CSPS | jpps Journal of Pharmacy and Pharmaceutical Sciences

TYPE Review
PUBLISHED 11 November 2026
DOI 10.3389/jpps.2026.17484

Check for updates

OPEN ACCESS

EDITED BY
Reza Mehvar,
Chapman University, United States

*CORRESPONDENCE
Neal M. Davies,
ndavies@ualberta.ca

RECEIVED 29 July 2026
REVISED 20 October 2026
ACCEPTED 26 October 2026
PUBLISHED 11 November 2026

CITATION
Davies NM and Loebenberg R (2026) Survival of the fittest tool: choosing between open and closed artificial intelligence in pharmaceutical science. *J. Pharm. Pharm. Sci.* 29:17484. doi: 10.3389/jpps.2026.17484

COPYRIGHT
© 2026 Davies and Loebenberg. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Survival of the fittest tool: choosing between open and closed artificial intelligence in pharmaceutical science

Neal M. Davies* and Raimar Loebenberg

Faculty of Pharmacy and Pharmaceutical Sciences, University of Alberta, Edmonton, AB, Canada

Three years ago, this journal published the editorial *Technological Darwinism*, urging pharmaceutical scientists to adapt to artificial intelligence (AI) rather than resist it. Since then, the question has shifted from *whether* to use AI to *which* AI to use, under what circumstances, and with what evidence. This review addresses that question by mapping current AI tools across the five canonical stages of drug development, from discovery to post-market pharmacovigilance. It is a narrative review, not a systematic one, and it treats openness as a spectrum rather than a binary. It compares two broad categories: open systems free, often open source or open weight, inspectable, and locally deployable and closed systems, which are proprietary, commercial, and delivered as a service. The central conclusion is pragmatic. For routine laboratory, educational, and research tasks, capable open tools now exist for almost every application, including manuscript drafting, literature review, reference management, statistical analysis, pharmacokinetic and pharmacodynamic modelling, calibration and standard-curve fitting, chemical structure drawing and SMILES handling, image and figure generation, and prediction of physicochemical, ADMET (absorption, distribution, metabolism, excretion, and toxicity), formulation, and excipient compatibility properties. Free web services, however, may require confidential structures to be uploaded to a third party, so openness is distinguished here from private deployment, and independent head to head evidence against commercial tools is uneven. In contrast, integrated end to end drug discovery and validated regulatory workflows remain dominated by proprietary platforms. During the period covered by this review, the first generative AI-designed molecule targeting a generative AI-derived protein reported positive Phase IIa results and entered Phase III in July 2026; protein structure prediction was recognized with the 2024 Nobel Prize in Chemistry, shared by Hassabis, Jumper and Baker and became partly open; an autonomous frontier language model improved a longstanding medicinal chemistry reaction through thousands of wet-laboratory experiments; and regulatory agencies issued their first guidance on AI credibility. The review argues that the most enduring competency is not mastery of any single AI platform but the ability to select appropriate tools, critically validate their outputs, and document their credibility in proportion to the decisions they inform. Accordingly, open AI systems should be the default for routine, low-risk, and educational applications, whereas proprietary platforms currently retain advantages for integrated discovery and regulated development. Credibility is defined operationally, human oversight is treated as a condition of use, and the review states where comparative evidence is lacking.

KEYWORDS
ADMET, artificial intelligence, credibility, drug development, human oversight

Introduction

The 2023 editorial in the Journal of Pharmacy and Pharmaceutical Sciences read AI through the lens of evolution: techniques that improve efficiency survive, and the pharmaceutical scientists who adapt to them prosper [1]. That framing has aged well. In the few years since, large language models (LLMs) have moved from novelty to daily instrument, protein structure prediction has hardened into infrastructure, and generative medicinal chemistry has produced its first clinical candidates. The present review is narrower and more practical than a general survey. Its aim is to name the current programs a pharmaceutical scientist actually meets, to place each on the spectrum from open to closed, and to show where it earns its keep across the development pipeline and across the working life of a laboratory and a classroom. Accordingly, the aim here is not to crown a universal winner, but to define when openness, reproducibility, support and regulatory familiarity each matter most.

The phrase itself is Herbert Spencer's, not Darwin's. Darwin adopted "survival of the fittest" only in the fifth edition of On the Origin of Species (1869), at Alfred Russel Wallace's urging, as a shorthand for natural selection [2]. The borrowing is apt, because fitness depends on context: in software as in biology, the tool that wins is the one best matched to its environment, not the one that is strongest in isolation. The phrase also carries a risk, because it is almost circular. In markets as in nature, what survives reflects the environment as much as merit, so network effects, switching costs, data advantages and regulatory familiarity can lock in a tool that is not in fact the best. The same logic works in the other direction, and it is the premise of this review: an open tool can be the better choice where transparency and reproducibility matter, even when a closed rival is more capable overall.

Darwin met the obverse of this danger on the Beagle voyage. The Falkland Islands wolf, the warrah, was the archipelago's only native land mammal, and having evolved without predators it grew so tame that sailors could approach and kill it by hand; Darwin foresaw that it would soon be classed with the dodo, and by 1876 it was, the first canid known to go extinct in historical time [3]. A lineage exquisitely fitted to a sheltered niche had lost the wariness a changing environment would demand. The lesson for pharmaceutical software is exact: a tool that dominates a protected niche, shielded by switching costs and the absence of credible rivals, can forfeit the adaptability that survival later requires, and be displaced quickly once the environment opens.

The review follows the conventional five stages of drug development, namely discovery and development, preclinical research, clinical research, regulatory review, and post-market safety surveillance, and it treats the last stage not as an ending but as the recovery of real world safety and effectiveness information. Figure 1 summarizes the AI applications and the representative open and closed tools at each stage. Comprehensive surveys of the underlying methods exist elsewhere [4–7]; the emphasis here falls on choice, access and evidence.

Scope, methods and selection criteria

This is a narrative review and not a systematic one. Literature from 2023 to October 2026, with earlier foundational papers where a method originates before that date, was identified through PubMed and the arXiv, bioRxiv and ChemRxiv preprint servers, and was supplemented by regulator and standards body websites and developer documentation for licences and release status. Search terms combined artificial intelligence, machine learning or large language model with the pipeline stages, open source, open weight and tool names. Tools were included when they were publicly available, maintained, documented in a peer reviewed paper, regulatory document or repository, and relevant to a task a pharmaceutical scientist performs, and each was classified as open source, open weight, free web service or closed using the criteria in Section *Open and closed AI: a working taxonomy*. Performance claims that come from developers or from unreviewed preprints are labelled as such. Clinical, regulatory and licensing statements are given as at October 2026 and were checked against trial registries, company disclosures, regulator publications and developer repositories, and a positive pivotal readout is distinguished from regulatory approval. The search was not recorded in a form that would allow it to be called systematic, and no such claim is made. The grades in Figure 2 and the verdicts in Table 1 are qualitative judgments by the authors and not empirical audits.

Open and closed AI: a working taxonomy

"Open" and "closed" are not a clean binary but a spectrum (Figure 3). The scheme is a working framework that builds on earlier open versus closed distinctions and is not offered as the first taxonomy of its kind. It separates two properties that are often conflated, namely whether a tool is open (its code or weights can be inspected and reused) and whether it can be deployed privately (it can be run where the data reside). At one end sit fully open-source tools whose code, model weights and, ideally, training data are public under a permissive licence; the cheminformatics toolkit RDKit [18], the deep-learning framework DeepChem [19] and the property prediction package Chemprop [20] are representative. Next are open weight models, whose parameters can be downloaded and run locally but whose licences are often restricted to non-commercial or research use, and whose training data are usually undisclosed. AlphaFold 3 illustrates the nuance: its code was released in November 2024 and its weights are available on

Abbreviations: ADMET, absorption, distribution, metabolism, excretion and toxicity; AI, artificial intelligence; BCS, Biopharmaceutics Classification System; CMC, chemistry, manufacturing and controls; GAMP, Good Automated Manufacturing Practice; ICH, International Council for Harmonisation; LLM, large language model; MIDD, model-informed drug development; ML, machine learning; NLP, natural language processing; PAT, process analytical technology; PBBM, physiologically based biopharmaceutics model; PBPK, physiologically based pharmacokinetic; PK-PD, pharmacokinetic-pharmacodynamic; PopPK-PD, population pharmacokinetic-pharmacodynamic; QSAR, quantitative structure-activity relationship; RAG, retrieval-augmented generation; SMILES, simplified molecular-input line-entry system; SOUP, software of unknown provenance; VBE, virtual bioequivalence.

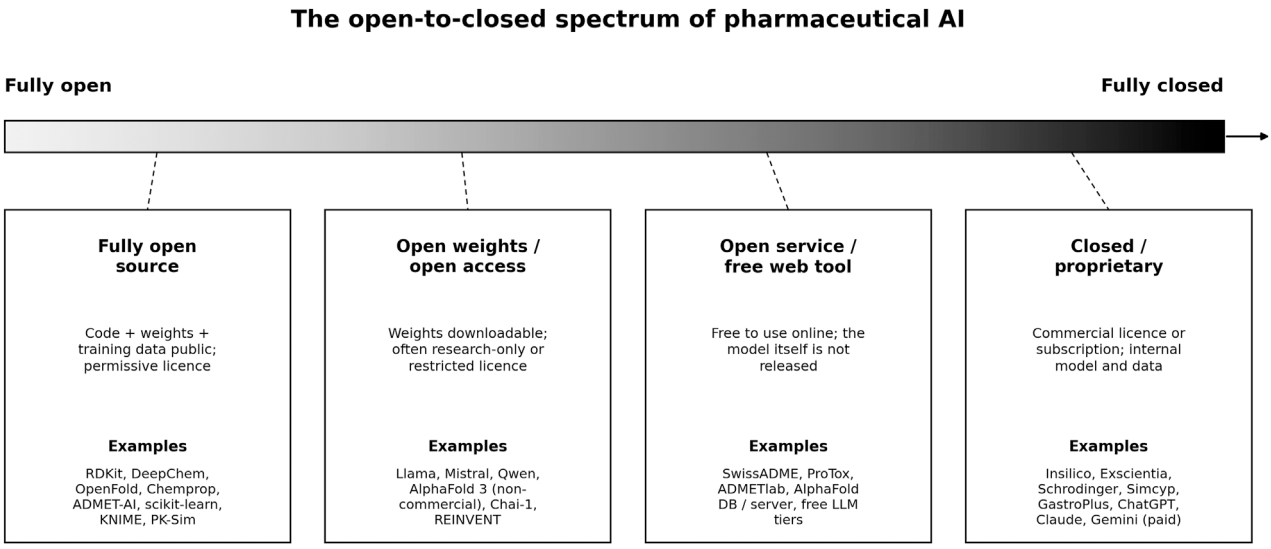


FIGURE 1
Artificial intelligence applications and representative open and closed tools across the five stages of drug development. Tool names are illustrative rather than exhaustive. Sources for the tools shown are given in Tables 4, 6 and Supplementary Table S1.

