Published in final edited form as:
Sci Transl Med. 2023 September 13; 15(713): eadi0336. doi:10.1126/scitranslmed.adi0336.

No Time to Lose: Stronger Artificial Intelligence Regulation is Needed in the US, Especially for Biomedicine

Artem A. Trotsyuk[1,†], **Carole A. Federico**[1,†], **Mildred K. Cho**[1], **Russ B. Altman**[2,3], **David Magnus**[1,*]

[1]Stanford Center for Biomedical Ethics, Stanford University, Stanford, USA

[2]Department of Genetics, Stanford University, Stanford, USA

[3]Departments of Bioengineering, Stanford University, Stanford, USA

One-Sentence Abstract:

Regulatory agencies have ample room to make a significant impact on assuring the safety and equity of AI innovation, and the time to do so is now.

Artificial Intelligence (AI) plays a growing role across a range of sectors – including health care, manufacturing, and food production – and every week that passes brings a new, alleged achievement or dire warning. But with every advancement in AI technology, the consequences for equity, privacy and national security remain unresolved. The recently released Blueprint for an AI Bill of Rights (BORs) by the White House Office of Science and Technology Policy (OSTP) identifies five principles and practices to guide the design, use, and deployment of AI. These include:

1. Protection from safe or ineffective systems, such as avoiding inappropriate, low-quality data; risk identification and mitigation; independent evaluation; ongoing monitoring, etc.

2. Protection from discrimination by algorithms and systems used and designed in an equitable way. This includes proactive assessment of equity in design; disparity assessment and mitigation; guarding against proxies, etc.

3. Protection from abusive data practices and agency over data usage, such as data collection and use-case scope limits, privacy-preserving security; use-specific consent, etc.

[*] Corresponding author. dmagnus@stanford.edu.
[†] These authors contributed equally to this work.
Author contributions:
Conceptualization: AAT, CAF, MKC, RBA, DM
Writing – original draft: AAT, CAF
Writing – review & editing: AAT, CAF, MKC, RBA, DM

Competing interests: RBA consults with companies creating diagnostics and therapeutics, including GSK USA, Personalis, BridgeBio, Tier1 Bio, Benevolent AI, InsightRx, Myome and Withhealth. DM is the Vice Chair of the IRB for the All of Us Research program.

4. Notification and explanation that an automated system is being used which encompasses communications tailored to the purpose and level of risk, etc.
5. Ability to opt out and access to persons for remediation, outlined with brief, clear opt-out instructions, timely and not burdensome human alternative, etc.

We commend the efforts of the OSTP and the National AI Research Resource (NAIRR) Task Force, which issued its own report calling for the expansion of AI resources and democratization of innovation to ensure long-term U.S. competitiveness in this critical technology area, but suggest they lack sufficient strength and enforceability; we desperately need regulatory mechanisms that shape, monitor, control and modify activities. This is especially true in biomedicine, where AI has implications for everyone and everything (1, 2). Whereas clinical research and biomedical product development have traditions of taking responsibility for evaluating and protecting the public from harm, AI lacks such scaffolding. Thus, robust oversight is necessary for principles to be put into practice, and the time to do so is now.

The AI BORs underscore that AI tools can do good and harm. For example, in drug discovery, AI tools can predict toxicity of novel chemicals to identify those that are safe. Such algorithms, however, can be used instead to prioritize toxicity, facilitating development of chemical weapons (2). In genetics, AI can estimate inherited susceptibility for traits (i.e., polygenic scores) to enhance disease risk prediction and deploy precision therapeutics (3). However, given that these algorithms are typically built on data from people of European ancestry, they can exacerbate health inequities and lead to employment and health insurance discrimination (4). If unaddressed, unintended harms from dual use applications of AI pose significant risks. Now is the time to create regulatory mechanisms, coordinated across multiple domains and agencies, for research and clinical practice involving AI that protect the principles espoused by the OSTP (Table 1).

FDA and IRB Mechanisms are Not Designed to Regulate AI

FDA

The FDA is responsible for ensuring the safety and efficacy of medical devices, as advocated by the AI BORs’ first principle. FDA-regulated medical devices using AI, such as disease-diagnosing software, are categorized as “Software as a Medical Device” (SaMD). The FDA’s oversight of AI is nascent, however, and is challenged by rapid AI advances in biomedicine. Unlike drugs and devices – but perhaps similar to cell-based therapies – AI technologies can evolve after they are implemented, making regulation difficult. Currently, standards for updating AI code post-FDA-authorization are based on hardware medical devices policies, leaving it to authorization holders, i.e., the developers themselves, to determine if changes related to patient care meet “significant change, effectiveness, or algorithmic risk” criteria, allowing developers wide discretion. While the FDA categorizes SaMDs differently from other medical devices and sought feedback on regulatory frameworks, no further implementation has been proposed. Furthermore, AI is increasingly used to assist medical decision-making, yet is excluded from regulatory scrutiny, leaving safety and efficacy unexamined.

