collectively secured involving the providers of data sets (healthcare systems or others), developers, clinicians, and AI system providers. We need a circle or Ecosystem of trust where trust is collectively built and maintained.
- Trust can mean different things to different people. How can developers and adopters (purchasers, funders, and end users) meet in the middle to establish trust in the use of AI solutions?
- Companies are already developing ethical practices for AI development. Could these converge to become industry wide standards?
- Industry is currently largely governing itself, e.g., through the Data and Trust Alliance (https://dataandtrustalliance.org/) setting the bar higher than current legislation, and with global consistency and is involved in the wider multi-stakeholder AI Governance Alliance of the World Economic Forum.
- Could the health sector and the AI health industry co-develop a set of practices and standards that will achieve consumer trust?
- The ability to trust and verify AI is important for acceptance, not only technically but for non-technical people to have confidence in the solution, including in the ways that the health data were used.
- The ethical use of personal health data is one important area requiring individual and societal trust.

- There is a growing opportunity to use of synthetic data during AI development, which is more readily available, has fewer data protection constraints, and could reduce data bias compared to real world data.

Use Cases & Factors

Developing the right use cases

- AI should not be a solution looking for a problem but be driven by a market pull: by scenarios of clinical decision making or patient care that are challenging today.
- AI developments must be driven by health systems needs and not by what data is readily available or by computer science challenges.
- Clinicians and patient groups should help identify the decision-making challenges and use cases for which AI could be most useful.
- Keep the use cases compact and the benefits easily measurable, to establish evidenced success, and build further use cases on top of this initial confidence.
- Powerful proof of concepts with narrow use cases are important to gain confidence (trust confidence and market confidence).
- We need to acknowledge that there are trade-offs between risk and benefits e.g. under some circumstances some risks might be worth taking because of the benefits the use case may bring.
- Value has very different meanings amongst stakeholders so any discussion about the “value” of AI needs clear communication of the benefits (rewards) and risks.

Demand side Factors

The need for demand side maturity

- The “demand side” for AI still needs maturing to drive a real market pull for AI targeted at the most needed use cases.
- A lot of digital innovation struggles to go beyond pilots and/or a single reference site. AI has that risk as well.
- Small AI projects can often be initiated and taken to a pilot stage by enthusiasts: clinical enthusiasts working with technical enthusiasts, but without the involvement of health decision makers to sustain the innovation beyond its pilot stage.

- Procurement expertise and organisational business case development skills are needed.
- Monitoring conformance will need skills training and investment.

Challenges to overcome/constraints.

- No common definitions of responsible AI and responsible use. Will these always be subjective?
- Need a balanced and proportionate risk/reward approach. But how to define “balanced” and “proportionate”?
- Too easy to generalise about AI risks and what is required is a use case by use case approach to risk.
- There will always be the usual curve to adoption and scaling as well as different appetites to risk based on behavioural elements of the end user which need to be recognised
- There is a risk that junior clinicians will always accept AI output leading to lost learning through experiences they would otherwise have had but this could be mitigated by adapting their training.
- No current methodology for how to assess self-learning medical devices in Europe (assessed pre and post market)
- Post market surveillance requires a detailed plan and important conversations are needed with providers (hospitals) and regulators about whether this requires reporting by industry on users and if so on exactly what outliers. This involves data-centric conversations about variation and clinical use understanding population error rate and the balance between false positives and false negatives and when further investigation is required. Indeed, there may be perfectly reasonable explanations for outliers. Whatever, regulation is passed all stakeholders will need an experiential period to learn and adjust AI processes, models, and systems.
- Notified bodies are going to have skills gaps and Europe will need to plan to address these.
- AI development and use involves a multi-disciplinary team which needs to be cohesive but who will be the leader and accept responsibility for promoting and deploying AI?
- Resistance maybe the wrong term as people may not resist the change but the challenges may arise from lack of useability or not facilitating fully the change (e.g., training, and protected time). Also pace of adoption is a better term to use.
- AI is capable of upskilling HCP so not necessarily a threat (deskilling and radiologists for example where there is evidence that they are not threatened)
- Training data sets to reflect local populations maybe a challenge.
- HCP may want to use AI models and systems but the barrier maybe their IT systems or procurement barriers.

- We need to remember the future of how AI will develop is uncertain so not possible to forecast with any real clarity or exactness now.
- To date AI has not created structural organisation changes but autonomous AI could or will (e.g., remote monitoring and guidance to patients).
- The distinction between autonomous and not autonomous AI is not binary and there is a sliding scale already between them (e.g., parts of tasks may already be autonomous) Autonomous AI needs to be positioned as freeing up HCP time to allow them to undertake more valuable work. This positioning needs to be nuanced. This is not a cultural revolution for the workforce as already other technologies only have HCP involved at escalation points. See the diagram below on levels of automation which illustrates the distinction in more detail.
- Fully autonomous AI (as opposed to humans always in the loop) raises both regulatory and legal issues outside the scope of the Round Table.
- Research is needed to investigate how AI could augment the knowledge and skills of others not just HCP including patients.

