Available online www.ijpras.com

International Journal of Pharmaceutical Research & Allied Sciences, 2026, 15(2):1-9
https://doi.org/10.51847/uGYERdBsld

Original Article

ISSN : 2277-3657
CODEN(USA) : IJPRPM

Autonomous Translational Pharmacology Requires Human-Defined Experiment Rights, Evidence Thresholds, Escalation Rules, and Irreversible Stop Conditions for High-Consequence Decisions

Lucas Meyer[1]*, Anna Schmid[2], Stefan Braun[1], Laura Keller[3]

[1]*Department of Autonomous Translational Pharmacology and Human Oversight, Faculty of Pharmacy, ETH Zurich, Zurich, Switzerland.*
[2]*Department of Experiment Rights and Evidence Thresholds, Faculty of Pharmaceutical Sciences, University of Bern, Bern, Switzerland.*
[3]*Department of Escalation Rules and Stop Conditions, Faculty of Pharmacy, EPFL Lausanne, Lausanne, Switzerland.*

***Email:** lucas.meyer@ethz.ch*

ABSTRACT

Autonomous experimental systems are progressing from isolated automation toward integrated platforms capable of proposing hypotheses, selecting tools, executing computational or physical procedures, interpreting observations, and modifying subsequent actions. In translational pharmacology, these capabilities create governance problems that cannot be resolved by model accuracy, conventional laboratory safety controls, general human oversight, or reporting requirements alone. This Original Autonomous-Governance Theory Article proposes a Governance Constitution for Autonomous Translational Pharmacology in which experimental capability remains subordinate to human-defined scientific purposes, enumerated experiment rights, consequence-sensitive evidence thresholds, escalation rules, meaningful intervention authority, and non-overridable stop conditions. The proposed Experiment-Rights Ledger specifies the data, materials, instruments, operations, scale, duration, reversibility, and evidence state within which an autonomous system may act. Evidence gates distinguish hypothesis generation and sandboxed computation from restricted physical experimentation, adaptive closed-loop activity, and translationally consequential action. Escalation is required when novelty, uncertainty, distribution shift, conflicting observations, protocol deviation, tool substitution, cumulative risk, or loss of provenance exceeds the authorized envelope. Irreversible stop conditions prohibit actions involving unauthorized biological targets, harmful optimization, containment failure, unapproved scale expansion, non-reconstructable activity, or patient-affecting decisions without explicit authorization. Append-only logging, independent replay, dissent recording, and challenge procedures make autonomous activity contestable rather than merely observable. Governance stress tests are proposed to examine adversarial manipulation, false evidence maturity, stop circumvention, automation bias, and interactions among models, tools, instruments, and human organizations. The framework is an original conceptual synthesis that requires prospective technical, experimental, human-factors, translational, and institutional validation. It does not establish clinical utility, regulatory acceptability, universal thresholds, legal responsibility, or deployment readiness.

Key words: *Translational pharmaceutical research, Translation readiness, Reproducibility, Manufacturing feasibility, Clinical relevance, Evidence maturity*

Received: 15 January 2026; **Revised:** 30 March 2026; **Accepted:** 03 April 2026

INTRODUCTION

Reliable translational progression requires more than a promising molecular mechanism or favorable experimental result. It depends on reproducible scientific practices, explicit assessment of the evidence needed to advance a therapeutic hypothesis, and multidimensional decisions that integrate efficacy, exposure, safety, patient selection, and development feasibility [1-3]. These requirements are already difficult to maintain across conventional pharmaceutical research, where evidence is distributed among models, assays, laboratories, disciplines, and organizational decision processes.

Autonomous experimentation changes this problem because a computational system may no longer function only as an analytical instrument. It may participate in goal decomposition, experimental selection, protocol generation, tool invocation, physical execution, interpretation, and iterative modification of its own research trajectory. A system that can generate consequential actions therefore raises a different question from whether its predictions are statistically accurate: under what authority may it act, within which scientific boundaries, and subject to what evidence and interruption conditions?

Prevailing governance approaches address only fragments of this question. Laboratory safety procedures regulate known physical hazards; model-reporting standards support appraisal; reproducibility practices enable scrutiny; human-in-the-loop arrangements preserve nominal supervision; and translational frameworks organize development evidence. None alone defines a coherent relationship among autonomous capability, permission to experiment, evidence maturity, human authority, escalation, and mandatory termination. Treating any single component as sufficient risks confusing technical competence with legitimate experimental authority.