Maturity of AI use by method and pipeline stage

	Discovery	Preclinical	Clinical	Regulatory	Post-market
Protein / structure prediction	mature	established	emerging	emerging	
Generative de novo design	mature	emerging			
Virtual screening & docking	mature	established			
QSAR / ADMET property models	mature	mature	emerging	emerging	emerging
PBPK & population PK-PD modelling	emerging	mature	mature	established	emerging
Statistics & data analysis (ML)	established	mature	mature	established	mature
NLP / large language models	established	established	mature	mature	mature
Computer vision (imaging, path.)	emerging	mature	established	emerging	emerging
Pharmacovigilance signal mining		emerging	established	established	mature

Qualitative synthesis of the cited literature; labels reflect relative breadth and maturity of reported use, not validated performance.

FIGURE 2
Qualitative maturity of AI methods by pipeline stage, synthesized from the cited literature. Cell labels indicate the breadth and maturity of reported use, not validated performance. Grades reflect our qualitative judgment of the breadth of reported use, the availability of independent peer reviewed validation and the degree of regulatory acceptance [8–11], and were not derived from a numerical scoring system.

request for academic, non-commercial use under restrictive terms, which is more open than a pure service but, as the journal Nature noted, is open-weight rather than strictly open-source [21, 22]. It is therefore treated here as restricted and not as unqualifiedly open, and the public AlphaFold server, which sends submitted sequences to a third party and bars commercial use, is treated as a free web service. The open large language models commonly cited in this category, such as Llama, Mistral and Qwen, differ in their licences, which should be read model by model and rechecked at the source because they change. Llama is distributed under a custom community licence with conditions of use, many Mistral and Qwen models are released under the permissive Apache 2.0 licence while others carry research or custom terms, and DeepSeek R1 is open weight, with weights and code under an MIT licence but without the training data. Few language models are fully open in the sense of the Open Source Initiative definition, which also requires sufficient information on the training data and the training code [23]; the OLMo family from the Allen Institute for

TABLE 1 The 2023 editorial's expectations [1] measured against developments to 2026, as a qualitative assessment by the authors.

2023 expectation (technological darwinism)	What happened, 2024 to 2026	Verdict
Language models would aid literature work, data analysis, modelling and interpretation	Now routine and embedded in writing, coding and analysis, with both closed assistants and locally deployable open-weight models	Confirmed; faster than expected
AI would accelerate drug discovery and development	Generative design produced clinical candidates; an AI-derived target and molecule reached phase IIa in fibrosis [12] and entered phase III in July 2026	Confirmed; reached the clinic
Tools carry bias and limitations; use with human expertise	Hallucination and silent error remain real; human oversight now formalized in guidance [10]	Confirmed; still the central caution
Output must be checked before entering the published record	Verification is now a documented expectation in regulated use via the FDA credibility framework [8]	Confirmed; formalized by regulators
Journals would need transparent AI-use policies	Disclosure and authorship rules have been issued by editors and publishing bodies	Confirmed [13–15]
Adaptability, not resistance, would decide who benefits	The tool landscape churned (mergers, pipeline cuts [16]) yet adaptable groups persisted and advanced	Confirmed; the meta-point held
(Not anticipated in 2023) openness would become a property of the models themselves, not only of the published papers	Until recently, openness in this field meant open-access papers. Open-weight releases, such as AlphaFold 3 (17,18) and the Llama, Mistral and Qwen language models, shifted it to the models too, by making the trained parameters themselves downloadable and runnable locally	New direction; not foreseen in 2023
(Not anticipated in 2023) AI would move from advising at the desk to acting in the laboratory	Frontier models coupled to high-throughput automation began running near-autonomous experimental loops; one improved a decades-old medicinal-chemistry reaction in a screen of 10,080 microscale reactions, as reported in a developer preprint, with human chemists steering and validating [17]	New direction; not foreseen in 2023

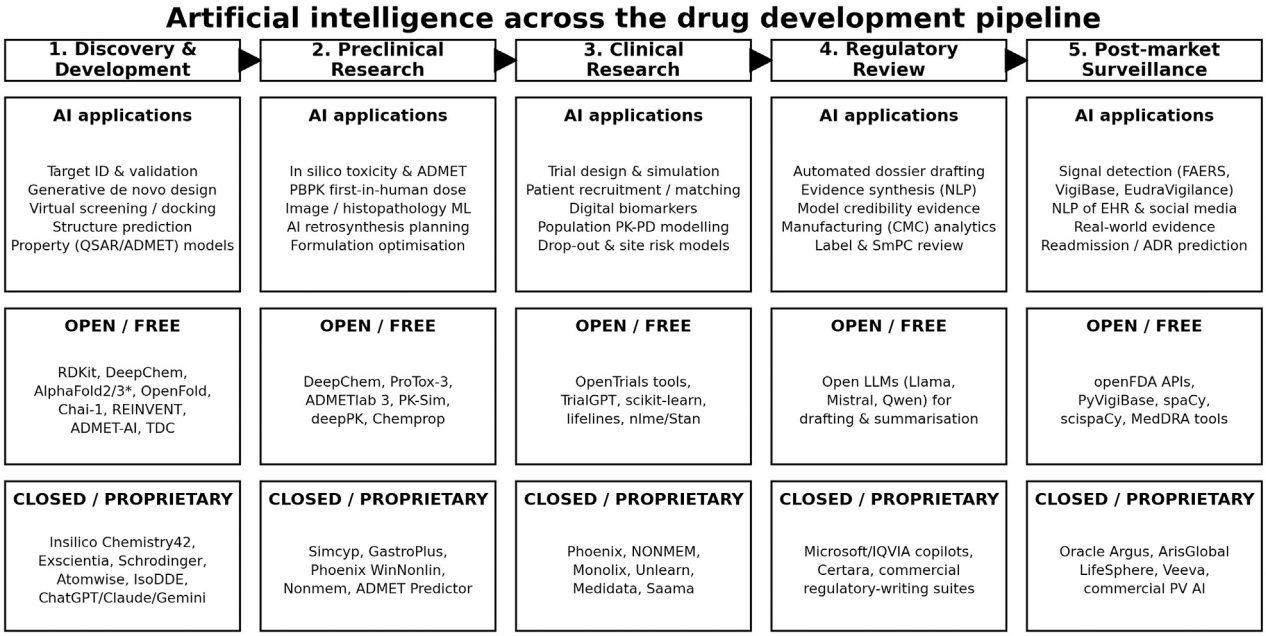


FIGURE 3
The open to closed spectrum of pharmaceutical AI, with representative tools at four reference points. Examples and their sources are given in Tables 2, 4, 5.

TABLE 2 Characteristics and trade-offs of open and closed AI systems in pharmaceutical science.

Dimension	Open/free (open-source, open-weight, free service)	Closed/proprietary (commercial, subscription)
Cost	No licence fee; cost is computer and staff time	Licence or subscription; can be substantial at scale
Transparency	Code and often weights inspectable; methods auditable	Model and data usually opaque; "black box"
Reproducibility	Exact re-execution possible; supports replication [26]	Versioning controlled by vendor; results may shift silently
Openness versus private deployment	Open code or weights can be inspected; only locally installed tools are also private; free web services are neither inspectable nor private	Not inspectable; private deployment is possible only through enterprise or on premises agreements
Data governance	Can run locally or on premises; data need not leave the institution; a free web service is not private and receives the submitted structure	Data typically routed to vendor cloud; raises confidentiality and privacy questions; zero retention or regional enterprise terms can mitigate this
Validation status	Variable; community validated but rarely regulator recognized out of the box	Often built for and familiar to regulated workflows; vendor support for qualification
Support and maintenance	Community based; self-hosting effort required	Professional support, service levels and updates
Typical examples	RDKit [18], DeepChem [19], Chemprop [20], ADMET-AI [27], OpenFold [28], PK-Sim/MoBi [29], scikit-learn; open-weight LLMs (Llama, Mistral, Qwen)	Insilico, Recursion (merged with Exscientia), Schrödinger [16], Simcyp, GastroPlus, Phoenix WinNonlin; ChatGPT, Claude, Gemini (paid)

AI, which releases weights, data and code, is the clearer example of full openness. Further along are free web services, where the tool is usable at no cost online but the model itself is not released; SwissADME [24], ProTox [25] and the public AlphaFold server belong here. These services cannot be inspected or installed locally, they can change without notice, and they may require confidential structures to be uploaded to a third party, so they carry both confidentiality and reproducibility risks. At the closed end are proprietary, commercial systems delivered by subscription or licence, including the integrated discovery platforms of Insilico Medicine, Recursion (which absorbed Exscientia in a 2024 merger), Schrödinger and Isomorphic Labs [16], the validated modelling suites used in regulated development, and the leading conversational assistants on their paid tiers. Although large language models are not drug discovery platforms *per se*, they function as proprietary AI services and therefore fit naturally within the open versus closed framework developed here. Zero data retention or regional enterprise deployments of closed models can mitigate some governance concerns, although they do not make a model inspectable, and they do not remove the need for institutional review and scrutiny of contract terms.

The distinction matters for four reasons that recur throughout this review: cost, reproducibility, data governance and validation. Open tools cost nothing to license, can be inspected and rerun, and, when run locally, keep sensitive structures or patient data on local hardware, which is decisive for confidential or regulated work (a free web service offers no such protection); their weaknesses are support, maintenance and the engineering effort of self-hosting. Closed tools offer polish, performance and, increasingly, regulatory familiarity, at the price of recurring fees, opacity and the routing of data through a vendor. Their strongest claim, though, is not polish but accountability: a validated development lineage and qualification package, vendor indemnification that shifts part of the liability, and professional security patching and support. Wherever a result must withstand audit these can outweigh the licence price, and in regulated industrial settings the value of a validated workflow and long term vendor support often does so decisively; the true cost of self-hosting and maintaining an open stack, in staff time and compute, is easily underestimated against a subscription. Table 2 sets out the trade offs.

In practice, the choice between open and closed tools is best made by asking four questions: can the code or model be inspected, can the result be reproduced later, can the data remain within institutional control, and does the use case require vendor support or formal qualification? Where the first three answers are yes, open tools are usually preferable; where the fourth dominates, closed tools may justify their cost. Table 3 renders this logic as a stepwise decision aid for routine practice. Free, open, inspectable, locally deployable and validated for a given context of use are separate attributes and none implies another: a free service may be neither inspectable nor private, an open model may be too costly to run locally, a locally deployed model is not thereby validated, and a tool validated for one context of use is not thereby fit for another. Confidential structures should not be uploaded to free web services, or to consumer tiers of closed assistants, unless the terms of use, data retention and training policies have been checked against institutional rules.