Adoption challenges & pace of adoption

- AI solutions need now to be delivered at scale, and accepted by users – which emphasises again the importance of trust, and communication to the public and professionals about the basis for that trust.
- We need communication campaigns for healthcare professionals and patients, undertaken by neutral parties, to explain the benefits for everyone including easing the healthcare professional burden, appointment backlog etc.
- Advocating for more AI in health systems is politically challenging: there is still a reluctance for decision makers to promote solutions that might deflect from the primacy and prioritisation of doctor contact time.
- Adoption of AI involves organisational change/change management and much learning around this topic exists as well as an AI case study with AZ (Mökander & Floridi Oxford Internet Institute). This involves conversations with HCP for engagement, feedback, useability as well as listening. The main conclusion of the AZ case study is the introduction of AI is essentially no different to other change management programmes, so success factors as known (e.g., leadership, clear scoping, communication, and responsibilities)
- Radiology has been using AI for some time so is probably relatively mature and the most developed use case to learn from. AI in radiology is a complimentary tool and the pace of adoption will vary from use case to use case. Skilful introduction and management is required to achieve optimal pace of adoption by use case.

- Is there some pattern recognition between the introduction of AI and nurse practitioners in terms of workforce impact (where doctors were concerned their jobs might be at risk or their role diminished, or patient safety jeopardised)? These concerns have not been realised and many benefits realised from multi professional teams. This provides opportunities for us to learn how to introduce AI and accelerate the pace of change.
- Change management expertise is needed which faces both into healthcare organisations and AI developers.

Education, Skills & Training

- Education about how to use AI responsibly and safely should now form part of professional qualifications and/or continuing education requirements.
- Educating & training for clinicians about AI should include:
    - AI in general (benefits, use, not replacing professional knowledge but enhancing it and AI as one source of evidence for clinicians to consider not as the single source of truth)
    - AI Models and how they work
    - Limitations of AI Models
    - AI Systems and the importance of the necessary human decisions (see later proposed definition of AI Systems in the Need for Agreed Terms section)
    - Explainability
    - How Clinicians should use AI and work with AI
- Education and information is needed for patient organisations and patients including:
    - AI in general
    - How AI is being used in Health care
    - AI systems and the interaction between the technology and humans/clinicians.
    - Regulatory protections on the use on AI in healthcare
    - Limitations of AI models
    - Explainability
- Significant investment in education and training will also be required for the developer workforce in order to uphold the Regulation and ethical principles.
- Similar training as above for policy makers, regulators and health service managers is needed.

Assurance (assessors, regulators, adopting communities and providers)

- There needs to be traceability of data sets used to create AI Models and Systems which requires technical standards as well as human decision making.
- There are a plethora of risk stratification and mitigation strategies issues for data protection, information security, research governance, medical devices and now AI. There is a need for better alignment between these components or there is a risk that end user confidence is undermined.
- AI evaluation needs to be dynamic and continuous not static (at one point in time) as AI systems continuously learn, adapt, and evolve.
- AI systems need to be auditable, and a discussion is required with providers, industry, and regulators about the technical means to facilitate auditability.
- AI trust needs to involve a standardised risk assessment system that travels through the life cycle from ideation to use which allows the management of both risk and liability. This life cycle management identifies who manages what risk and who is responsible for controlling those risks.
- Patients and decision makers need to be certain that an AI component/solution is safe, has been tested and verified, and meets standards such as the Medical Device Regulation.
- Purchasers want certainty about the safety and effectiveness of novel technology, ideally through certification.
- Regulatory and HTA assessment of AI solutions would have a powerful influence on adoption confidence. European alignment is needed on AI approvals.
- A mixed model of self-regulation, co-opetition, legislation, and policy will be required for AI solutions, for some years ahead.

Need for agreed Terms

There is a need for agreed definitions of commonly used terms. The following terms were discussed and proposed definitions created.

Transparency: How the AI system is working. More thought is needed as to what exactly amounts to transparency from the perspectives of developers, clinicians, patients, and regulators (what is proportional, practical, and reasonable)

Explainability: Understanding why the decision was made (which involves factors such as counter factual and quantifying the uncertainty in the system)

AI Model: is an algorithm trained on data that predicts an outcome.

AI System: is all the elements that come together to create the experience of the end-user including the AI model and the human decisions about a) the deployment parameters (e.g., how many false positives are acceptable), b) how to display the results to clinicians and c) the final clinical decision to change a patient's treatment or not.

