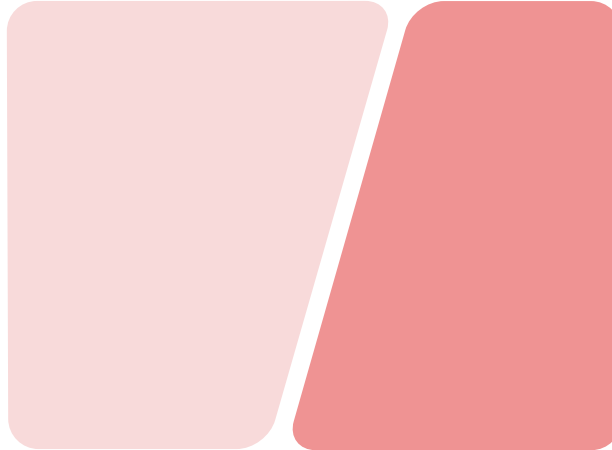


Views and Opinions

Legal perspective



AI in clinical research: Regulatory and legal uncertainties in protecting participants' rights

Author: Audrey Lebreton

Affiliations: Associate Professor of Biolaw and Digital Health at the Faculty of Law, Criminal Sciences and Public Administration (FDCA) and the Institute of Humanities in Medicine (IHM) at the Faculty of Biology and Medicine (FBM) of the University of Lausanne (UNIL) and Co-founder of Lab' Santé et Droit

doi: [10.54920/SCTO.2026.RAWATCH.11.28](https://doi.org/10.54920/SCTO.2026.RAWATCH.11.28)

In the context of clinical research, machine learning algorithms can improve participant recruitment, enhance patient safety, accelerate clinical trials, and reduce associated costs. However, the growing integration of artificial intelligence (**AI**) also raises a number of legal and ethical concerns, including issues related to transparency, data quality, algorithmic bias, and the protection of participants' consent. Moreover, the development and deployment of AI systems raise complex questions regarding the allocation of responsibility among the various actors involved in the AI development chain. The current legal framework remains fragmented and does not provide a comprehensive response to all these challenges. This fragmentation may, to some extent, undermine the coherence of AI regulation in clinical research. Normative uncertainties also arise where the protection of patients' and participants' rights, including data protection and non-discrimination, requires reconciling competing legally protected interests. These issues are illustrated by the regulation of health data and by the challenges posed by algorithmic bias in clinical research.

Legal and regulatory landscape surrounding AI in clinical research

The legal and regulatory framework governing artificial intelligence (AI) in clinical research is fragmented in both Europe and Switzerland. While the European Union (EU) has taken an important regulatory step by adopting the [Artificial Intelligence Act \(AI Act\)](#), the specific challenges posed by medical AI are not addressed in a dedicated framework. The AI Act must therefore be interpreted alongside other risk-based instruments such as the EU's [Medical Device Regulation \(MDR\)](#) and the [European Health Data Space \(EHDS\) Regulation](#) as well as rights-based instruments such as the EU's [Charter of Fundamental Rights](#).

Although Switzerland has legislation that adapts the MDR framework to its domestic legal order, it has not mirrored the EU approach for AI and does not currently plan to adopt a comprehensive regulatory framework for artificial intelligence. However, Switzerland recently signed the first international treaty on AI and human rights, which will also apply to health care: the [Council of Europe Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law](#). This treaty requires contracting states to implement concrete measures to safeguard fundamental rights such as autonomy, non-discrimination, and digital literacy. While this international instrument establishes general principles, it may eventually be complemented by more specific domestic regulations, which remains an option for medical AI in Switzerland.

Beyond AI-specific regulatory frameworks, the field remains subject to a range of pre-existing general legal norms. Yet while the regulatory debate primarily focuses on how to regulate AI prospectively, certain gaps and shortcomings have already arisen from these existing legal frameworks, with potential implications for the rights and interests of patients and research participants. One important issue concerns the capacity of current frameworks to effectively prevent algorithmic bias that may lead to discriminatory outcomes.

Risks of bias and discrimination

The risks of biases in AI systems are now well documented and have multiple sources.^{1,2} In the context

of clinical research, biased datasets may emerge from long-standing research practices, such as the quasi-systematic exclusion of certain categories of participants. For instance, people over 65 years old are regularly excluded from clinical trials, so they are often under-represented in Alzheimer's trials.³ In addition, women have historically been under-represented in many clinical studies due to study exclusion based on gender-specific factors such as pregnancy, childbearing age, and menopause.⁴ Such disparities in participation can lead to biased training data and consequently to qualitative disparities in algorithmic predictions. When these inequities amount to unjustified differences in treatment based on legally protected characteristics, they violate the principle of non-discrimination (regardless of whether discriminatory intent existed) and potentially constitute direct discrimination.

However, risks of algorithmic biases are not limited to situations in which legally protected groups are explicitly excluded. They may also arise from the definition of objectives, such as longevity matching in healthcare resource allocation algorithms, which may disproportionately affect older patients, chronically ill individuals, or persons living with disabilities.⁵ These risks often stem from political and design choices made upstream of the deployment of machine learning systems.

Whether biases result from a data quality problem (the well-known "garbage in garbage out" issue) or from design choices, they may be partly mitigated by increasing fairness awareness and developing fairness metrics. The EU's AI Act, for example, requires that particular attention is paid to data governance, including data selection and preprocessing. However, other legal standards designed to protect individuals may sometimes affect how inclusive and representative health datasets can be obtained to train fair AI systems. In such situations, legal trade-offs may be made, potentially leading to new forms of bias.