AI across the drug development pipeline

Figure 2 summarizes, as a qualitative maturity matrix, where each family of methods is most heavily used today, and the pattern is orderly. The grades are our judgment of three criteria, namely the breadth of reported use, the availability of independent validation in the peer reviewed literature and the degree of regulatory acceptance. They are not derived from a quantitative scoring system. Structure prediction, generative design and virtual screening are concentrated and mature in discovery; physiologically based and population modelling peak in pre-clinical and clinical development; statistics, natural language processing and computer vision run broad but

TABLE 3 A practical decision framework for choosing between open and closed AI tools, to be read as a sequence of questions resolved in order.

Question	If yes	If no
Is the output intended to support a regulatory submission or a publication-critical claim?	Prefer a qualified, validated platform and document its credibility for the specific context of use [8] this may be a commercial platform or an open tool pinned to a version and qualified by the user	Go to the next question
Must confidential or patient data remain under institutional control?	Prefer an open, locally deployable tool that keeps data on local hardware. Do not use free web services for proprietary structures	Go to the next question
Is exact reproducibility or auditability essential?	Prefer an open tool pinned to a fixed version, with inspectable code [26]	Go to the next question
Is vendor support, indemnification or formal qualification [30, 31] required?	A closed, supported platform may justify its subscription cost	Go to the next question
Otherwise, for educational, exploratory, routine or low-risk work	A capable free or open tool is in our assessment usually sufficient (section *Accuracy versus accessibility: where the performance gap lies*) and should be the default	End

TABLE 4 Representative AI tools by pipeline stage, classified as open or proprietary.

Stage	Open/free tools	Closed/proprietary tools	Primary AI task
Discovery	RDKit [18], DeepChem [19], Chemprop [20], ADMET-AI [27], AlphaFold 2, AlphaFold 3 (restricted, non-commercial) [21], OpenFold [28], Chai-1 (32), Protenix [32], AlphaFold DB, REINVENT, MolGPT, AutoDock Vina, SwissDock, CB-Dock2, SwissTargetPrediction, SwissADME [24], ProTox [25], PubChem, ChEMBL, DrugBank, STRING, TDC [33], COCONUT [34], LOTUS [35], Boltz-2 (37), DiffDock [36], PLIP [37], ASKCOS [38], vNN-ADMET [39], Open Targets Platform [40], PharmGKB [41]	Insilico (Chemistry42), Recursion (merged with Exscientia), Schrödinger [16], Atomwise, Isomorphic IsoDDE, Molecule.one (Maria) [17]	Target ID, generative design, docking and screening, property prediction, autonomous reaction optimization [17], genetics-led target prioritization, synthesis planning
Development and preclinical	DeepChem [19], ProTox [25], ADMETlab, pkCSM [42], BioTransformer 3.0 (45), admetSAR, DruMAP, PK-Sim/MoBi [29], deepPK, computer-vision libraries, AI-PBPK platform [43], FormulationDE [44] FormulationAI [45]	Simcyp, GastroPlus, ADMET Predictor, Phoenix	In silico ADMET and toxicity, metabolite prediction, PBPK dosing, imaging
Clinical	Scikit-learn, lifelines, nlme, Stan, Pumas (Julia), TrialGPT-style tools	NONMEM, Monolix, Phoenix WinNonlin, Unlearn, Medidata, Saama	Trial design and simulation, PopPK-PD, recruitment
Regulatory	Open-weight LLMs (Llama, Mistral, Qwen) for drafting and summarization	Certara and commercial regulatory-writing suites, enterprise copilots	Dossier drafting, evidence synthesis, CMC analytics
Post-market	openFDA APIs, spaCy/scispaCy, scikit-learn, MedDRA tooling	Oracle Argus, ArisGlobal LifeSphere, Veeva, commercial PV AI	Signal detection, NLP of reports and EHR, RWE

uneven across the whole pipeline; and pharmacovigilance methods come into their own only after approval. The subsections that follow walk through the five stages in turn, naming the open and closed tools at each and the evidence behind them. Table 4 consolidates the tools by stage.

Discovery and development

Discovery is where AI is most visibly transformative and where the open and closed worlds diverge most sharply. Four tasks dominate. First, target identification and validation increasingly use protein structure prediction: AlphaFold and its successors predict three dimensional structures and, in AlphaFold 3, protein interactions with ligands, nucleic acids and ions, a body of work recognized by the 2024 Nobel Prize in Chemistry, shared by Hassabis and Jumper for protein structure prediction and by Baker for computational protein design [21, 46]. The model is partly open [22], and independent open models such as OpenFold [28], Chai-1 [47] and Protenix [32], the last two so far described only in preprints, give the academic laboratory genuinely free options, while Isomorphic Labs commercializes a closed successor engine. A newer open-weight entrant, Boltz-2, jointly

predicts complex structure and small molecule binding affinity under a permissive MIT licence and, according to its developers, is the first deep learning model to approach the accuracy of physics based free-energy perturbation calculations on the benchmarks they report, extending the open structure prediction ecosystem from pose to potency [48]. This is a developer reported claim from a preprint that has not yet been peer reviewed or independently replicated at scale. Upstream of structure, the open Open Targets Platform integrates human genetics, functional genomics and clinical evidence to score and rank target disease associations, and PharmGKB curates variant drug associations and clinical dosing guidelines, together giving academic groups a free route to target prioritization and pharmacogenomic risk assessment before a structure is ever modelled [40, 41]. Second, generative *de novo* design uses deep generative models to propose novel chemotypes; the field reached proof of concept when a model identified potent DDR1 inhibitors in 21 days [49], and open toolkits such as REINVENT and the transformer based MolGPT let academic groups run comparable workflows without licence cost. Third, virtual screening and molecular docking are accelerated by machine learned scoring, with open stacks built on RDKit [18] and DeepChem [19] and free docking engines such as AutoDock Vina, the SwissDock server and the automated CB-Dock2 cavity detection server competing against proprietary physics plus machine learning platforms; the chemical matter for these workflows comes overwhelmingly from the open PubChem, ChEMBL and DrugBank repositories, which catalogue structures, bioactivities, targets and interactions for essentially every known compound, while free reverse screening tools such as SwissTargetPrediction and the Similarity Ensemble Approach predict the likely targets of a structure and support drug repurposing. Diffusion based docking, exemplified by the open DiffDock, reframes pose prediction as a generative rather than a search problem and needs no predefined binding pocket, while the open Protein-Ligand Interaction Profiler (PLIP) automatically extracts and visualizes the hydrogen bonds, hydrophobic contacts and π-stacking that a docked or crystallographic pose implies [36, 37]. An independent evaluation with the PoseBusters test suite found that deep learning docking methods, DiffDock among them, did not yet outperform classical engines such as AutoDock Vina and the commercial CCDC GOLD in physical plausibility or in generalization to novel complexes [50]. Downstream of a proposed structure, the open ASKCOS suite combines retrosynthetic route prediction, reaction outcome prediction and condition recommendation in a single freely available tool, closing the loop from generative design to a plausible synthetic route [38]. Fourth, early property prediction (quantitative structure-activity relationships and ADMET) is now quite robust and largely open: ADMET-AI [27], built on Chemprop and trained on the Therapeutics Data Commons [33], was reported by its developers to rank highest on the Therapeutics Data Commons ADMET leaderboard at the time of its publication, a developer reported and date dependent result, and free web tools such as SwissADME [24] and ProTox [25] cover drug likeness and toxicity. The variable nearest neighbour vNN-ADMET server adds cytotoxicity, CYP450 inhibition, mutagenicity and liver injury endpoints from a distance weighted similarity model [39]. The integrated commercial platforms (Insilico, Recursion, which merged with Exscientia in 2024, and Schrödinger) bundle these steps end to end and have produced the clinical candidates discussed in Section *Clinical research* [16].

FormulationAI is a web based AI platform that supports *in silico* pharmaceutical formulation design by predicting key formulation properties from basic drug and excipient information. This enables rapid identification of promising formulations while reducing reliance on trial and error experimentation, shortening development timelines, and lowering costs [45]. Complementing this approach, FormulationDE applies machine learning to predict drug excipient compatibility and identify potential interaction risks based on molecular functional groups. By providing rapid, interpretable predictions, it streamlines pre-formulation studies and reduces the need for labour-intensive compatibility testing [44]. As web services, these tools carry the confidentiality caution of Section *Open and closed AI: a working taxonomy*.

AI is also transforming pharmacokinetic (PK) prediction. Integrated artificial intelligence physiologically based pharmacokinetic (AI–PBPK) platforms can predict *in vivo* drug disposition directly from molecular structures by estimating key physicochemical and ADME properties, including solubility, permeability, protein binding, plasma clearance, and tissue partitioning. These parameters are incorporated into PBPK models to forecast PK profiles, with validation against more than 670 human intravenous and oral datasets demonstrating robust predictive accuracy. By accurately predicting tissue distribution and organ selectivity, AI–PBPK models can accelerate candidate selection and prioritize compounds with favourable PK characteristics, reducing dependence on extensive *in vitro* and *in vivo* studies [43]. The formulation and AI–PBPK platforms described in this section are free web services, so proprietary structures should not be uploaded to them.

Beyond computational prediction, AI is beginning to automate experimental chemistry. In 2026, OpenAI and Molecule.one reported the first publicly documented use of a frontier language model as a near-autonomous agent in a real medicinal chemistry workflow. Coupled to an agentic system and a high-throughput robotic laboratory, GPT-5.4 identified TEMPO as a mild oxidant capable of improving the yield of the challenging Chan–Lam coupling reaction for primary sulfonamides. In two microscale screening campaigns comprising 10,080 reactions, the optimized TEMPO condition raised the mean estimated product yield from 16.6% to 25.2% and increased the fraction of reactions exceeding a 30% yield from 15.6% to 37.5%, and bench scale experiments confirmed the benefit in 11 of 14 representative substrate pairs [17]. The screening yields were estimated by liquid chromatography with absorbance correction and not isolated, the mechanism remains unresolved, and the result applies to one reaction class on one platform. Although the work is a preprint authored by the developers that has not been peer reviewed, and required human oversight for experimental design and validation, it demonstrates how AI can extend beyond *in silico* prediction to assist with autonomous laboratory experimentation while maintaining scientific accountability [17].

A sixth task, less often itemized but central to phytomedicine and natural product research, is the systematic mining of curated databases of bioactive natural products for candidate scaffolds. The open option is COCONUT (COlleCtion of Open NatUral

prodUcTs), a free, login free web resource whose 2024 overhaul aggregates more than sixty natural product collections and supports name, structure, substructure and similarity searching, bulk download and a REST API, with precomputed natural product likeness, synthetic accessibility and drug likeness scores to help prioritize compounds [34]. Machine learning has been built directly on it: a support vector machine trained on COCONUT structures assigned biological origin (plant, fungal or bacterial) at 89 to 94 percent accuracy, illustrating how an open database also serves as a substrate for predictive models [51].