Levels of automation of medical artificial intelligence systems

	Assistive AI algorithms		Autonomous AI algorithms		
	Level 1	Level 2	Level 3	Level 4	Level 5
	Data presentation	Clinical decision-support	Conditional automation	High automation	Full automation
Event monitoring	AI	AI	AI	AI	AI
Response execution	Clinician	Clinician and AI	AI	AI	AI
Fallback	Not applicable	Clinician	AI, with a backup clinician available at AI request	AI	AI
Domain, system, and population specificity	Low	Low	Low	Low	High
Liability	Clinician	Clinician	Case dependent	AI developer	AI developer
Example	AI analyses mammogram and highlights high-risk regions	AI analyses mammogram and provides risk score that is interpreted by clinician	AI analyses mammogram and makes recommendation for biopsy, with a clinician always available as backup	AI analyses mammogram and makes biopsy recommendation, without a clinician available as backup	Same as level 4, but intended for use in all populations and systems

From: Bitterman D, Aerts H, Mak R. Approaching autonomy in medical artificial intelligence. The Lancet 2020:2(9);E447-E449. https://doi.org/10.1016/S2589-7500(20)30187-4

Responsible AI use

There have been a lot of publications, guidelines and more recently legislation specifying the measures required to ensure responsible AI development, including ethical principles, risk assessments and mitigation measures, transparency and explain ability, the importance of checking for data bias and for algorithm drift, as examples. We believe this prior work should be used to create a definition of "Responsible AI Development".

In contrast, little has been written to date about the responsibility of users of AI solutions.

- Responsible AI Use combines
    - how clinicians use the AI data/prediction
    - parameters set by clinician tools for clinicians (e.g., for visualisation, debugging models, data isolation and filtering) with Responsible AI Developed models.
- Responsible co-development and use includes making sure of clinical utility and that the user experience is good.
- AI solutions must be embedded within the care and provider organisational ecosystem and models of care, respecting the users and the impact on the clinical workflows.
- A responsible AI solution must help the end user to perform better or provide convenience.
- For example, an AI powered diabetes system can take a psychological pressure off the patient by managing the hourly and daily stability of their blood glucose.
- An AI powered device can provide a patient with permanent access to specialist expertise that would not be realistic to provide through a clinician specialist.
- AI ideally demonstrates how AI and humans can maximize performance together.

Economics & benefits evidence

- AI seems to be very competitive at present.
- The balance between user needs and commercial interests will need to be carefully managed.
- A bottom-up co-operation on metrics and evaluation methods does not appear to be taking place, but would be desirable for transparency.
- The development of standard metrics is occurring through the HTA network at the European level, but not presently through industry driven cooperation. This might be an area to explore further for potential alignment opportunities.

The importance of partnership

- AI innovations should ideally be a co-creation between developers and clinician/patient users.
- Co-creation partnerships take time and effort to build.
- An iterative dialogue with the people on the ground requires the developer team to be broader than pure AI scientists.
- Activities like clinical workflow adaptation and organisational change requires specialist skills and dedicated resources.

AI Liability

A strong consensus, if not unanimity, that the question of liability is very complicated and important.

Trust in AI is key and who is liable if things go wrong is an important success factor in AI being embraced by Clinicians, healthcare providers and patients. What are the considerations that need to be discussed or addressed to resolve this challenge or is it so complicated it may be necessary to have a case-by-case approach?

The Round Table discussed:

- Traditionally the liability for healthcare services was clear but the increasing use of technology has muddied the waters especially the introduction of smart tools and AI.
- Liability conversations need to be open and honest reflecting the genuinely collaborative nature of the development of health AI models and systems.
- Who could be liable? Potentially:
    - Software developer
    - Data set providers and any data curators including clinicians?
    - AI system seller
    - Providers who use the AI or AI system
    - Procurers?
    - Regulators?
- The interaction between these stakeholders and professions and/or dependencies complicates the issue further. Shared liability between the parties listed above is likely. It is likely that over time organisations that contract for collaborative work will seek to list their respective responsibilities and liabilities at least to each other.
- What Liability? Liability for negligence and personal injury (tortious) or civil liability (product and services)? Risks will vary between research and everyday clinical services to patients.
- A very expansive collection of expertise is involved in AI, a true multi-disciplinary team (of stakeholders) involved in developing and using AI and substantial interactions and dependencies between them. Industry itself needs to provide good communication internally between its own multi-disciplinary staff to ensure the different areas of knowledge are applied to the question of liability.
- Any discussion about liability needs all the right stakeholders around the table to ensure the right questions are discussed.