IRB

Institutional Review Boards (IRBs) play a role in addressing the AI BORs' concerns about safety, efficacy, and privacy. While IRBs evaluate AI research risks and benefits, they are not authorized to consider broader societal implications, such as risks posed by dual use (5). For instance, an IRB cannot reject a research protocol based on its potential weaponization. IRBs also may not have the resources or authority for post-approval monitoring and oversight, especially in rapidly evolving fields like AI. In addition, IRB review is required only for activities that are federally funded or FDA regulated, but much of AI development for biomedicine is occurring in the private sector where products do not require FDA approval (6). Finally, IRBs' closed-door review may contribute to AI's reputation of lacking transparency and traceability.

Thus, the current FDA and IRB review processes are necessary but insufficient to fulfill the AI BORs' goal of protecting American citizens from AI technology risks, and dual-use concerns are unregulated altogether. Effective regulation of AI in biomedicine will require collaboration between groups like academic institutions and the FDA, as well as government and state-level regulators. What follows describes the functions and capabilities necessary to regulate AI in research and clinical care and suggestions for meeting them, some of which can be rapidly implemented.

A Policy Framework Must Address the Use of AI in Research and Clinical Care

Despite the authority afforded to the FDA and IRBs to regulate certain aspects of AI for biomedicine, additional regulatory levers are necessary to ensure that the principles and practices described by the AI BORs are both protected and enforceable. Fortunately, there already exist precedents from synthetic biology and recombinant DNA technologies for creating a regulatory framework for AI that protects individuals and mitigates downstream risks, such as those posed by dual use (7).

Research

First, we recommend that the US Department of Health and Human Services (HHS) quickly convene a panel of experts in AI, health care, and ethics to provide guidance and synthesize recommendations for the oversight of AI in biomedicine. Such a group was convened during the pandemic to develop new standards for assessing risks associated with gain-of-function research. Additionally, the HHS should update the mandate of the National Science Advisory Board for Biosecurity (NSABB), which oversees research that poses a biologic threat to public health and/or national security, to explicitly include in their mandate the consideration of biosecurity risks posed by AI.

Second, as oversight by the FDA shapes the research agenda, clarity about the approval process for regulated AI products and the precise circumstances under which reapplication is necessary after algorithmic change is needed, as is a call for increased transparency. For example, the quantity and quality of data used for training and validation – such as input/output variables, data collection process used, patient population targeted and number of data points per patient, and the amount of missing data – should be disclosed. Moreover, AI programs, often used as classifiers, need thorough performance explanation for patients as part of the informed consent process; a tool like the confusion matrix, which assesses classifier performance along multiple dimensions, could be used to inform trial design, and its assessment shared with the FDA and patients. For non-classifier algorithms, analogous assessment strategies should be developed. Moreover, these transparency requirements should also extend to AI algorithms employed in clinical practice. The FDA could also require accreditation or pre-certification of AI developers in biomedicine, which emphasizes process evaluation for a "culture of quality and organizational excellence" versus product evaluation (6); even in an environment of increased regulation, developers have responsibilities to ensure their products are ethical. The FDA should also consult with the public to inform the design, implementation, deployment, acquisition, and maintenance of AI technologies, as described by the AI BORs. Overall, the FDA's Good Machine Learning Practice for Medical Device Development: Guiding Principles is a good first step, but too general, and oversight should cover the entire development pipeline.

Third, we recommend creating a special review mechanism as part of federal grant reviews, like the Embryonic Stem Cell Research Oversight Committee, to ensure that AI research for biomedicine has appropriate, expertise-specific oversight. The Ethics and Society Review at Stanford University could serve as a model for institutions to rapidly implement, as it successfully engaged researchers seeking internal AI funding in ethical and societal reflection early on in their projects (5). Although scaling may be difficult, it could serve as a temporary solution until other regulatory mechanisms are established.

