

ACTIONABLE RECOMMENDATIONS FROM THE  
EIT HEALTH THINK TANK ROUND TABLE SERIES

---

# **Liability and managing risk**



## If AI applications are not explainable with full data transparency, then people are less likely to trust their outcomes.

Round Table participants endorsed the findings of the EIT Health and McKinsey & Company report that one of the most common questions that arises in relation to liability is: who is responsible if the algorithm prediction is wrong – the hospital, the doctor, the researcher, or the company which developed the AI tool?

AI applications in health are intended as decision support systems, so it was agreed that the ultimate decision rests with the clinician.

In addition, relating to liability and risk, there are a range of different challenges that will require different approaches. For example, those that an AI system presents because it is poorly implemented or governed (such as privacy issues) versus issues that are inherent to a more automated system (such as robustness and cybersecurity) and challenges that are inherently human, such as oversight and how to work with AI systems in order to achieve the best possible outcomes.

This highlights the complex liability issues linked to AI implementation in healthcare and the need for standardisation and agreement.

It was suggested that AI systems should be given an appropriate legal context with clear direction on whether it is obligatory to make decisions based on the information generated by AI systems that are used to support the human decision-making process. It is just as important to think about the impact of not using an algorithm as much as the impact of using it.

**“ A framework to work within that would protect healthcare professionals and organisations from reputational, legal and financial risk when using AI ...that would make people feel a lot more comfortable. ”**

**Mark Kelly, Director & Chief Customer Officer, Allidus International Consulting**

The need for ‘explainable AI’ in relation to healthcare – understanding how the AI application has reached its outcome or decision, as opposed to the ‘black box’ situation where the input and output from the system or program can be seen, but there is no insight into the processes and workings in between and how the output has been arrived at – was also a subject of considerable debate in relation to liability and risk.

In general, Round Table participants considered that if AI applications are not explainable with full data transparency, then people are

less likely to trust their outcomes. In addition, from a regulatory standpoint, AI algorithms and the data pipeline need to be auditable. At the Round Table Meeting in Spain it was suggested that safety-by-design – the integration of risk assessment methods early in the design process in order to eliminate or minimise harm – and ethics-by-design – the integration of ethics into the design process from the beginning – should be incorporated into the development of AI applications from the very beginning.

AI algorithms also need to be trustworthy. In order to support

accurate clinical decision-making, the datasets that algorithms use for learning must be sufficiently large and representative of the population being tested (demographics, gender, etc.). Both Round Table Meetings in France and Germany suggested that once deployed in clinical settings, long-term monitoring of AI applications should be undertaken to enable ongoing assessment of the benefit/risk balance. Recording and tracking error states should be automated through logging and traceability concepts in order to gradually reduce possible application risks.