This article develops an original, non-empirical governance theory for autonomous translational pharmacology. Its principal contribution is a proposed constitution comprising a human sovereign layer, an Experiment-Rights Ledger, evidence-dependent action gates, escalation rules, intervention rights, irreversible stop conditions, provenance requirements, and scientific challenge procedures. The framework is intended to organize testable governance propositions and validation requirements. It does not claim empirical validation, exhaustive coverage, clinical applicability, regulatory endorsement, or universal suitability across research institutions and jurisdictions.

Why autonomous experimentation changes translational governance

Autonomous laboratory systems have demonstrated that experimental selection, physical execution, measurement, and iterative updating can be integrated within closed research workflows [4-6]. Mobile robotic platforms can navigate laboratory environments and choose subsequent experiments; tool-enabled language-model agents can translate scientific objectives into chemical procedures; and closed-loop laboratories can combine planning, synthesis, characterization, and learning. These demonstrations remain bounded by their tasks and domains, but they establish that autonomous experimentation is not equivalent to passive prediction.

The governance consequence arises from the transition between producing information and generating action. A conventional computational model may rank compounds or estimate an endpoint while a human decides whether and how to proceed. An autonomous platform may instead select the next procedure, alter experimental conditions, invoke a different analytical tool, or continue an iterative search without a new human decision at every step. Governance must therefore attach not only to the quality of outputs but also to the authority to transform those outputs into computational or physical consequences.

Autonomous translational pharmacology is defined here as the delegated use of computational agents, laboratory robotics, adaptive experimental design, and integrated data systems to conduct pharmacologically relevant inquiry with varying degrees of independent action. “Autonomous” does not imply consciousness, moral agency, or complete independence from human-built infrastructure. “Translational” does not mean that an experiment is clinically relevant merely because it concerns a therapeutic target. Translation readiness remains conditional on reproducibility, biological relevance, exposure, safety, patient context, manufacturing feasibility, and implementation evidence.

The central governance problem is consequently one of bounded delegation. Autonomy may increase speed, consistency, search breadth, or experimental continuity, but those properties do not establish that the system should be permitted to expand its action space. Alternative explanations for apparent autonomy must also be considered: many platforms remain highly scripted, depend on narrow instrument interfaces, or transfer crucial judgment to upstream human choices. The proposed constitution therefore regulates the actions a system is authorized to perform rather than assigning governance status from an autonomy label alone.

Scientific roles, authority, and accountability

Self-driving laboratories comprise interacting modules for goal specification, experimental planning, execution, measurement, data analysis, learning, and feedback, with substantial variation in how decision functions are distributed between humans and machines [7]. This architectural variability makes role definition essential. Technical capability describes what a platform can perform; scientific authority describes which actions it is permitted to initiate; accountability identifies the people and institutions answerable for defining, supervising, reviewing, and terminating those actions.

The proposed governance model separates operational participation from independent challenge. System developers, platform operators, principal investigators, safety specialists, translational reviewers, data stewards, and institutional authorities may contribute different expertise, but the same group should not exclusively define permissions, assess compliance, and adjudicate failures. Independent auditing provides a mechanism for examining whether a system's documented controls, behavior, and evidence correspond to declared requirements [8]. Audit remains evidentiary rather than exculpatory: passing an audit cannot establish the absence of untested failure modes.

Human accountability must likewise extend beyond placing a person nominally "in the loop." Responsible use of machine-supported decisions requires attention to data quality, intended use, evaluation, deployment conditions, monitoring, and consequences rather than technical performance alone [9]. In autonomous pharmacology, accountable humans must therefore define legitimate purposes, approve experiment classes, specify evidence gates, review escalations, authorize consequential transitions, and investigate incidents. The autonomous system is treated as a delegated operational actor, not as the scientific or institutional principal.

Authority should be distributed but not ambiguous. Scientific investigators may define hypotheses and interpret biological significance; platform specialists may verify system capability; safety authorities may prohibit hazardous operations; translational reviewers may evaluate patient relevance and evidence maturity; and institutional leadership may accept residual organizational risk. The article does not determine jurisdiction-specific legal liability. It proposes that each consequential right must have a named granting authority, an identifiable reviewer, a revocation mechanism, and an accountable institutional owner.