The related LOTUS initiative, disseminated through Wikidata, adds more than 750,000 referenced structure organism pairs and is the natural companion for ethnopharmacological and traditional medicine work [35]. The proprietary comparators are the Dictionary of Natural Products and the natural product content of subscription chemistry databases, so the open versus closed contrast of this review extends to natural product space as well. Downstream of scaffold selection, the multi-target mechanisms of natural products and repurposed drugs are increasingly dissected by network pharmacology, for which the open STRING protein interaction database, the Cytoscape network visualization platform and the MetaboAnalyst metabolomics server form a free and widely used toolchain.

Preclinical research

In pre-clinical work AI supplements rather than replaces wet lab toxicology and pharmacokinetics. In silico toxicity and ADMET models triage liabilities before synthesis [25, 27, 52]; computer vision reads histopathology and high content imaging, a use that is mature in this stage; and physiologically based pharmacokinetic (PBPK) modelling supports first in human dose projection. PBPK is the clearest example of an open option matching a commercial one: the Open Systems Pharmacology suite (PK-Sim and MoBi) is free and open source [29], standing alongside the proprietary Simcyp and GastroPlus that are standard in industry. Machine learned synthesis planning and formulation optimization round out the stage. Reviews of AI for pharmacokinetics [52] and for pharmacology practice [53] document the breadth of use.

Clinical research

Clinical development uses AI for trial design and simulation, patient recruitment and matching, digital biomarkers, population pharmacokinetic-pharmacodynamic (PopPK-PD) modelling, and prediction of drop out and site risk. The pharmacometric workhorses here are largely closed (NONMEM, Monolix, Phoenix WinNonlin), although open alternatives in R (nlme, the Stan ecosystem) and Julia (the Pumas project) are growing. The headline development of the review period is clinical: Insilico Medicine's rentosertib (ISM001-055), a TNIK inhibitor in which both the target and the molecule were generated by AI, reported positive phase IIa results in idiopathic pulmonary fibrosis and entered a phase III program in July 2026 (NCT07687459) [12, 16], while Exscientia's DSP-1181 was the first AI-designed molecule to enter human trials, it was discontinued after phase I, which is a reminder that accelerated discovery does not guarantee clinical fitness, and the physics based computational design platform of Schrödinger, applied at Nimbus Therapeutics, produced zasocitinib (TAK-279), a TYK2 inhibitor that reported positive pivotal phase III results in plaque psoriasis in December 2025 and superiority over deucravacitinib in a head to head phase III study in June 2026 [16, 54, 55]. These results validate the discovery platforms but remain early, and whether any AI-originated drug has completed phase III depends on the definition used. We use AI-originated for a drug whose target and molecule were both identified by AI, as for rentosertib, and computationally designed for a molecule optimized with physics based or machine learning tools against a conventionally selected target, as for zasocitinib. On that definition no AI-originated drug has yet completed phase III or reached the market, whereas a computationally designed drug has reported pivotal phase III results.

Regulatory review

AI is entering the regulatory stage from two directions. Sponsors use natural language processing and LLMs to draft and quality check dossiers, synthesize evidence and review labelling, and use analytics for manufacturing and for chemistry, manufacturing and controls (CMC) (Table 5). Regulators, in turn, have begun to govern AI used to support decisions. In January 2025 the US Food and Drug Administration issued its first dedicated draft guidance, *Considerations for the Use of Artificial Intelligence to Support Regulatory Decision-Making for Drug and Biological Products*, proposing a seven step, risk based credibility assessment framework anchored on the model's context of use [8]. The guidance reflects the agency's experience with more than 500 drug and biologic submissions containing AI components since 2016 and has itself been the subject of critical appraisal [56]. As far as we could determine, the FDA still listed this document as a non-binding draft guidance in September 2026 [10]. For the pharmaceutical scientist, the practical message is that an AI result intended to support a regulatory claim must carry documented evidence of its credibility for the specific question it addresses. Parallel activity is under way internationally: the European Medicines Agency adopted a reflection paper on AI across the medicinal product lifecycle in September 2024 [9]; the FDA and EMA jointly published ten guiding principles of good AI practice in drug development in January 2026, stressing human centric design, a risk based approach and a clear context of use [10]; ICH adopted its M15 guideline on model informed drug development in January 2026, with a framework of model influence, risk and impact [11]; the MHRA and Health Canada have, with the FDA, identified good machine learning practice principles, and Health Canada issued pre market guidance for machine learning enabled medical devices in April 2026, although these address devices and not medicines directly [57, 58]; and the European Union Artificial Intelligence Act (Regulation (EU) 2024/1689) adds a horizontal legal layer applicable to many AI systems used in drug development, so the United States focus of this section reflects the most developed example rather than the only one.

Post-market surveillance and the recovery of safety information

After approval, AI is used to recover safety and effectiveness signals from data at a scale beyond manual review. Spontaneous reporting databases (the FDA Adverse Event Reporting System, the

TABLE 5 The large language model landscape relevant to pharmaceutical work: open-weight versus closed.

Attribute	Open-weight (e.g. Llama, Mistral, Qwen, DeepSeek)	Closed (e.g. ChatGPT, Claude, Gemini)
Access	Weights downloadable; can self-host	Service only, via web or API
Licence	Varies by model: custom community licence (Llama); Apache 2.0 for many Mistral and Qwen models; MIT licence on weights and code for DeepSeek R1; fully open release for OLMo [23]	Proprietary terms of service
Cost	Free weights; pay for own compute	Free tier limited; paid subscription or per-token
Data privacy	Data can stay on local hardware	Data routed to vendor unless enterprise terms apply; zero retention or regional deployment options exist
Customization	Full fine-tuning possible	Limited to provider features
Capability	Strong and closing the gap; behind on the hardest reasoning	Generally leads on reasoning, coding and polish
Best fit	Confidential or regulated text; reproducible local pipelines	Highest-quality drafting and analysis when data sharing is acceptable

World Health Organization VigiBase, and the European Medicines Agency EudraVigilance) receive millions of reports annually, and machine learning and natural language processing now augment classical disproportionality analysis for signal detection [59–61]. Open building blocks include the openFDA application programming interfaces, the spaCy and scispaCy text-processing libraries, and MedDRA-aware tooling; closed safety platforms (for example Oracle Argus and ArisGlobal LifeSphere) embed AI for case processing in industry. Generative models are being explored for narrative summarization and causality assessment [62], and narrative reviews emphasize the need for explainability and human oversight in this safety critical setting [63].

Biologics and protein design

The open versus closed contrast extends to biologics, although this review concentrates on small molecules. Protein language models such as the ESM family [64], diffusion based backbone generators such as RFdiffusion [65] and inverse folding tools such as ProteinMPNN [66] are released openly, with licence terms that vary by model, whereas antibody design is served by both academic releases and proprietary platforms whose comparative evidence is mostly developer reported. Predicted structures and sequences still require experimental expression, binding and developability testing, and immunogenicity cannot be predicted reliably from sequence. Maturity and evidence differ substantially between protein design, antibody engineering and manufacturing analytics (Section *Manufacturing, chemistry, manufacturing and controls, and process analytical technology*).

Manufacturing, chemistry, manufacturing and controls, and process analytical technology

Manufacturing and CMC are where AI is least visible in the open literature and where closed, validated systems dominate, because the data are proprietary and the systems operate under good manufacturing practice. Machine learning predicts critical quality attributes from composition and process data and underpins soft sensors and multivariate models on process analytical technology (PAT) signals, including in continuous manufacturing, which ICH Q13 addresses [67]. The FDA has sought comment on integrating AI into its manufacturing framework [68], and the joint FDA and EMA principles span manufacturing as well as development [10]. A model that releases product or controls a process is a high risk context of use whose lifecycle, change control and oversight must be qualified within the sponsor quality system.

AI at the bench and in teaching

The editorial that prompted this review was addressed to the working pharmaceutical scientist [1], and so is this section. It answers the everyday question of which tools actually help with the tasks a pharmaceutical scientist or student traineee performs from 1 day to the next. The encouraging answer is that, for routine bench and classroom work, a capable free or open tool now exists for very nearly every one of them. Table 6 is organized by task rather than by vendor. Many of the tools below are free web services, and Section *Open and closed AI: a working taxonomy* explains why proprietary structures should not be uploaded to them.

Statistics and data analysis

General statistics, regression, survival analysis and machine learning are well served by free, open and reproducible environments: R (with packages such as nlme, survival and tidymodels), Python (scikit-learn, statsmodels, lifelines), and the point and click JASP and jamovi for teaching. These are open source and citable, which clearly matters for method transparency in a manuscript. Conversational assistants are useful as statistical tutors and code generators but should not be trusted to perform or report analyses without verification, a caution echoed across the AI in pharmacology literature [53, 72].

Pharmacokinetic and pharmacodynamic modelling

For non-compartmental analysis, compartmental PK-PD and population modelling, the open options include R (nlme, nlmixr2,

TABLE 6 A free starter kit at a glance: one recommended open or free starting tool for each common pharmaceutical sciences task. Tools are illustrative, and the discipline of recording the exact version used applies throughout.

Task	Recommended free starting tool
General statistics and teaching	JASP or jamovi, point-and-click and built on R
Scripting, reproducible analysis and machine learning	R (tidymodels, survival) or Python (scikit-learn, statsmodels)
Non-compartmental and population pharmacokinetics	R (nlme, nlmixr2 (70,71), mrgsolve)
Physiologically based PK and biopharmaceutics	PK-Sim and MoBi (open systems pharmacology) [29]
Allometric scaling and first-in-human dose	PKSmart [69] or DallphinAtoM [70] R, Python (power-model regression), PK-Sim and MoBi [29]
Calibration and standard curves	R (drc) or Python (lmfit, SciPy)
Physicochemical and ADMET prediction	SwissADME [24], ADMET-AI [27], ProTox [25], ADMETlab, pkCSM [42], admetSAR (web services; use a local installation for proprietary structures)
Metabolite prediction	BioTransformer 3.0 (45) (web service; use for non-proprietary structures)
Provisional BCS class	FormulationBCS [71] (web service)
Formulation	FormulationDE [44], FormulationAI [45] (web services)
Structure drawing and SMILES	RDKit [18], with Ketcher or JSME for sketching and format conversion; open Babel for interconverting file formats
Molecular docking	AutoDock Vina, or the diffusion-based DiffDock [36]
Protein structure	The public AlphaFold server [21, 22], or OpenFold [28] (the server is restricted to non-commercial use and receives the submitted sequence)
Natural-product scaffolds	COCONUT [34]
Confidential writing and coding	A locally run open-weight model (Llama, Mistral or Qwen)
Reproducible workflow assembly	Jupyter or KNIME

Web services (SwissADME, ProTox, the public ADMET-AI server, FormulationBCS, FormulationAI, FormulationDE and the AlphaFold server) receive the submitted structure or sequence, so proprietary or unpublished structures should not be submitted to them. Locally installed packages (RDKit, ADMET-AI, OpenFold, PK-Sim and MoBi) keep data on local hardware.

mrgsolve), the Julia-based Pumas project, and, for mechanistic PBPK, the open-source PK-Sim and MoBi suite [29]. The commercial standards remain Phoenix WinNonlin, NONMEM, Monolix, Simcyp and GastroPlus [72]. Machine learning is increasingly used to predict PK parameters directly from structure [52]; this is an area where a free model can give a useful first estimate that a validated tool later confirms. Free chatbots have also been evaluated as substitutes for specialized pharmacokinetic software in teaching: prompt-engineered free versions of ChatGPT and Gemini produced accurate, reproducible patient only and Bayesian estimates of volume of distribution and elimination rate constant, making Bayesian analysis more accessible to educators and students, though the results clearly depended on prompt design and required verification [73].

