

Feedback From swissethics



Ethical challenges related to artificial intelligence in clinical research

Author: Pietro Gervasoni

Affiliations: Managing Director at swissethics

doi: [10.54920/SCTO.2026.RAWATCH.11.21](https://doi.org/10.54920/SCTO.2026.RAWATCH.11.21)

While existing ethical principles and regulations are fully applicable to clinical research on or using artificial intelligence (**AI**), research in this field introduces new aspects that challenge the current standard approaches to the oversight, accountability, and governance responsibilities of researchers, sponsors, and ethics committees. In recent years, several frameworks and guidance documents have appeared that offer practical, decision-oriented tools to support researchers, sponsors, and ethics committees in their roles. Yet given the dynamic nature of developments in AI, it can be difficult to maintain a clear overview of AI's ethical implications in clinical research. This article highlights the most important ethical challenges related to AI and clinical research.

Preliminary question: Research focus or tool?

To begin with, we must first differentiate between:

1 artificial intelligence (**AI**) as the subject of an investigation *in research* (i.e. an AI system is the focus of research and is the intervention under investigation, an AI system is investigated systematically and intended to contribute to generalisable knowledge, etc.) and **2** AI used as a tool *to administer research* (i.e. AI supports the operative aspects of research and is used, for example, to draft the research protocol, optimise patient recruitment, enhance safety monitoring, etc.).

Answering this preliminary question is critical in order to determine whether a given clinical research activity involving AI classifies as research under Switzerland's Human Research Act (HRA).

When AI is the focus of a trial and an intervention under investigation, ethics committees evaluate, among other aspects, the AI's provenance (i.e. the documented history, origin, and lifecycle of the AI system, including the data, algorithms, and hardware used to create and operate it) as well as its intended use, validation status, and risk profile. These topics will be addressed in more detail in the following sections. When AI is limited to operational support, approval by an ethics committee for that specific use is not required under Swiss law, but the ethical challenges remain. Researchers and sponsor(s) of a clinical trial are ultimately responsible for proper governance, fairness/equitability, and transparency as well as human oversight.

In practice, AI is frequently used in activities that straddle the boundary between research, quality improvement, and operational analytics – which raises different ethical questions and concerns than in clinical research without AI. Many frameworks, guidelines, templates, and other resources have been created to help researchers, sponsors, and ethics committees approach AI in clinical research (see **Resources and further reading** at the end of this article).

Ethical questions and concerns

When AI systems are the focus of clinical research, ethical issues, use practices, and regulatory status evolve as the AI systems move from conceptual development to real-world clinical use.

In the *discovery stage*, AI systems are typically trained and developed using existing datasets. During this early stage, ethical questions and concerns are related to data provenance, whether data was originally collected with appropriate consent, whether secondary data uses are ethically and legally justified, and study participants' identifiability. Given modern re-identification techniques, participants' data, even if nominally anonymised, may remain linkable to their IDs. Today, there is general agreement that we need to move beyond binary notions of "identified" versus "anonymous" data and instead assess the realistic risk of linkage.

In the *translation stage*, researchers focus on validation, performance assessment, and risk characterisation. An AI system's intended use must be clearly specified and aligned with the objectives of the research. Researchers must also consider whether future applications might fall under the scope of the Medical Devices Ordinance (MedDO) or the Ordinance on In Vitro Diagnostic Medical Devices (IVDO). Performance metrics – such as accuracy and precision, sensitivity and specificity, usability, etc. – are not merely technical or operational matters but ethical indicators of potential harm, bias amplification, injustice, and inequitable impact.

The *deployment stage* represents the highest risk context. During this stage, AI systems directly influence clinical decision-making, diagnosis, treatment, and/or study participants' outcomes. Researchers and sponsors must ensure high-level human oversight, real-time monitoring, clarification of the regulatory status of AI (e.g. software as a medical device), and procedures for responding to performance drift. AI systems used as companion diagnostics or clinical decision support tools are both high value and high risk, and they require ongoing accountability and oversight rather than merely one-time (initial) ethical approval.

These three stages can certainly coexist in one research project.

Algorithm stability and adaptive learning

Traditional clinical trials are event-based and built around research protocols with static interventions. AI systems, by contrast, may continuously learn, update, or adapt in response to new data, and researchers and sponsors must define if they are “locked” or if they are adaptive algorithms. Adaptive algorithms introduce the possibility that risk profiles may change during the course of the research. In this case, researchers and sponsors must predefine limits to the changes allowed as well as define monitoring protocols and triggers for audits, rules for rollbacks, and triggers for amendment submissions. Researchers and sponsors must embed these elements into their research protocols.

Data identifiability, governance, and secondary use

With AI research, attention must be given to the increasing feasibility of re-identification, especially when large datasets are combined or processed using generative AI (**GenAI**) and large language models (**LLMs**). Ethics committees must assess not only whether data protection legislation and standards are being complied with but also how likely it is that someone could re-identify study participants. Safeguards – such as data coding, secure IT environments (e.g. through the use of the Swiss Personalized Health Network’s (**SPHN’s**) [Excel Use Case Evaluation and Risk Assessment Template](#)), and contractual restraints (e.g. data transfer and use agreements (**DTUAs**)) – are essential elements of ethical AI research.