Proposed experiment-rights framework

Assessment frameworks can characterize dimensions of autonomous laboratory capability, but an assessment of what a platform can do does not establish what it should be permitted to do [10]. The proposed framework therefore places authorization between capability and execution. **Figure 1** presents the original conceptual synthesis for governance constitution for autonomous translational pharmacology, showing how the article's principal components, evidence relationships, uncertainties, and decision boundaries are connected.

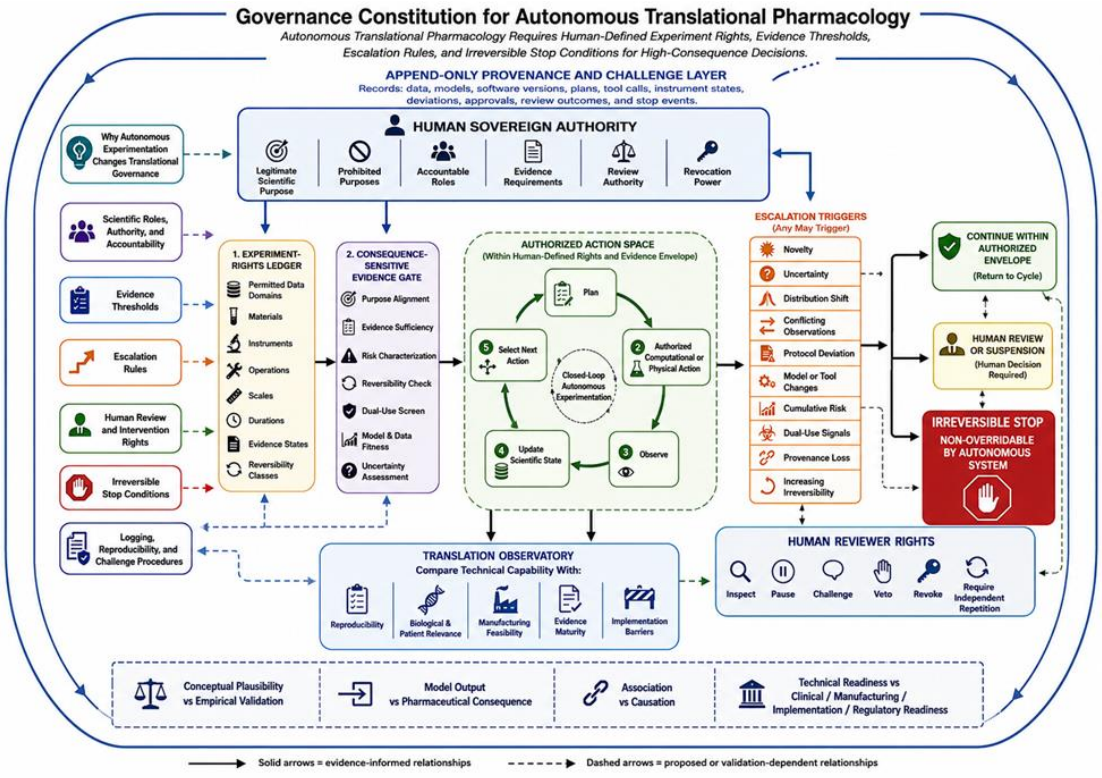

Figure 1. Governance Constitution for Autonomous Translational Pharmacology

The principal unit of authorization is the experiment right: a conditional permission to perform a defined computational or physical action. Robotic synthesis systems demonstrate that machine-selected operations can be constrained to explicit experimental choices and updated from observed outcomes [11]. The proposed framework extends this bounded-action principle by requiring each right to identify the allowed operation, material or data domain, instrument, maximum scale, duration, endpoint, evidence state, reversibility class, and responsible human authority.

Machine-executable experimental languages further demonstrate that laboratory operations can be represented as formal instructions rather than informal prose alone [12]. Formalization creates an opportunity to attach permissions, constraints, verification checks, and stop conditions to individual operations. It does not guarantee that a protocol is biologically meaningful, safe, reproducible, or translationally justified. A perfectly executable procedure may still rest on an invalid hypothesis, inappropriate model, weak evidence base, or unacceptable objective.