Calibration and standard curves

Fitting calibration and standard curves (linear, weighted linear and four-parameter logistic) needs nothing proprietary. R (the drc package), Python (SciPy and lmfit) and open spreadsheet tools fit these models, report weighting and goodness of fit, and can be scripted for reproducibility. An LLM is helpful for generating the fitting script and for explaining why a weighting scheme is appropriate, provided the numerical result is checked against the open library.

Structure drawing, SMILES and chemical notation

Drawing structures, generating and parsing SMILES, canonicalizing and converting between formats, and computing descriptors are core open-source strengths. RDKit [18] handles SMILES parsing, canonicalization, two-dimensional depiction and descriptor calculation programmatically; Open Babel converts between formats; and free editors such as the Ketcher and JSME web sketchers produce SMILES for teaching and manuscripts. LLMs can read and generate SMILES and explain notation, but they make silent errors on stereochemistry and valence, so generated structures should always be validated in RDKit before use.

Physicochemical and ADMET property prediction

Prediction of logP, solubility, pKa, permeability and the wider ADMET panel is now a free-tool strength. SwissADME [24] gives drug-likeness and pharmacokinetic summaries from a SMILES string in seconds; ADMET-AI [27] provides fast, accurate, batch

predictions with an open Python package; ProTox [25] predicts toxicity endpoints; pkCSM [42], from the University of Cambridge and the University of Melbourne, uses graph based signature machine learning models to predict ADMET properties, including CYP substrates and inhibitors, from a SMILES string or a two dimensional structure; BioTransformer 3.0 [74] combines machine learning with knowledge based rules to predict mammalian, CYP450 and phase II metabolism and to generate the structures of the predicted metabolites from a SMILES string or an SDF file; and the Therapeutics Data Commons [33] supplies the datasets and benchmarks. Commercial predictors (ADMET Predictor, the Schrödinger suite) add breadth and support. In our assessment the free tools are adequate for most teaching, ranking and early stage decisions, although independent head to head comparisons with commercial predictors are limited (Section *Accuracy versus accessibility: where the performance gap lies*). SwissADME, ProTox, pkCSM, BioTransformer 3.0 and the public ADMET-AI server are web services that receive the submitted structure, so proprietary or unpublished structures should be predicted with a locally installed package such as ADMET-AI. pkCSM is free for academic use, whereas non-academic use requires a server licence.

Scientific writing, literature work and teaching

Research has little effect until it is communicated, and much of a pharmaceutical scientist's working life is spent turning results into figures, slides, posters, abstracts, manuscripts and regulatory documents. AI now assists at every step of that dissemination: generating and refining figures, images and presentation slides, checking grammar and references, drafting and restructuring text, and helping to assemble and format formal submissions such as investigational new drug applications (Table 5). The gain is in speed and polish rather than in substance, and the responsibility for accuracy stays with the author. LLMs are now embedded in writing, summarization, coding and tutoring. The pragmatic division is that closed assistants (ChatGPT, Claude, Gemini) in our assessment currently lead on reasoning and polish, while open-weight models (Llama, Mistral, Qwen) can be run locally when confidentiality forbids sending text to a vendor. In teaching, these tools generate worked examples, quiz items, case vignettes and code, and they lower the barrier to the computational methods above. Dedicated literature search assistants such as Elicit, Consensus and the Semantic Scholar research tools apply similar language model techniques specifically to systematic search, evidence extraction and citation graph navigation, and are increasingly used alongside general purpose chatbots when scoping a review. Two cautions, consistent with the original editorial and with the pharmacology literature, apply throughout: outputs must be verified because models fabricate plausible references and results, and use must be disclosed in line with journal and institutional policy [1, 53]. A related development is the emergence of AI enabled electronic laboratory notebooks (for example Benchling AI and eLabNext AI), which summarize experiments, extract and structure protocols and build searchable knowledge graphs, extending this assistance from the manuscript to the day to day laboratory record and broadening its industrial relevance.

Biopharmaceutics: solubility, permeability and BCS classification and formulation development

Biopharmaceutics adds a specific and high value case to the property prediction of Section *Physicochemical and ADMET property prediction*. The Biopharmaceutics Classification System (BCS) sorts an oral drug into one of four classes by aqueous solubility and intestinal permeability [75], and a BCS assignment can support a biowaiver of *in vivo* bioequivalence studies under the harmonized ICH M9 guideline for class I and class III drugs [76]. The two discriminating measurements, solubility and permeability, are precisely the properties that machine learning now estimates quickly from structure. Free web tools give a usable first pass: SwissADME reports lipophilicity, aqueous solubility and a gastrointestinal absorption flag from a SMILES string [24], and ADMET-AI returns calibrated aqueous solubility and Caco-2 permeability among its endpoints [27]. A dedicated open platform, FormulationBCS, curates log S, log P, log D and apparent permeability datasets and predicts BCS class directly, showing that *in silico* provisional classification is now feasible at high throughput for early developability and formulation decisions [71]. The practical division is the same as elsewhere in this review: a free model is well suited to ranking candidates, triaging a library or teaching the BCS, whereas a regulatory biowaiver still rests on the experimental solubility and permeability determinations specified by ICH M9, for which a predicted class is a hypothesis to be confirmed rather than evidence in itself [76]. FormulationAI and FormulationDE, described in Section *Discovery and development*, extend this approach to formulation design and drug excipient compatibility [44, 45].

In silico bioequivalence and digital twins

Two model-based methods extend the previous subsections from prediction toward evidence, and AI assists in both. The first is virtual bioequivalence (VBE). Where Section *Biopharmaceutics: solubility, permeability and BCS classification and formulation development* used machine learning to assign a BCS class, VBE goes a step further and simulates the bioequivalence study itself: a physiologically based biopharmaceutics model (PBBM), built on the same physiology as PBPK, generates a virtual population, propagates formulation and physiological variability, and predicts whether a test and a reference product would fall within bioequivalence limits [77]. Regulators including the FDA, EMA, MHRA, Health Canada, ANVISA and PMDA now actively discuss VBE for formulation bridging, dissolution safe space justification and selected biowaivers [77], and the open PK-Sim and MoBi suite [29] offers a free entry point to the underlying PBPK, although validated commercial platforms such as Simcyp and GastroPlus remain standard for regulatory submissions. As with the BCS biowaiver of Section *Biopharmaceutics: solubility, permeability and BCS classification and formulation development*, revisited in Section *Software of unknown provenance (SOUP) and the governance of AI in biopharmaceutics*, a VBE result intended to support a regulatory claim is a high-stakes context of use whose model must be qualified and documented [8, 77].

The second method moves to the clinical trial. An AI digital twin is a model, trained on historical patient data, that forecasts a participant's likely trajectory under control or standard of care from their baseline characteristics alone [78]. Used as a prognostic covariate rather than as a synthetic control, the digital twin preserves randomization while increasing statistical power, which can in turn permit a smaller control arm. This approach, prognostic covariate adjustment, received a qualification opinion from the European Medicines Agency in 2022 and underlies the commercial platform already listed for clinical development in Table 4 (Unlearn) [79]. Digital twins make the same point as the rest of this review: the value lies not in replacing the experiment but in using a well governed model, with a documented context of use, to make the experiment more efficient while a human remains accountable for the inference [8].

Allometric scaling and human pharmacokinetic prediction

Predicting human pharmacokinetics and a first in human dose from preclinical data is one of the few areas of drug development where the free ecosystem is unusually deep, and it spans both mechanistic and machine learning approaches. Classical interspecies allometric scaling, the log-linear regression of clearance and volume of distribution on body weight, needs nothing proprietary and runs in R or Python. Mechanistic physiologically based extrapolation is served by the open-source Open Systems Pharmacology suite, whose PK-Sim and MoBi components carry species specific physiological databases, support cross species scaling and first in human projection, and are used across academia, industry and regulatory science [29]; a freely reusable library of qualified PBPK models shortens model development for new compounds. AI–PBPK platforms, described in Section *Discovery and development*, predict profiles directly from structure [43]. A purpose-built free tool, DallphinAtoM (Drugs with Allometry and Physiology Inside, Animal to huMan), integrates published PBPK and allometric methods in a transparent graphical interface and predicts human PK parameters and concentration-time profiles from physicochemical properties, *in vitro* ADME data and animal PK data, with the explicit aim of letting the user move between allometric scaling and PBPK as the available data allow [70]. The machine learning option is PKSmart, which predicts human clearance, volume of distribution and half-life directly from molecular structure together with predicted animal PK, with accuracy that its developers report as comparable to industry standard approaches, in a peer reviewed but developer reported evaluation while remaining open-source and freely web-hosted [69]; structure based ADMET predictors such as ADMET-AI return related estimates [27]. The proprietary comparators are the validated suites of Simcyp, GastroPlus and Phoenix WinNonlin, whose advantage lies in regulatory acceptance and support rather than in availability. The division is the familiar one: these free tools are well suited to ranking candidates, projecting an initial human dose and teaching the methods, whereas a dose that enters a regulatory submission still rests on a qualified model and experimental confirmation.

A free starter kit to try

If a reader adopts only a handful of free tools this year, the starter kit in Table 6 spans the recurring tasks of Sections *Statistics and data analysis, Pharmacokinetic and pharmacodynamic modelling, Calibration and standard curves, Structure drawing, SMILES and chemical notation, Physicochemical and ADMET property prediction, Scientific writing, literature work and teaching, Biopharmaceutics: solubility, permeability and BCS classification and formulation development, In silico bioequivalence and digital twins, Allometric scaling and human pharmacokinetic prediction*, and every item can be cited by version in a manuscript. Statistics and scripting rest on R, Python and the point and click JASP and jamovi; pharmacokinetics on the R packages nlme, nlmixr2 and mrgsolve and on the open source PK-Sim and MoBi suite [29]; chemistry on RDKit [18] with the Ketcher and JSME sketchers; and reproducible pipelines on Jupyter and KNIME. For physicochemical, ADMET and biopharmaceutics screening, SwissADME [24], ADMET-AI [27], ProTox [25] and FormulationBCS [71] return estimates in seconds, but they are web services unless installed locally, and proprietary structures should not be submitted to them. For protein structure, the public AlphaFold server and the open reproductions OpenFold, Chai-1 and Protenix [22, 28, 32, 47] avoid licence costs, although the server is restricted to non-commercial use and receives the submitted sequence. For confidential writing and coding, a locally run open weight model keeps text on local hardware. Recording the exact version used is the reproducibility discipline most worth teaching a trainee. The full list of tools with access addresses is provided as Supplementary Table S1.