Data protection, consent models, and selection bias

Respecting and protecting the various rights of patients and research participants may generate tensions between competing rights, ultimately

requiring balance and compromise. The legal framework governing the collection and use of health data, which constitute personal data protected by law, provides a good illustration of these tensions. In Switzerland, the development of AI systems within the context of clinical research falls within the scope of the Human Research Act (HRA), which applies to health-related personal data unless they have been effectively anonymised (HRA, Art. 2).

As a general rule, individuals participating in a research project must be informed about the nature, purpose, and duration of the project as well as its foreseeable risks and expected benefits, and their informed consent must be obtained (HRA, Art. 16). When research is conducted on data that were originally collected for a different purpose (i.e. in the case of secondary or further use of data), consent must be obtained at the time of the original health data collection and individuals must be informed of their right to object (HRA, Art. 17). In addition, Article 33 of the HRA provides two possible legal bases for the reuse of non-genetic health data: *explicit consent* (informed consent), in which case the data may remain uncoded, and *implicit consent* (when the person concerned has been duly informed and has not objected), in which case the data must be coded. The further use of genetic data, by contrast, requires informed consent regardless of whether the data are coded or not (HRA, Art. 32). The concrete modalities of consent and the scope of information to be given to participants are further specified in the Human Research Ordinance (HRO), Chapter 3).

While these provisions are essential for protecting individuals' self-determination, they may also introduce selection biases depending on the consent model used. Empirical evidence suggests that participation patterns differ across consent models. In particular, opt-in mechanisms may lead to the over-representation of certain population groups depending on their linguistic, age-related, or socio-economic characteristics.⁶ More broadly, systematic reviews have shown that different consent procedures can significantly affect participation rates and introduce consent-related selection bias in research datasets.⁷ Such disparities in model training, which result from strong protections of participant autonomy, may therefore reduce the volume and diversity of data available for training machine learning

systems, especially for groups that are already under-represented. This can potentially lead to inequitable outputs.

Challenges of algorithmic decision-making

These legal trade-offs reflect the necessary compromises that legal and regulatory frameworks must strike. Yet beyond considering these trade-offs, legal scholarship has only modestly begun to address the deeper conceptual challenges posed by algorithmic decision-making. This is particularly evident in situations where algorithms identify correlations and construct algorithmic groups to which harmful decisions may subsequently be applied. This may occur when AI systems classify individuals as dog owners, sad teens, or single parents in order to allocate resources.⁸ Individuals may then suffer harm not because they belong to historically protected groups but because they are classified within newly constructed algorithmic categories. Such forms of algorithmic harm may escape traditional anti-discrimination frameworks because legislation is primarily designed to protect groups that have been historically stigmatised in the social world.⁶ While legal protections must remain attentive to social realities, the increasing digitisation of health care and the opacity of many algorithmic systems suggest that legal and regulatory frameworks should also adapt to these emerging algorithmic realities and develop tools capable of addressing new forms of algorithmic harm.

Conclusion

The challenges presented by AI in clinical research do not stem solely from the absence of AI-specific regulation. They also reveal tensions and limitations within existing legal frameworks, particularly where the protection of autonomy, data governance, and non-discrimination intersect. As AI systems become increasingly embedded in health research, paying greater attention to these tensions and to the interpretation of rights in the digital context may help ensure that innovation and the protection of research participants' rights evolve in a mutually reinforcing manner.

References

- ¹ Barocas S and Selbst AD (2016) Big data's disparate impact. *California Law Review* 104:671–732. doi: [10.2139/ssrn.2477899](https://doi.org/10.2139/ssrn.2477899)
- ² Obermeyer Z et al. (2019) Dissecting racial bias in an algorithm used to manage the health of populations. *Science* 366(6464):447–453. doi: [10.1126/science.aax2342](https://doi.org/10.1126/science.aax2342)
- ³ Banzi R et al. (2016) Older patients are still under-represented in clinical trials of Alzheimer's disease. *Alzheimer's Research & Therapy* 8(32). doi: [10.1186/s13195-016-0201-2](https://doi.org/10.1186/s13195-016-0201-2)
- ⁴ Criado Perez C (2019) *Invisible women: Exposing data bias in a world designed for men*. New York: Abrams Press.
- ⁵ Lebet A (2023) Allocating organs through algorithms and equitable access to transplantation: A European human rights law approach. *Journal of Law and the Biosciences* 10(1):lsad004. doi: [10.1093/jlb/lsad004](https://doi.org/10.1093/jlb/lsad004)
- ⁶ Bosmani C et al. (2023) REPRESENT: REPresentativeness of RESearch data obtained through the "general informed consENT". *BMC Medical Ethics* (24)10. doi: [10.1186/s12910-022-00877-7](https://doi.org/10.1186/s12910-022-00877-7)
- ⁷ de Man Y et al. (2023) Opt-in and opt-out consent procedures for the reuse of routinely recorded health data in scientific research and their consequences for consent rate and consent bias: Systematic review. *Journal of Medical Internet Research* 25:e42131. doi: [10.2196/42131](https://doi.org/10.2196/42131)
- ⁸ Wachter S (2022) The theory of artificial immutability: Protecting algorithmic groups under anti-discrimination law. *Tulane Law Review* 97:149. doi: [10.2139/ssrn.4099100](https://doi.org/10.2139/ssrn.4099100)