The proposed Experiment-Rights Ledger is a machine-readable and human-auditable record of granted, withheld, suspended, expired, and revoked permissions. Rights are non-transitive: permission to analyze a dataset does not imply permission to generate a physical intervention, and permission to conduct a small-scale assay does not imply permission to alter material class, organism, instrument, dose, scale, or translational endpoint. Scope expansion requires a new authorization rather than autonomous reinterpretation of an existing right. Rights also expire when their evidence basis, software environment, instrument configuration, or intended purpose materially changes.

The framework yields a rights-before-action proposition: no autonomous operation should occur unless an enumerated right covers the specific action and its context. It also establishes a capability–authority separation proposition: increasing technical competence does not automatically broaden permission. These propositions remain conceptual and require empirical testing through formal rights languages, machine-verifiable containment, attempted scope violations, human-factors evaluation, and independent challenge. The framework does not establish that all experimental activity can be completely specified in advance or that a ledger can prevent every technical, organizational, or adversarial route around its constraints.

Evidence thresholds for computational and physical action

Consequential model use requires a sufficiently complete evidence record, stage-appropriate evaluation, and transparent reporting that permits independent appraisal [13-15]. These requirements are necessary but do not independently authorize experimental action.

For computational activity, evidence should establish data provenance, intended use, model version, validation design, uncertainty behavior, domain limits, and reproducibility [13]. Prediction quality must remain distinct from biological truth and decision authority.

Physical experimentation requires additional evidence concerning instrument qualification, material identity, containment, protocol executability, reversibility, and failure recovery. Evaluation should become progressively more prospective and human-centered as consequences increase [14].

The framework therefore proposes states from conceptual reasoning and sandboxed computation to restricted physical execution, adaptive experimentation, and translationally consequential action. These are organizational states, not validated universal maturity levels or regulatory thresholds.

Escalation rules under novelty, uncertainty, and risk

Uncertainty must alter whether an autonomous system is permitted to continue. Machine-supported decisions require explicit attention to uncertainty when errors may produce scientific, biological, or downstream clinical consequences [16].

Escalation should occur when uncertainty exceeds the authorized envelope, cannot be estimated reliably, or conflicts with observed evidence. Novel materials, endpoints, tools, or experimental scales similarly require renewed authorization.

Communicating uncertainty is useful only when it changes action. Operational responses should include clarification, additional measurement, abstention, independent review, protocol suspension, or a second scientific opinion [17].

Escalation does not imply failure. It is a governance transfer through which authority returns to humans when novelty, cumulative deviation, distribution shift, or consequence exceeds the conditions supporting delegated action.

Human review and intervention rights

Human presence alone does not constitute effective oversight. Decision-makers can remain susceptible to incorrect machine advice even when they are expected to exercise independent judgment [18].

Meaningful review therefore requires access to relevant evidence, adequate competence, sufficient time, and practical authority. Reviewers must be able to inspect, pause, modify, veto, revoke, and require independent repetition.

Human–computer collaboration may improve performance under some conditions, yet erroneous assistance can also misdirect experts [19]. Intervention rights must therefore remain usable before, during, and after execution.

A reviewer's disagreement should be logged rather than silently overridden. The framework also protects scientific challenge by allowing minority opinions, unresolved uncertainty, and competing interpretations to accompany progression decisions.

Irreversible stop conditions and prohibited actions

Some actions should terminate authorization rather than merely trigger temporary review. Drug-discovery objectives can be redirected toward harmful optimization, demonstrating that beneficial scientific capability may also create dual-use risk [20].

AI-scientist risks may emerge through interactions among models, tools, experimental environments, and delegated autonomy [21]. Stop conditions must therefore address compound failures rather than isolated component performance alone.

Proposed irreversible stops include harmful or prohibited objectives, containment failure, unauthorized biological targets, unapproved scale expansion, unvalidated tool substitution, provenance loss, attempted safeguard circumvention, and patient-affecting action without authorization.

"Irreversible" means that the autonomous system cannot restore the terminated permission. Resumption requires a new human authorization, documented investigation, corrected evidence basis, and confirmation that the original stop condition has been resolved.

Table 1 organizes the evidence, constructs, relationships, uncertainties, and boundary conditions required for integrated construct, components, evidence requirements, failure modes, and decision boundaries for autonomous translational pharmacology requires human-defined experiment rights..