- Everyone needs clarity on liability apportionment but unless strict liability (where no fault needs to be proved just that harm has happened) is imposed for AI liability then proof is required that the AI has caused harm which is a question of fact based on expert evidence.

- No concluded view on whether the lack of clarity about liability is hampering the development of AI. Without clarity there is a risk some AI will not be deployed. Likely speed of deployment and/or types of AI deployed may vary geographically commensurate with legal risks (scale of damages risks and whether liability is covered by commercial insurance or public sector risk pools like the NHS).

- AI is characterised by continual development and learning so evaluation too must be dynamic not static. A strong emphasis is need on providing the tools for clinicians to have oversight of this continual development both pre and post market. Another complicating factor is AI models developed by one team of developers may be adapted or developed further another team of developers especially if open-source coding is used.

- There was support for the development of communities of practice and user groups which in turn could foster the development of Guidance on Responsible AI and Use and assist with building and maintaining trust in AI use by end users.

- An important requirement is the development of Testing facilities for AI experimentation and testing. In Europe the EC issued a tender for 5 different sector testing and experimentation centres and for healthcare the winning consortium is led by CharitéIt will develop a single EU-wide entry point through Ebrians.eu and will provide a range of services including physical and virtual testing, education, standards and quality, certification and legal. Another research project relates to digital twins where there may again be helpful knowledge in due course.

- Sandbox environments generally are important for AI use cases to be demonstrated on real or synthetic data.

- Legal analysis of the liability issues is beyond the scope of this Round Table. The EU has created a tripartite model for the regulation of AI. Firstly, AI Act is aimed at ensuring AI is developed safely. Secondly, proposed amendments to the Product Liability Directive which provides for strict liability for physical harm extended to cover software including AI systems and data loss harm. Thirdly, the proposed AI Liability Directive provides for common rules for
  fault-based liability regime for damage caused by AI including in certain circumstances a rebuttal presumption that a defendant breached a relevant duty and another rebuttable presumption of casual link between defendant fault and act or omission of AI system giving rise to damage.

Glossary

AI Act: The Artificial Intelligence Act of the European Union (2024/1689(COD))

AI: Artificial Intelligence

Artificial Intelligence System:

1) The draft AI Act had the following definition- software that is developed with one or more of the techniques and approaches listed in Annex I and can, for a given set of human-defined objectives, generate outputs such as content, predictions, recommendations, or decisions influencing the environments they interact with;

Annex 1 lists:

(a) Machine learning approaches, including supervised, unsupervised and reinforcement learning, using a wide variety of methods including deep learning.

(b) Logic- and knowledge-based approaches, including knowledge representation, inductive (logic) programming, knowledge bases, inference and deductive engines, (symbolic) reasoning and expert systems.

(c) Statistical approaches, Bayesian estimation, search and optimization methods.

NB This was the definition from the draft AI Act and used in the Round Table discussions.

2) The AI Act subsequently to the Round Table meetings uses the following revised definition;

An AI system is a machine-based system designed to operate with varying levels of autonomy and that may exhibit adaptiveness after deployment and that, for explicit or implicit objectives, infers, from input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments.

HCP: Health Care Professional

HTA: Health Technology Assessment

MDR: Medical Device Regulation

Contributors

Name	Organisation
Dipak Kalra (Co Lead)	The European Institute for Innovation through Health Data
Bleddyn Rees (Co Lead)	The Digital Health Society
Dr Richard Baumartner	MSD Biomedical Group
Maria Christofidou	The European Institute for Innovation through Health Data
Dr Kathrin Cresswell	University of Edinburgh
Alan Cross	Diveplane
Javier Rinćon Cruz	MSD
Elisabetta Gatti	Johnson & Johnson
Gordon Johnston	Johnson & Johnson
Adrian Jonas	NHS England
Nathan Lee	The European Institute for Innovation through Health Data
Karolina Mackiewicz	European Connected Health Alliance
Jelena Malinina	EURORDIS
Dr George Margelis	Chair Aged Care Industry Information Technology Council (Australia)
Stuart McCann	Roche
Dr Cecily Morrison	Microsoft Research Cambridge
Daniel Corredera Nadal	Johnson & Johnson
Professor Tudor Pitulac	Opensky Data Systems
Richard Rak	Digital Europe
Hannah Richardson	Microsoft
Tjade Stroband	Microsoft
Michael Strübin	Digital Europe
Dr Anja Thieme	Microsoft Research Cambridge
Danny Van Roijen	Agoria

the digital health society
The data arm of ECHAlliance Group

iHD

Creating an environment of Trust and responsible AI adoption and Use

2023 RECOMMENDATIONS BASED ON CALLS TO ACTION ON HEALTH DATA ECOSYSTEMS

THIS INITIATIVE IS SUPPORTED BY

Johnson & Johnson

Microsoft

MSD