To see how these free tools combine in practice, consider a graduate student building a first physiologically based pharmacokinetic model. The student might use a conversational assistant to draft and annotate initial R code, SwissADME or ADMET-AI to obtain first-pass physicochemical and permeability estimates from the candidate structure, PK-Sim and MoBi to run the mechanistic simulations, and an established pharmacometric platform to confirm the completed workflow before any result is reported, checking each tool against the next as the analysis proceeds. The pipeline is assembled entirely from freely available components, yet every step remains the responsibility of the modeller rather than of any single program.

Revisiting technological darwinism: which 2023 expectations held?

Because this review is a sequel, it is worth asking plainly how the earlier editorial's predictions have fared [1]. Written within weeks of the public release of ChatGPT, that editorial made several claims: that language models would assist with literature review, data analysis, modelling and interpretation; that AI would accelerate drug discovery and development; that the tools carry biases and limitations and must work alongside human expertise and critical thinking; that their output must be checked for accuracy before entering the published record; that journals would need policies for transparent disclosure; and, through the Darwin framing, that adaptability rather than resistance would decide who benefits [1]. Table 1 sets each expectation against what actually happened

between 2024 and 2026, as a qualitative assessment by the authors and not an empirical audit.

The balance is striking. The directional calls were right, and in two respects the pace exceeded what a cautious reader in 2023 would have predicted: generative design reached the clinic with an AI originated target and molecule in phase IIa [12, 16], and regulators moved from silence to a dedicated credibility framework within 2 years [8]. The editorial also under anticipated one development, namely that openness would migrate from the publishing layer, where this journal has been open access since 1998, to the models themselves through open weight releases such as AlphaFold 3 [21, 22]. What it got most right is the part that has not changed: models still fabricate plausible output, so human verification and proportionate oversight remain the load-bearing requirement [8, 63].

Discussion

Three limitations cut across every stage and every tool. The first is opacity: closed models cannot be inspected, and even open weight models seldom disclose their training data, which complicates the assessment of bias and the work of validation [22, 63]. The second is reproducibility: a result from a versioned commercial service may not be reproducible once the version moves on, whereas an open tool pinned to a release can be rerun deterministically, given fixed seeds and a pinned software environment, and allowing for residual hardware level numerical non-determinism, which is itself an argument for preferring open tools in published methods. As prompt engineering increasingly influences AI output, documenting prompts may become as important for reproducibility as documenting software versions. The third is reliability: generative models fabricate references, structures and numbers with fluent confidence, so any output that informs a decision must be checked against an authoritative source [1, 53]. Retrieval-augmented generation (RAG), in which models ground responses in external databases, offers one strategy for reducing 'hallucination' but still requires verification. The FDA credibility framework formalizes this discipline for the regulated setting, tying the evidence required to the model's context of use and to the consequences of error [8, 56]; the same logic is sound everywhere, namely to match the rigour of validation to the stakes of the decision and to keep a human accountable for the result. A further caveat is user-dependent: even the best tool can be misapplied by a user who lacks basic statistical, pharmacometric or regulatory literacy. A final caveat is temporal: because the landscape shifts rapidly through model updates, acquisitions and licensing changes, the specific tools named in this review should be read as representative snapshots rather than permanent rankings.

These general limitations are borne out by documented failures. A cross-disciplinary audit found data leakage affecting close to three hundred studies across seventeen fields, in several cases producing markedly overoptimistic performance that did not survive correct evaluation [26]. In medicinal chemistry, quantitative structure-activity and ADMET models that score well in cross-validation often generalize poorly to external chemical series, partly because public training sets over represent well studied chemotypes, and a strong discovery metric is no guarantee of clinical success, as the discontinuation of the first AI-designed candidate after phase I shows [16]. Silent updates to commercial models compound the problem by changing outputs without notice. The practical lesson is that a benchmark result is a claim to be checked rather than a guarantee, and that external validation and the honest reporting of negative results matter as much as headline accuracy [26].

These considerations point to a concrete reporting practice. Whenever an AI tool informs an analysis or a manuscript, the methods or Supplementary Material should record the tool name and version, the model or package release date, the access mode (open-source, open-weight, free web service or proprietary), the date of use, the exact prompt or query where a language model was used, the key parameters, seeds and software environment where relevant, the source of any external data or retrieval step, and the human verification step together with final responsibility for the result. Recording these items converts the reproducibility discussion above into a checklist that a reader or auditor or regulator can follow. Underlying all of this is data quality, since an AI tool can be currently no better than the data it learns from, and adherence to the FAIR principles, that data be findable, accessible, interoperable and reusable, may prove as important to dependable performance as advances in model architecture [80]. In practice the records above are best preserved with the ordinary tools of reproducible computing, namely version control such as Git, containerization with Docker, pinned package environments such as Conda, and shareable notebooks such as Jupyter, which together let an analysis be re-executed long after it was first run. Where confidentiality prevents pooling data across institutions, federated learning offers a privacy preserving alternative that trains a shared model on data which never leaves each site, an approach of growing interest for multi-centre clinical and pharmaceutical datasets [81].

Software of unknown provenance (SOUP) and the governance of AI in biopharmaceutics

The reproducibility and opacity concerns above have a concrete regulatory analogue that pharmaceutical scientists will increasingly meet by name. In medical device software, any component that is already developed and generally available, and that was not produced under a controlled development lifecycle, is termed software of unknown provenance (SOUP) under the international standard IEC 62304 [30]. SOUP is not prohibited; instead the manufacturer must inventory it, specify what it is required to do, and manage the residual risk through testing and monitoring proportionate to the harm a failure could cause [30]. The same posture has been proposed for third-party machine-learning models, which are SOUP with the added difficulties that their training data are usually undisclosed and their behaviour can drift after deployment [82]. For computerized systems in regulated manufacturing and laboratory practice, the parallel framework is the risk-based, supplier-leveraging approach of GAMP 5 [31], now being extended to adaptive AI. Adoption therefore depends not only on technical merit, but also on governance, training, procurement and local IT constraints.

This matters directly for the biopharmaceutics tasks of Section *Biopharmaceutics: solubility, permeability and BCS classification and formulation development*. An open predictor used to rank candidates

or teach the BCS is low stakes and needs little beyond sensible checking. The moment a model output is offered in support of a regulatory claim, for instance a predicted BCS class advanced toward an ICH M9 biowaiver, that output becomes a high-stakes use of software whose provenance, version, training data and validation must be documented, exactly as the FDA credibility framework requires for the model's context of use [8, 71]. The open versus closed distinction is then not merely a question of cost. An open tool pinned to a known version, with inspectable code, is easier to qualify and to defend as fit for purpose than a closed service that may change silently, which is the practical reason transparency and reproducibility, rather than licence price, should drive tool choice wherever a result must withstand audit [8, 30, 82]. This does not make open tools automatically qualified: a commercial platform brings a vendor qualification package whereas an open tool must be qualified by the user, and both routes are workable. Neither an open nor a commercial label establishes qualification by itself, which depends on the defined context of use, change control, verification, validation and governance. R based analysis packages have passed FDA review in the R Consortium submissions pilots [83], the open PK-Sim platform has a published automated qualification framework with a geometric mean fold error of 1.4 for CYP3A4 mediated interaction predictions [84], and nlmixr2 has been described and benchmarked in peer reviewed work [85, 86]. Table 3 therefore asks for a qualified platform, commercial or open.

What credibility means: an operational definition

In this review credibility means trust, established through evidence proportionate to risk, in the performance of an AI model for a specific context of use, following the FDA draft guidance [8]. It is not a single statistic. Credibility is not an intrinsic property of a model and is not established by a high coefficient of determination or by agreement with a training set. It should be distinguished from three related concepts: verification, which establishes that an implementation operates according to its specification; validation, which tests whether predictions of the quantities of interest are accurate and reliable enough for the intended use; and reproducibility, which asks whether results can be regenerated under documented conditions. Credibility integrates these and other lines of evidence to decide whether reliance on the model is justified for a specific decision. We divide it into five measurable dimensions (Table 7): predictive performance and robustness, reproducibility, data quality and provenance, regulatory and governance alignment, and human oversight and accountability. Alignment with regulatory standards means that the evidence is organized as the relevant framework expects, for example the FDA seven step process built on model influence and decision consequence [8] or the model risk and model impact assessment of ICH M15 [11]. The rigour required scales with the stakes, so a model that ranks compounds for teaching needs far less evidence than one that supports a dose or a biowaiver. Thresholds are specific to the context of use and should be prespecified in the credibility assessment plan, and Table 7 therefore gives examples of metrics and conventions, such as the two fold error criterion for pharmacokinetic predictions, and not universal limits. Depending on the use, the evidence may need to cover calibration, discrimination or prediction error, uncertainty, applicability domain coverage, external validation, data provenance and representativeness, version control, known limitations, and the consequences of false positive and false negative results. Not every metric suits every model, and the criteria must be matched to the scientific task.

Human oversight and the modular design of AI workflows

Human oversight is a condition for using AI in pharmaceutical science, because the consequences of error fall on patients. The FDA

TABLE 7 **Operational dimensions of credibility for an AI model in pharmaceutical science, with examples of measurable indicators. Thresholds are context specific and should be prespecified in a credibility assessment plan.**

Dimension	Question it answers	Examples of measurable indicators and evidence
Predictive performance and robustness	Is the model accurate, calibrated and stable for the intended inputs?	Error and discrimination metrics on external, scaffold split or time split data (for example geometric mean fold error and the proportion of predictions within two fold for pharmacokinetics; area under the ROC curve and calibration slope for classification); applicability domain coverage; observed coverage of nominal prediction intervals; sensitivity to input perturbation; comparison with a simple baseline [26]
Reproducibility	Can the result be regenerated and independently confirmed?	Identical output under a pinned version, seed and environment, or reported variance across repeated runs; archived code, data and prompts; independent replication or experimental confirmation of key predictions [26]
Data quality and provenance	Are the training and test data fit for the context of use?	Documented sources and curation; leakage checks; representation of the relevant chemical or patient space; FAIR metadata; assessment of bias [26, 80]
Regulatory and governance alignment	Is the evidence organized as the relevant framework expects?	Defined question and context of use; model influence and decision consequence; credibility assessment plan and report; change control and lifecycle monitoring; supplier or SOUP assessment [8, 11]
Human oversight and accountability	Who can detect and override an error?	Named accountable person; defined review points and override authority; decision logging; reviewer training; monitoring for automation bias [87–89]

and EMA principles place human centric design first [10], the EMA reflection paper stresses a risk based and human centred approach [9], and Article 14 of the EU Artificial Intelligence Act requires effective human oversight of high risk systems [90]. Three arrangements should be kept distinct: human in the loop, where a person approves each output before it takes effect; human on the loop, where a person supervises an autonomous process and can intervene, as in a self driving laboratory with defined stopping rules; and human in command, where a governing body decides whether, where and how a system is deployed and retired. The arrangement should follow model influence and decision consequence, so that dose selection, batch release and regulatory claims call for review of each output while low risk screening can run under supervision. Oversight in this sense covers more than checking a final output: it spans the selection of data, the formulation of the scientific question, the evaluation of model suitability, the interpretation of predictions, verification against experimental or authoritative evidence, and authorization of consequential decisions. Literature organization or preliminary screening can accept light touch, periodic review, whereas predictions that inform toxicological risk, clinical dosing, manufacturing release or regulatory claims need stricter verification and documented accountability.