Table 1. Integrated construct, components, evidence requirements, failure modes, and decision boundaries for Autonomous Translational Pharmacology Requires Human-Defined Experiment Rights.

Component or scientific dimension	Problem addressed	Inputs or determinants	Proposed mechanism or relationship	Expected contribution	Evidence required	Failure or uncertainty risk	Boundary statement
Human sovereign authority	Technical capability may be confused with legitimate authority	Scientific purpose, institutional responsibility, affected stakeholders	**Proposed:** humans define purposes, exclusions, permissions, and revocation	Clarifies accountability	Named authorities, independent review, auditable decisions [8, 9]	Diffuse or ceremonial responsibility	Does not determine legal liability
Experiment-Rights Ledger	Autonomous scope may expand implicitly	Action, material, instrument, scale, duration, endpoint	**Proposed:** only enumerated rights permit execution	Makes delegation explicit	Capability assessment and machine-readable operations [10-12]	Ambiguous rights or circumvention	Does not guarantee complete specification
Evidence gates	Reporting may be mistaken for validation	Data quality, intended use, validation, uncertainty	**Proposed:** stronger evidence is required as consequence increases	Separates analysis from action authority	Minimum information and staged evaluation [13-15]	False evidence maturity	No universal cutoff is established
Escalation engine	Novelty and uncertainty may be ignored	Uncertainty, drift, conflicting observations, deviations	**Proposed:** exceeding the envelope returns authority to humans	Supports abstention and review	Uncertainty assessment and communication [16, 17]	Unrecognized uncertainty	Cannot eliminate unknown hazards
Human intervention rights	Nominal oversight may not prevent error	Reviewer competence, information, time, authority	**Proposed:** inspect, pause, veto, revoke, and repeat	Makes oversight operational	Human-response and collaboration evidence [18, 19]	Automation bias or delayed intervention	Human involvement does not guarantee correctness
Irreversible stop layer	High-consequence activity may continue after a critical breach	Dual-use signals, containment loss, unauthorized actions	**Proposed:** specified permissions terminate without agent override	Limits known catastrophic pathways	Dual-use and autonomous-agent risk evidence [20, 21]	Incomplete stop coverage	Does not establish fail-safe containment
Provenance and challenge	Actions may be unreconstructable or uncontestable	Data, models, software, plans, tools, instrument states	**Proposed:** append-only records support replay and challenge	Improves accountability and reproducibility	Transparency and FAIR workflow evidence [22, 23]	Incomplete or misleading logs	Logging does not establish validity
Governance stress testing	Normal operation may conceal failure	Adversarial inputs, leakage, bias, tool changes	**Proposed:** test attacks, deviations, and stop bypass	Evaluates specified failure scenarios	Adversarial and methodological evidence [24-26]	Tests may omit unknown failures	Passing tests does not prove general safety

Translation observatory	Technical success may be confused with readiness	Reproducibility, patient relevance, manufacturing feasibility, implementation	**Proposed:** monitor readiness dimensions separately	Prevents premature progression	Translational and implementation evidence [2, 3, 27]	Readiness dimensions may conflict	Does not confer clinical or regulatory readiness

Logging, reproducibility, and challenge procedures

Autonomous activity must be reconstructable. Missing methods, code, transformations, and model details prevent independent scrutiny and can undermine reproducibility [22].

The required record should include objectives, data lineage, model and software versions, plans, prompts, tool calls, instrument states, observations, deviations, approvals, escalations, and stop events.

FAIR principles support interoperable provenance across data, models, software, and workflows [23]. Nevertheless, accessible records do not demonstrate that an experiment was valid, justified, or correctly interpreted.

Challenge procedures should permit independent replay, dissent recording, counterfactual review, replication requests, and post-incident investigation. Confidentiality and security restrictions should be documented rather than used to eliminate scrutiny.

Evaluation scenarios and governance stress tests

Evaluation should examine governance under abnormal conditions rather than only successful operation. Adversarial manipulation can produce unsafe machine behavior despite apparently strong conventional performance [24].

Stress scenarios should include corrupted inputs, conflicting sensors, malicious goal changes, out-of-domain proposals, unauthorized scale increases, instrument substitution, model updates during execution, and attempted stop bypass.

Methodological weaknesses, biased datasets, and inadequate external validation can produce misleading performance claims [25]. Governance testing must therefore assess the evidence supporting authorization, not only whether controls execute mechanically.