Finally, federal research agencies should consider what types of research to support. The EU Artificial Intelligence Act, for example, bans research posing unacceptable risk, such as social scoring algorithms used by governments. We recommend that the NIH/NSF, and other federal agencies, follow suit, considering the EU's risk stratification and the NSABB's recommendations (or its AI equivalent). The private sector should implement other mechanisms, such as independent analytical validation of SaMDs or patent clocks tied to SaMD approval or independent validation, versus application. For papers using AI technologies that pose more than minimal risk, the International Committee of Journal Medical Editors could require proof of specialized review, which would provide a strong incentive for authors to meet standards (8).

Clinical Practice

Implementation of AI technologies into clinical practice, which is happening at an increasing rate, raises ethical and safety concerns. The AI BORs states, "…[y]ou should not face discrimination by algorithms and systems should be used and designed in an equitable way." However, in practice we see bias in AI products, due to biased training data, which increases disparities in health care access, quality, and outcomes (1). Despite the recognition about patient harm, regulators in some states, such as California, are still in information gathering mode and there remains no nationwide requirements for AI oversight in clinical care. Moreover, at the implementation phase in hospitals, there are no standards for when or whether AI tools are put into use, or under what conditions they should be employed (e.g., to improve efficiency vs. accuracy). Nor is there guidance about who should be involved in such decision-making, which may be fraught with conflicts of interest (9, 10).

First, as with research, the FDA should generate an explicit and transparent regulatory framework for the clinical evaluation of SaMDs that ensures their equity, safety, efficacy, and performance. In contrast to drugs and devices, this evaluation should require independent testing for approval purposes and mandate post-approval oversight, including the evaluation of real-world performance data – the disclosure and implications of which should be made during discussions with patients – and the impact of continuous learning on safety, efficacy, and performance metrics.

Second, the Centers for Medicare & Medicaid Services (CMS) should establish reimbursement policies for AI-based medical services to ensure that they are safe, effective, and meet the standard of care, mitigating risks associated with the use of AI-based products and services for patient care. Reimbursement policies provide a financial incentive for AI developers to prioritize safety and efficacy and for clinicians to employ only those technologies that have been vetted; they also ensure that the benefits of AI technologies are shared equitably by guaranteeing coverage for those that could not otherwise afford access. Finally, CMS should promote the use of AI in health care through the development of alternative payment models, reimbursing health care providers for the quality of care they provide, rather than the volume of services they deliver, incentivizing the use of AI technologies that improve patient outcomes and reduce costs. A similar approach can be used with value-based care, a form of reimbursement that ties payment of care delivery to the value of care provided, rewarding providers for efficiency and effectiveness. By doing so, CMS could spur innovation in the field of AI, while mitigating risks.

Third, state-level regulations should establish best practices for hospitals and health care providers. The Joint Commission, responsible for the accreditation of hospitals, should require institutions to have policies and procedures in place for AI in health care, including guidelines for data governance, privacy, security, and bias evaluation. Current reimbursement mechanisms are tied to quality metrics, and hospitals that fail to perform can be subject to loss of accreditation or funding, fines, and lower reimbursement rates. States and state medical boards should also enact laws that specify what hospitals and health care providers must do to use AI in the provision of care, such as requirements for training, assessment of competency and ongoing monitoring of AI products and services. These measures would ensure that physicians are equipped with the knowledge, skills, and tools to appropriately employ AI-based products and services in an effective manner aligned with patient preferences. State-level legislation should also impose penalties, such as revocation or suspension of licenses for physicians, who fail to comply with these requirements.

And finally, once best practices for the implementation of AI in medicine are established and implemented, they can become subject to tort laws. In the case where a hospital or health care provider fails to follow such practices, they can be held liable for negative consequences or harms resulting from their actions, including financial penalties and damages for individual patients who are harmed.

Conclusions

Regulating the use of AI in biomedicine, including research and clinical practice, is a pressing need that requires a comprehensive approach involving academic institutions, federal agencies, and state-level regulators. It must be informed by experts in AI, health care, ethics, and the public, and provide clearly defined standards and consequences for noncompliance. While the NAIRR Task Force presents a roadmap for cultivating responsible AI innovation in the US and makes many of the same recommendations, including ethics review mechanisms, required training and resources to support AI trustworthiness, we offer concrete means to achieve these goals. Moreover, we present multiple loci of responsibility, ranging from individual AI developers to institutional bodies, all of which have a role to play in our proposed framework. The principles and practices introduced in the AI BORs will be achievable, and concerns about potential downstream risks of AI technologies will be identified and mitigated, only if regulatory mechanisms are quickly empowered to shape, monitor, control and modify research and clinical care activities.