Oversight is only as good as its design. Reviewers need the expertise, time and authority to disagree, otherwise automation bias turns review into a formality, and routine reliance on AI may deskill the workforce so that oversight becomes nominal [88]. Human involvement alone does not guarantee safety, because reviewers also need access to the evidence behind a prediction and clear responsibility for the decision, and plausible but wrong outputs are hard to detect without domain knowledge. Practical safeguards are independent verification against an authoritative source, audited review, documented override and decision logs consistent with data integrity expectations, and reviewer training, as proposed for human in the loop use in ongoing process verification under EU GMP Annex 22 and ICH Q9 [87]. The EMA accepted in 2025 clinical trial evidence generated by an AI tool supervised by a human pathologist, which shows that an explicit oversight arrangement can make AI output acceptable to a regulator [91]. A recent perspective applies the human in the loop principle to drug design, clinical trials, synthesis and manufacturing, analysis and quality control, and formulation, and assigns moderate to high oversight needs to classical machine learning and high oversight needs to deep learning and language models [89]. AI should augment scientific judgment without obscuring the responsibility of the investigators, developers, manufacturers and qualified professionals who rely on its outputs.

Whether each stage should use its own AI tool is a related question. A modular design gives each tool a narrow, validatable context of use, makes errors easier to localize, and creates a checkpoint at each hand off where a human can review and sign off, which strengthens human authority. It also permits a stage specific choice between open and closed tools (Table 4). Separate specialized tools could support target identification, molecular design, retrosynthetic planning, analytical characterization, formulation development, preclinical prediction, pharmacokinetic modelling, clinical trial analysis and manufacturing quality assessment, each with a specified intended use, defined inputs and outputs, and independently assessable performance criteria. The costs are interface and format mismatches, errors that propagate silently when modules are validated only in isolation, and the loss of the cross stage learning that integrated platforms exploit. We therefore recommend a modular architecture with explicit human review at each hand off and end to end validation of the pipeline for the decision it supports. Interfaces should carry explicit acceptance criteria and version control so that an uncertain or invalid output is not propagated automatically, and responsibility for the final decision should remain with qualified scientific and regulatory personnel. Modularity is not synonymous with safety, because errors can accumulate across components and integration introduces risks of its own.

Accuracy versus accessibility: where the performance gap lies

The case for open tools in this review rests on cost, transparency and reproducibility and is not a claim of equal performance. Commercial platforms bring integrated and optimized models, proprietary training data that open tools cannot match, support and indemnification, and in several domains they may remain more accurate (for example molecular simulation, free energy estimation, physiologically based pharmacokinetic modelling and regulated pharmacometric analysis, where tightly integrated workflows, specialized physics based calculations and curated reference datasets are consequential). Where comparative evidence exists it is uneven (Table 8). An independent docking benchmark found that deep learning methods did not outperform classical engines, open (AutoDock Vina) or commercial (CCDC GOLD), in physical plausibility or on novel complexes [50]. The open weight Boltz-2 is reported by its developers to approach free energy perturbation at a small fraction of the cost, but this rests on a preprint and developer chosen benchmarks [48]. A simvastatin case study found comparable predictive performance in the open PK-Sim and the commercial Simcyp, although model development strategies differed [92]. Docking pose validity, ADMET prediction and PBPK simulation test different things and are not interchangeable forms of validation. Leaderboard performance of ADMET-AI on public data [27] does not show that a model generalizes to the chemical series of a company, where proprietary models trained on internal data may do better. Neither licensing model guarantees scientific validity: an open model may be poorly calibrated or used outside its applicability domain, whereas a proprietary model may be difficult to scrutinize independently or to reproduce after an undocumented update. Meaningful comparison therefore requires task specific benchmarking against appropriate reference data, prespecified performance criteria, assessment of uncertainty and applicability domain, and attention to the operational and regulatory requirements of the intended use.

Systematic head to head comparisons of free and commercial tools on identical independent datasets for ADMET, BCS classification, allometric scaling and virtual bioequivalence are largely absent. Our view that free tools are adequate for teaching, ranking and early stage work is therefore a judgment from the available evidence and practice and not a measured equivalence, and a dated, versioned, independent benchmark including both kinds of predictor would be the most useful way to settle the question. The

TABLE 8 Comparative evidence on open and commercial tools, with the status of that evidence. No independent head to head comparison was identified where this is stated.

Domain	Open or free option	Comparator	Evidence and its status
Protein ligand docking	AutoDock Vina; DiffDock [36]	CCDC GOLD (commercial)	Independent benchmark: classical engines, open and commercial, were ahead of deep learning methods in physical validity and on novel complexes [50]
Binding affinity	Boltz-2 (37) (open weight, MIT licence)	Physics based free energy perturbation	Developers report approaching free energy perturbation accuracy at roughly 1000 fold lower computational cost (preprint; independent replication limited)
Structure prediction	OpenFold [28], Chai-1 (32), Protenix [32]	AlphaFold 3 (restricted access)	Developer reported performance near AlphaFold 3; OpenFold peer reviewed, Chai-1 and Protenix preprints; independent comparisons limited
PBPK modelling	PK-Sim and MoBi [29]	Simcyp, GastroPlus	Simvastatin case study [92] found comparable predictive performance; automated qualification framework [84] published for PK-Sim; no broad independent series identified (peer reviewed)
ADMET prediction	ADMET-AI [27], SwissADME [24], ProTox [25]	Commercial predictors	ADMET-AI reports a leading rank on a public leaderboard at publication (developer reported, peer reviewed); no independent comparison with commercial predictors identified
Human PK from structure	PKSmart [69]	Allometric scaling; commercial suites	Developers report accuracy comparable to industry standard approaches (peer reviewed); independent evaluation not identified
Population PK estimation	nlmixr2	NONMEM, Monolix	Peer reviewed description and algorithm benchmarking of nlmixr [85, 86]; open R submission pilots [83]; commercial tools retain greater regulatory familiarity

practical objective is not to maximize openness or proprietary capability in isolation, but to select the tool that offers the most defensible combination of performance, transparency, reproducibility, integration and accountability for the decision at hand.

Security, intellectual property, sustainability and equity

Four further considerations bear on tool choice. Security: open models and pipeline packages are software components that can contain vulnerabilities or malicious code and whose dependencies may become unmaintained, and agentic systems that call external tools widen the attack surface [93], so pinned dependencies, verified checksums, isolated execution and the SOUP style inventory above apply to open and closed tools alike. Intellectual property: licences determine whether outputs, fine tuned weights and derived models can be used commercially, and the ownership and patentability of AI generated molecules and text depend on the jurisdiction and on the human contribution, so the terms of each tool should be read before it is built into a pipeline. A tool's licence is distinct from the ownership or protectability of what it generates. Sustainability: training and large scale inference consume substantial energy and carry a carbon cost, self hosting transfers that cost to the institution, and academic web services depend on grant funded servers that can be withdrawn, as the footnote to Supplementary Table S1 notes. Access to compute is itself unequal. Equity: free tools lower the cost barrier, but only where infrastructure, training and representative local data are also available (including language coverage, affordability and sustainable access in low and middle income settings), so open weights and open data are necessary but not sufficient for equitable access.

Outlook

Forecasting in a fast-moving field is hazardous, but the trajectory of 2024–2026 supports a few disciplined predictions. First, the rentosertib result will be tested by outcome: within this horizon at least one AI-originated drug should report a phase III readout, and the field's credibility will hinge on whether efficacy survives a larger and more heterogeneous population rather than on the discovery speed that first drew attention [12, 16]. Second, the open and closed worlds will continue to converge, with capable open-weight models closing the reasoning gap and structure prediction becoming routine infrastructure through open reproductions [22, 28, 32, 47]; the durable advantage of closed platforms will narrow to integration, support and regulatory familiarity rather than raw capability. Third, regulation will move from guidance to expectation, as the FDA credibility framework and its analogues become a standard part of submissions and SOUP-style inventory and qualification of AI components migrate from medical device practice into routine biopharmaceutics and manufacturing workflows [8, 30, 31, 82]. Fourth, biopharmaceutics will see *in silico* solubility, permeability and BCS classification used earlier and more confidently for triage and developability, while the biowaiver decision itself remains anchored to experimental confirmation [71, 76]. Fifth, the binding constraint will shift from model capability to data and verification: curated, well governed datasets such as the Therapeutics Data Commons, and the discipline of checking and documenting outputs, will matter more than access to any particular model [33]. A further structural shift, still nascent in pharmaceutical work, is the move from standalone language models toward agentic systems that plan and execute multi-step tasks by invoking external tools and data, a transition that emerging open standards such as the Model Context Protocol are beginning to formalize [93]. The logical

TABLE 9 A longer view: a vision of artificial intelligence in pharmaceutical science from 2026 to 2050, offered as a disciplined extrapolation rather than a forecast.

Horizon	Plausible trajectory	What is likely to remain human
2026 to 2035: convergence	Open-weight models match closed ones on most reasoning tasks; the first AI-originated drugs complete phase III and reach the market [16]; self-driving laboratories automate design make test analyze cycles in early discovery; PBPK, virtual bioequivalence and formal credibility assessment become routine in submissions	Causal reasoning, mechanism and accountability for the result; experimental confirmation still anchors approval
2035 to 2050: Integration	End to end generative pipelines link target discovery, molecular design, formulation and trial simulation; multimodal foundation models trained on genomic, proteomic and real-world data propose and rank mechanisms; digital twins shrink early-phase cohorts; manufacturing becomes adaptive and continuously qualified	Validation, provenance governance and safety oversight scale up rather than fade; rare and heterogeneous populations still require real-world confirmation

TABLE 10 Core artificial-intelligence competencies for the pharmaceutical curriculum [97], with the reason each matters.