Passing a finite scenario suite establishes performance only for the conditions tested. Independent teams should design additional challenges, including compound failures involving human delay, software error, instrument malfunction, and organizational ambiguity.

Limitations and research priorities

The framework is conceptual and may omit relevant technical, legal, institutional, cultural, or jurisdictional constraints. It has not been tested in an operating autonomous translational-pharmacology platform.

Data leakage can create apparently strong but irreproducible scientific findings, making false evidence maturity a central validation threat [26]. Leakage controls require independent evaluation rather than developer assertion.

Clinical impact also depends on prospective evaluation, workflow integration, usability, ethics, monitoring, and implementation conditions beyond technical accuracy [27]. These requirements cannot be inferred from laboratory autonomy.

Research priorities include formal rights languages, verifiable containment, stop-reliability testing, cross-site replay, human-factors studies, independent governance audits, and validation within specific computational, experimental, manufacturing, and translational contexts.

CONCLUSION

Autonomous translational pharmacology requires governance of action, not merely assessment of prediction. The proposed constitution subordinates experimental capability to human-defined purposes, enumerated experiment rights, consequence-sensitive evidence gates, escalation rules, meaningful intervention powers, non-overridable stops, and reconstructable challenge procedures. It offers a testable conceptual architecture while remaining unvalidated, context-dependent, and insufficient by itself to establish clinical utility, regulatory acceptability, legal responsibility, universal applicability, or deployment readiness.

ACKNOWLEDGMENTS: None

CONFLICT OF INTEREST: None

FINANCIAL SUPPORT: None

ETHICS STATEMENT: None

REFERENCES

1. Munafò MR, Nosek BA, Bishop DVM, Button KS, Chambers CD, Percie du Sert N, et al. A manifesto for reproducible science. Nat Hum Behav. 2017;1:0021. doi:10.1038/s41562-016-0021
2. Emmerich CH, Gamboa LM, Hofmann MCJ, Bonin-Andresen M, Arbach O, Schendel P, et al. Improving target assessment in biomedical research: The GOT-IT recommendations. Nat Rev Drug Discov. 2021;20(1):64-81. doi:10.1038/s41573-020-0087-3
3. Morgan P, Brown DG, Lennard S, Anderton MJ, Barrett JC, Eriksson U, et al. Impact of a five-dimensional framework on R&D productivity at AstraZeneca. Nat Rev Drug Discov. 2018;17(3):167-81. doi:10.1038/nrd.2017.244
4. Burger B, Maffettone PM, Gusev VV, Aitchison CM, Bai Y, Wang X, et al. A mobile robotic chemist. Nature. 2020;583(7815):237-41. doi:10.1038/s41586-020-2442-2
5. Boiko DA, MacKnight R, Kline B, Gomes G. Autonomous chemical research with large language models. Nature. 2023;624(7992):570-8. doi:10.1038/s41586-023-06792-0
6. Szymanski NJ, Rendy B, Fei Y, Kumar R, He T, Milsted D, et al. An autonomous laboratory for the accelerated synthesis of inorganic materials. Nature. 2023;624(7990):86-91. doi:10.1038/s41586-023-06734-w
7. Tom G, Schmid SP, Baird SG, Cao Y, Darvish K, Hao H, et al. Self-driving laboratories for chemistry and materials science. Chem Rev. 2024;124(16):9633-732. doi:10.1021/acs.chemrev.4c00055
8. Falco G, Shneiderman B, Badger J, Carrier R, Dahbura A, Danks D, et al. Governing AI safety through independent audits. Nat Mach Intell. 2021;3(7):566-71. doi:10.1038/s42256-021-00370-7
9. Wiens J, Saria S, Sendak M, Ghassemi M, Liu VX, Doshi-Velez F, et al. Do no harm: A roadmap for responsible machine learning for health care. Nat Med. 2019;25(9):1337-40. doi:10.1038/s41591-019-0548-6
10. Salazar-Villacis P, Benyahia B. The ADePT framework for assessing autonomous laboratory robotics. Commun Chem. 2026;9(1):99. doi:10.1038/s42004-026-01932-9
11. Granda JM, Donina L, Dragone V, Long DL, Cronin L. Controlling an organic synthesis robot with machine learning to search for new reactivity. Nature. 2018;559(7714):377-81. doi:10.1038/s41586-018-0307-8
12. Steiner S, Wolf J, Glatzel S, Andreou A, Granda JM, Keenan G, et al. Organic synthesis in a modular robotic system driven by a chemical programming language. Science. 2019;363(6423). doi:10.1126/science.aav2211
13. Norgeot B, Quer G, Beaulieu-Jones BK, Torkamani A, Dias R, Gianfrancesco M, et al. Minimum information about clinical artificial intelligence modeling: The MI-CLAIM checklist. Nat Med. 2020;26(9):1320-4. doi:10.1038/s41591-020-1041-y
14. Vasey B, Nagendran M, Campbell B, Clifton DA, Collins GS, Denaxas S, et al. Reporting guideline for the early-stage clinical evaluation of decision support systems driven by artificial intelligence: DECIDE-AI. Nat Med. 2022;28(5):924-33. doi:10.1038/s41591-022-01772-9
15. Collins GS, Moons KGM, Dhiman P, Riley RD, Beam AL, Van Calster B, et al. TRIPOD+AI statement: Updated guidance for reporting clinical prediction models that use regression or machine learning methods. BMJ. 2024;385. doi:10.1136/bmj-2023-078378
16. Begoli E, Bhattacharya T, Kusnezov D. The need for uncertainty quantification in machine-assisted medical decision making. Nat Mach Intell. 2019;1(1):20-3. doi:10.1038/s42256-018-0004-1
17. Kompa B, Snoek J, Beam AL. Second opinion needed: Communicating uncertainty in medical machine learning. NPJ Digit Med. 2021;4(1):4. doi:10.1038/s41746-020-00367-3
18. Gaube S, Suresh H, Raue M, Merritt A, Berkowitz SJ, Lermer E, et al. Do as AI say: Susceptibility in deployment of clinical decision-aids. NPJ Digit Med. 2021;4(1):31. doi:10.1038/s41746-021-00385-9
19. Tschandl P, Rinner C, Apalla Z, Argenziano G, Codella N, Halpern A, et al. Human–computer collaboration for skin cancer recognition. Nat Med. 2020;26(8):1229-34. doi:10.1038/s41591-020-0942-0
20. Urbina F, Lentzos F, Invernizzi C, Ekins S. Dual use of artificial-intelligence-powered drug discovery. Nat Mach Intell. 2022;4(3):189-91. doi:10.1038/s42256-022-00465-9