Funding:

GSK.ai-Stanford Ethics Fellowship (AAT, CAF)

Stanford Clinical and Translational Science Award UL1TR003142 (MKC, DM)

National Institutes of Health grant R01HG010476 (MKC, RBA)

Chan-Zuckerberg Biohub (RBA)

US Food & Drug Administration grant FDA FD005987 (RBA)

NOTE: The funders had no role in the conceptualization, decision to publish, or preparation of the manuscript. The opinions and recommendations of this paper are those of the authors, and do not necessarily represent the official views of, nor an endorsement, by FDA/HHS, or the U.S. Government.

References and Notes

1. Obermeyer Z, Powers B, Vogeli C, Mullainathan S, Dissecting racial bias in an algorithm used to manage the health of populations. Science 366, 447–453 (2019). [PubMed: 31649194]
2. Urbina F, Lentzos F, Invernizzi C, Ekins S, Dual Use of Artificial Intelligence-powered Drug Discovery. Nat Mach Intell 4, 189–191 (2022). [PubMed: 36211133]
3. Polygenic Risk Score Task Force of the International Common Disease Alliance, Responsible use of polygenic risk scores in the clinic: potential benefits, risks and gaps. Nat Med 27, 1876–1884 (2021). [PubMed: 34782789]
4. Martin AR et al. Clinical use of current polygenic risk scores may exacerbate health disparities. Nat Genet 51, 584–591 (2019). [PubMed: 30926966]
5. Bernstein MS et al. Ethics and society review: Ethics reflection as a precondition to research funding. Proc Natl Acad Sci U S A 118, (2021).
6. Nichol AA et al. A Typology of Existing Machine Learning-Based Predictive Analytic Tools Focused on Reducing Costs and Improving Quality in Health Care: Systematic Search and Content Analysis. J Med Internet Res 23, e26391 (2021). [PubMed: 34156338]
7. Kelle A, Beyond patchwork precaution in the dual-use governance of synthetic biology. Sci Eng Ethics 19, 1121–1139 (2013). [PubMed: 22535577]
8. Moher D, Implementing Incentives and Rewards to Improve the Research Ecosystem. JAMA Netw Open 4, e2138622 (2021). [PubMed: 34846530]

9. Lu J et al. Considerations in the reliability and fairness audits of predictive models for advance care planning. Front Digit Health 4, 943768 (2022). [PubMed: 36339512]

10. Jung K et al. A framework for making predictive models useful in practice. J Am Med Inform Assoc 28, 1149–1158 (2021). [PubMed: 33355350]

Table 1.

Recommendations for stronger AI regulation, organized by type of oversight, that protect the principles espoused by the OSTP and are enforceable.

Guiding examples are provided, when available.

Type of Oversight	Recommendation	Examples
Policy Synthesis and Coordination	Convene expert panel to provide recommendations and for the oversight of AI and synthesize existing guidance; update existing National Science Advisory Board for Biosecurity (NSABB) mandate to include biosecurity risks posed by AI	NSABB panel on gain-of-function research
	Create coordinated framework for regulating AI across diverse agencies	Coordinated Framework for Regulation of Biotechnology
Developer Review	Mandate certification of organizational excellence	FDA’s Software Precertification (Pre-Cert) Pilot Program
Funding	Set acceptable risk levels for federally fundable research	Risk stratification in the Artificial Intelligence Act (EU)
	Tie release of institutional funding to specialized review mechanisms	Stanford’s Ethics and Society Review
Research Review	Create specialized review mechanism for federally funded research	Embryonic Stem Cell Research Oversight Committee
	Require independent analytical validation of regulated AI products	
Product Review	Provide clear guidance about approval processes for regulated AI products, including post-approval oversight	The FDA’s Good Machine Learning Practice for Medical Device Development: Guiding Principles
	Require independent clinical evaluation of regulated AI products	
Clinical Implementation	Tie patent clocks to regulated AI products’ approval	
	Mandate post-approval oversight, including evaluation of real-world performance and impact of continuous learning	
	Establish reimbursement policies for AI-based medical services that prioritize safety, efficacy, and equity	Alternative payment models, value-based care
	Establish best practices for hospitals and providers, including accreditation, training requirements, assessment of competency and ongoing monitoring, that are subject to penalties	The Joint Commission
Dissemination	Mandate proof of specialized review for publication	International Committee of Journal Medical Editors requirement for statement about IRB approval