WHEN AI GOES WRONG

Speakers: Stéphanie ALLASSONNIÈRE, Professor and Vice President at the Université Paris Cité | France, **Jean-Emmanuel BIBAULT,** Professor of Radiation Oncology at the Georges Pompidou European Hospital - HEGP | France, **David GRUSON,** Cofounder of EthikIA and Healthcare Director of Luminess | France, **Arnaud ROSIER,** CEO and Cofounder of Implicity | France

Debate led by **Alix PRADÈRE,** Managing Director Health France at Accenture | France

In the healthcare sector, Artificial Intelligence (AI) relies on two types of data

Depending on its type, AI does not require the same amount of data. AIs that require small amounts of data rely on statistical models, whereas those that require a large amount, such as Chat GPT, perform much better thanks to deep learning (an advanced AI learning technique). In the medical field, a major challenge lies in creating frugal and high-performing AIs with reduced databases given the sheer volume of available data in the field. In medicine, it is possible to identify two sorts of data:

- Clinical data, such as genome and imaging data. These are generally of good quality, easily extractable and exploitable.
- Data created from human classifications, which are less formal and therefore less structured, such as the GCS (Glasgow Coma Scale, an indicator of the level of consciousness) or the GIR (a French indicator of the degree of dependency).

AI assistance delivers more effective healthcare and frees up medical time

When AI draws on many different types of data (e.g., genetic and biological data coupled with medical imaging), it can outperform humans in identifying potential pathologies. While a doctor can recognize up to 60 rare diseases in a foetus, AI can diagnose 500 of them. AI is therefore essential in helping doctors make the right decisions and avoid errors.

AI is also crucial in rehumanizing medical practice. By relieving healthcare professionals of time-consuming technical tasks, it gives them more time to dedicate to patient care. For instance, at the Georges Pompidou European Hospital, the manual preparation and calibration of a radiotherapy session used to take two to three hours. Algorithms have reduced this time to two to three minutes.

AI must be controlled, but regulation should not hinder its development

AI is the fruit of a probabilistic science based on modelling and is therefore inherently a source of error, even if it tends to get closer to reality. Furthermore, the use of data introduces new risks of errors related to the nature or quality of the collected data.

Adhering to the Human Oversight ("*Garantie Humaine*") ensures the protection of citizens and limits AI's potential for errors. This regulatory principle helps avoiding ethical drifts by keeping humans at the centre of the decision-making process, by ensuring that the final decision is always taken by humans (i.e., doctors) and by guaranteeing that AI is supervised by representatives of healthcare professionals and patients. It has been incorporated into the 2021 French Bioethics Law and will also be included in the 2025 AI European Regulation.

However, overly restrictive regulations could impede the development of AI. While fast tracks have been implemented in the USA to validate algorithms quicker, the trend of adding regulatory constraints in Europe could limit innovation in the area.

If AI is not well-developed in Europe, two problems could arise:

- Europe would lose sovereignty and the ability to protect its data.
- AI developed outside Europe might not be designed for local populations.

The population's awareness and education on such tools need to be increased to ensure that AI is not mistrusted and that the usefulness, benefits and risks of health data are understood. While AI is a source of error, so are humans. And yet, no Human Oversight principle is applied.

Predictive AI is going to disrupt the field of healthcare

AI-powered productivity tools are widely used and support doctors in performing technical tasks such as analyzing electrocardiograms or imaging data. However, the development of predictive AIs will enable to perform tasks that the practitioner is unable to do alone such as anticipating the onset of disease or the chances of recovery. As of today, predictive AI models are already capable of assessing the risk of hospitalization for patients suffering from heart failure.

The emergence of such technologies raises the question of liability. When using AI for productivity, the ultimate responsibility lies with the practitioner who verifies the results produced by the tool. The same task cannot be easily carried out when using predictive AI. As a result, liability in the event of an error will be more unclear, and the principle of Human Oversight will no longer apply. Will responsibility in the event of an error lie with the software's developer, the hospital that purchased the software, or the doctor?