21. Tang X, Jin Q, Zhu K, Yuan T, Zhang Y, Zhou W, et al. Risks of AI scientists: Prioritizing safeguarding over autonomy. Nat Commun. 2025;16(1):8317. doi:10.1038/s41467-025-63913-1
22. Haibe-Kains B, Adam GA, Hosny A, Khodakarami F, Massive Analysis Quality Control (MAQC) Society Board of Directors, Waldron L, et al. Transparency and reproducibility in artificial intelligence. Nature. 2020;586(7829). doi:10.1038/s41586-020-2766-y
23. Huerta EA, Blaiszik B, Brinson LC, Bouchard KE, Diaz D, Doglioni C, et al. FAIR for AI: An interdisciplinary and international community building perspective. Sci Data. 2023;10(1):487. doi:10.1038/s41597-023-02298-6
24. Finlayson SG, Bowers JD, Ito J, Zittrain JL, Beam AL, Kohane IS. Adversarial attacks on medical machine learning. Science. 2019;363(6433):1287-9. doi:10.1126/science.aaw4399
25. Roberts M, Driggs D, Thorpe M, Gilbey J, Yeung M, Ursprung S, et al. Common pitfalls and recommendations for using machine learning to detect and prognosticate for COVID-19 using chest radiographs and CT scans. Nat Mach Intell. 2021;3(3):199-217. doi:10.1038/s42256-021-00307-0
26. Kapoor S, Narayanan A. Leakage and the reproducibility crisis in machine-learning-based science. Patterns (N Y). 2023;4(9):100804. doi:10.1016/j.patter.2023.100804
27. Kelly CJ, Karthikesalingam A, Suleyman M, Corrado G, King D. Key challenges for delivering clinical impact with artificial intelligence. BMC Med. 2019;17(1):195. doi:10.1186/s12916-019-1426-2