Attention must also be given to the ethical issue of data retention when participants withdraw from studies. Once data are incorporated into trained AI systems, removal of those data may no longer be technically feasible. This must be transparently disclosed to study participants during the informed consent process.

Additional ethical considerations

There are specific domains that require caution in clinical research when AI is the subject of an investigation: technical robustness and safety, confidentiality and data governance, transparency, representativeness and fairness, and informed consent. In all these domains, it is crucial to emphasise the importance of preserving human judgment. AI outputs should inform, not replace, clinical decision-making, and research protocols should specify how researchers and treating clinicians can override or question AI recommendations. Having training requirements, the ability to understand and explain decisions, and accountability are ethical safeguards rather than optional best practices.

Fairness and representativeness must receive sustained attention. It is evident that AI systems trained on unrepresentative datasets can exacerbate disparities. Thus, researchers and ethics committees must question dataset provenance and composition, assess differential impacts on underrepresented populations, and consider access barriers such as digital literacy and technology availability.

Finally, informed consent is an evolving process rather than a static document. Study participants must be meaningfully informed about how AI systems operate within the research they are involved in, how their data may be reused, and what risks arise from data linkage and algorithmic inference. Researchers and ethics committees must approach the informed consent process from the perspective of what a reasonable person would want to know about research on AI.

Conclusion

As AI becomes increasingly integral to the clinical research life cycle, from discovery to translation and deployment, the *informed* development and use of AI seem to be the most present ethical challenges for researchers, sponsors, institutions, funders, study participants, ethics committees, and regulatory authorities. It is an ethical imperative to ensure that those involved in research are properly trained so that they can develop AI systems and use AI tools in a responsible manner. This is to protect study participants, maintain public trust and ensure that innovation proceeds responsibly. All those involved in research are called upon to fulfil their responsibilities in full compliance with fundamental ethical principles and with the established principles of scientific integrity.

Acknowledgment: I would like to thank Dr Isabelle Wienand for her critical review of the manuscript and valuable comments.

Resources and further reading

- swissethics (2025) Artificial intelligence (AI) and research involving human beings: Issues to consider when submitting a project to a research ethics committee (version 1.0 dated 14 April 2025). Accessed 28 April 2026: https://swissethics.ch/assets/pos_papiere_leitfaden/web_swiss-ethics-position-statement_issues-to-consider-in-research-involving-ai_v1.0.pdf
- Swiss National Science Foundation (SNSF) (n.d.) Towards a responsible use of AI in research [web page]. Accessed 28 April 2026: <https://www.snf.ch/en/LE2hc62fQoNDMoFb/topic/towards-a-responsible-use-of-ai-in-research>
- United Nations Educational, Scientific and Cultural Organization (UNESCO) (2022) Recommendation on the ethics of artificial intelligence. Accessed 28 April 2026: <https://unesdoc.unesco.org/ark:/48223/pf0000381137> [Recommendation was adopted 23 November 2021]
- UNESCO (2023) Guidance for generative AI in education and research. Accessed 28 April 2026: <https://unesdoc.unesco.org/ark:/48223/pf0000386693>
- Organisation for Economic Co-operation and Development (OECD) (n.d.) AI principles [web page]. Accessed 28 April 2026: <https://www.oecd.org/en/topics/ai-principles>
- OECD (2019, amended 2024) Recommendation of the Council on Artificial Intelligence. Accessed 28 April 2026: <https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449>
- European Commission (2019) Ethics guidelines for trustworthy AI. Accessed 28 April 2026: https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=60419
- European Commission (2025) Living guidelines on the responsible use of generative AI in research (third version dated May 2026). Accessed 28 April 2026: https://research-and-innovation.ec.europa.eu/document/download/2b6cf7e5-36ac-41cb-aab5-0d32050143dc_en?file-name=ec_rtd_ai-guidelines.pdf
- European Commission (n.d.) The general-purpose AI code of practice [web page]. Accessed 28 April 2026: <https://digital-strategy.ec.europa.eu/en/policies/contents-code-gpai> [GPAI code was published 2024–2025]
- Council of Europe (2024) Council of Europe framework convention on artificial intelligence and human rights, democracy and the rule of law. Accessed 28 April 2026: <https://rm.coe.int/1680afae3c>
- Centre Responsible Digitality (ZEVEDI) (2023) Research ethics for AI research projects: Guidelines to support the work of ethics committees at universities (version 1.0 dated February 2023). Accessed 28 April 2026: https://zevedi.de/wp-content/uploads/2023/02/ZEVEDI_AI-Research-Ethics_web_2023.pdf
- Multi-Regional Clinical Trials (MRCT) Center (2025) Framework for review of clinical research involving artificial intelligence. Accessed 28 April 2026: <https://mrctcenter.org/download/24193/?tmstv=1752514597>
- World Health Organization (WHO) (2024) Artificial intelligence for health (2024). Accessed 28 April 2026: https://cdn.who.int/media/docs/default-source/digital-health-documents/who_brochure_ai_web.pdf
- Swiss Academies of Arts and Sciences (2021) Code of conduct for scientific integrity. Accessed 28 April 2026: <https://www.snf.ch/media/en/LlogiKBPPwpCrszc/Kodex-wissenschaftliche-Integritaet-en.pdf>