Competency	Why it matters
Prompt engineering	Eliciting useful, well-scoped output from a language model and making its use reproducible
Validation and verification	Detecting hallucinated references, structures and numbers before they inform a decision
Statistics and pharmacometrics	Judging whether an AI-generated analysis is correct and appropriately reported
Reproducibility	Recording versions, prompts and environments so a result can be re-run and audited [26]
Data governance and FAIR principles	Keeping data secure, well-described and reusable [80], and meeting regulatory expectations
Model bias and limitations	Anticipating failure modes such as poor external validation [26] and over-represented chemotypes
Regulatory and credibility assessment	Matching the evidence for a model to the stakes of the decision it supports [8]
AI ethics and accountability	Ensuring responsible use and that a human remains answerable for the result [88]

extension of agentic systems is largely automated scientific discovery, an early and still unreliable version of which has already been demonstrated end to end and published [94]. The one prediction offered with confidence is the one the original editorial made, namely that the scientists who prosper will be those who adapt, choosing tools by evidence and keeping a human accountable for the result [1]. Beyond this decade the trajectory can only be sketched, and Table 9 offers such a sketch out to 2050, as a disciplined extrapolation rather than a forecast; across every horizon the pattern is the same, namely that capability migrates steadily to the machine while judgment, accountability and the verification of results remain with the pharmaceutical scientist.

With respect to AI self-awareness we see no basis for predicting a date, because self-awareness and consciousness have no agreed operational definition or validated test. The most systematic current approach derives indicators from neuroscientific theories to inform credences, without concluding that present systems are conscious and without excluding that future ones could become plausible candidates [95]. Views differ on whether such features could ever suffice: computational functionalists hold that the right functional organization could be enough, whereas biological naturalists argue that consciousness depends on being a living organism, which makes artificial consciousness unlikely along current trajectories [95, 96]. Present pharmaceutical models are not candidates on either view. Structure predictors and generative design networks are function approximators that run in an input, computation and output cycle with no continuous internal state, and a language model that writes that it is analyzing a molecule reproduces the statistics of language and does not report an inner state. The transformation of drug discovery rests on capability, not awareness. For pharmaceutical use the relevant properties are more mundane and measurable, namely reliability, calibration, controllability and goal directed autonomy, and these are already changing the risk profile whether or not any system is aware. Systems that already plan multistep workflows, invoke external tools, generate hypotheses and coordinate computational and laboratory tasks show expanding operational capability, not subjective experience, and what they require is governance: explicit operating boundaries, traceable actions, independent verification, human authorization before consequential steps, and mechanisms for interruption or rollback. Even if a system were judged aware, accountability would not change, because sponsors, marketing authorization holders, investigators and qualified persons remain legally responsible and regulators assess a model for its context of use and not its inner life. Increasing autonomy, not awareness, is what should shape the oversight arrangements of Section *Human oversight and the modular design of AI workflows*,

and we would revisit the question if a validated and accepted test of machine consciousness emerged.

AI literacy as a core competency

The recurring lesson of this review, that judgment rather than tool use is the enduring skill, carries a clear educational implication. Just as contemporary pharmaceutical curricula incorporated statistics, pharmacokinetics and evidence-based medicine, future curricula will require formal training in AI literacy, model validation, reproducibility, algorithmic governance and regulatory evaluation. Equipping pharmacy undergraduate and graduate students to interrogate an AI tool, to recognize its failure modes and to document its credibility for the decision at hand is now as fundamental as teaching them to design an experiment or interpret a pharmacokinetic concentration vs. time plot, and it is the surest way to ensure that the next-generation of pharmacists and pharmaceutical scientists directs these tools rather than defers to them [1, 97]. Table 10 sets out a concise competency map for this purpose, pairing each skill with the reason it matters. Competence in designing and exercising human oversight, including awareness of automation bias, belongs on that list (Section *Human oversight and the modular design of AI workflows*).

Conclusion

The pharmaceutical scientist of 2026 is not short of AI tools but of clear criteria for choosing among them. The map drawn here supports three durable conclusions. First, for the bench and the classroom, the open ecosystem is now good enough for nearly everything routine, including statistics, PK-PD and PBPK modelling, calibration curves, structure and SMILES handling, and physicochemical and ADMET prediction, and it should be the default wherever transparency and reproducibility matter, provided that confidential inputs stay under local control and that the limits of the comparative evidence are acknowledged. Open tools are not automatically validated or private, and commercial tools are not automatically more accurate or regulator qualified, so the choice should rest on fitness for purpose, with validation and documentation proportionate to the decision supported. Second, closed platforms keep a real advantage in integrated end to end discovery and in the validated, supported workflows of late development and pharmacovigilance, and they have delivered the first clinical proof that generative design can reach patients [12, 16]. Third, the decisive competence is no longer fluency with any single program but the judgment to select the right tool, to validate its output, and to document its credibility in proportion to the decision it supports [8, 53]. Credibility should be defined operationally, as in Table 7, and human oversight should be designed in proportion to the risk of the decision and not added afterwards. That judgment, exercised with the adaptability the earlier editorial called for [1], is the one thing AI cannot supply and the one thing the pharmaceutical scientist must. As computation increasingly complements the bench, the pharmaceutical scientists who prosper will be literate not only in laboratory methods but in model validation and algorithmic governance. The most durable competitive advantage is not access to a particular model, but the ability to evaluate, document and govern whatever model is used. In the end the question is no longer whether artificial intelligence will be used in pharmaceutical science, but whether it will be used transparently, reproducibly and with enough human judgment to merit trust. The future belongs neither to open nor closed AI, but to pharmaceutical scientists who know when each is appropriate and who never delegate scientific judgment.

From so simple a beginning endless forms most beautiful and most wonderful have been, and are being, evolved. Charles Darwin, On the Origin of Species (1859) [98].

Author contributions

ND and RL wrote and edited the manuscript. All authors contributed to the article and approved the submitted version.

Funding

The author(s) declared that financial support was not received for this work and/or its publication.

Conflict of interest

The author(s) declared that this work was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declared that generative AI was used in the creation of this manuscript. The authors wrote the manuscript and carefully researched it and verfied every reference. We compiled all the tables but we did run the manuscript through free Chat GPT for spell checking and also through free AI tools Claude, Google Gemini, and Perplexity to also check grammar syntax and spelling accordingly.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontierspartnerships.org/articles/10.3389/jpps.2026.17484/full#supplementary-material

References

1. Davies NM. Adapting artificial intelligence into the evolution of pharmaceutical sciences and publishing: technological darwinism. *J Pharm Pharm Sci* (2023) 26:11349. doi: 10.3389/jpps.2023.11349

2. Darwin C. *On the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life*. 5th ed. London: John Murray (1869).

3. Darwin C. In: Beagle HMS, editor. *Journal of Researches into the Geology and Natural History of the Various Countries Visited*. London: Henry Colburn (1839).

4. Mottaghi-Dastjerdi N, Soltany-Rezaee-Rad M. Advancements and applications of artificial intelligence in pharmaceutical sciences: a comprehensive review. *Iran J Pharm Res* (2024) 23(1):e150510. doi: 10.5812/ijpr-150510

5. Malheiro V, Santos B, Figueiras A, Mascarenhas-Melo F. The potential of artificial intelligence in pharmaceutical innovation: from drug discovery to clinical trials. *Pharmaceuticals (Basel)* (2025) 18(6):788. doi: 10.3390/ph18060788

6. Dermawan D, Alotaiq N. From lab to clinic: how artificial intelligence (AI) is reshaping drug discovery timelines and industry outcomes. *Pharmaceuticals (Basel)* (2025) 18(7):981. doi: 10.3390/ph18070981

7. Comoglu T. The pioneering impact of artificial intelligence (AI) on pharmaceutical research and drug development (invited editorial). *Pharm Dev Technol* (2024) 29(8): 791–2. doi: 10.1080/10837450.2024.2411492

8. US Food and Drug Administration. Considerations for the use of artificial intelligence to support regulatory decision-making for drug and biological products. Draft guidance for industry. Silver Spring (MD): FDA (2025). In: *Docket FDA-2024-D-4689. Status: Listed by the FDA as Draft Guidance in September 2026*.

9. European Medicines Agency. *Reflection Paper on the Use of Artificial Intelligence (AI) in the Medicinal Product Lifecycle. EMA/CHMP/CVMP/83833/2023*. Amsterdam: EMA (2024).

10. US Food and Drug Administration. European medicines agency. In: *Guiding Principles of Good AI Practice in Drug Development*. Silver Spring, MA and Amsterdam, Netherlands: U.S. Food and Drug Administration (FDA) & European Medicines Agency (EMA) (2026). Available online at: https://www.fda.gov/about-fda/center-drug-evaluation-and-research-cder/artificial-intelligence-drug-development (Accessed November 2, 2026).

11. International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use. M15 general principles for model-informed drug development. In: *ICH Harmonised Guideline, Step 4 Version, Adopted 29 January 2026*. Geneva: ICH (2026).

12. Xu Z, Ren F, Wang P, Cao J, Tan C, Ma D, et al. A generative AI-discovered TNIK inhibitor for idiopathic pulmonary fibrosis: a randomized phase 2a trial. *Nat Med* (2025) 31(8):2602–10. doi: 10.1038/s41591-025-03743-2

13. International Committee of Medical Journal Editors. Recommendations for the conduct, reporting, editing, and publication of scholarly work in medical journals (2026). Available online at: https://www.icmje.org/recommendations/ (Accessed November 2, 2026).

14. Committee on Publication Ethics. Authorship and AI tools: COPE position statement (2023). Available online at: https://publicationethics.org/cope-position-statements/ai-author (Accessed November 2, 2026).

15. Zielinski C, Winker MA, Aggarwal R, Ferris LE, Heinemann M, Lapeña JF, Jr, et al. Chatbots, generative AI, and scholarly manuscripts: WAME recommendations on chatbots and generative artificial intelligence in relation to scholarly publications. *Colomb Med (Cali)* (2023) 54(3):e1015868. doi: 10.25100/cm.v54i3.5868

16. Dharmasivam M, Kaya B, Akinware A, Gholam Azad M, Richardson DR. Leading artificial intelligence-driven drug discovery platforms: 2025 landscape and global outlook. *Pharmacol Rev* (2026) 78(1):100102. doi: 10.1016/j.pharmr.2025.100102

17. Rzymkowski J, Zhang S, Chołuj A, Szkółka A, Dud M, Bruno-Kamiński M, et al. TEMPO improves generality and decreases oxidative deboronation in Chan–Lam couplings of primary sulfonamides. Preprint, not peer reviewed (2026). Available online at: https://cdn.openai.com/pdf/7136bb75-6d47-4834-8fff-c07c0e06708a/tempo-improves-generality-and-decreases-oxidative-deboronation.pdf (Accessed November 2, 2026).

18. Landrum G, the RDKit contributors. RDKit: open-source cheminformatics software (software) (2026). Available online at: https://www.rdkit.org (Accessed November 2, 2026).

19. Ramsundar B, Eastman P, Walters P, Pande V. *Deep Learning for the Life Sciences*. O'Reilly Media: Sebastopol, CA (2019). Available online at: https://deepchem.io (Accessed November 2, 2026).

20. Heid E, Greenman KP, Chung Y, Li SC, Graff DE, Vermeire FH, et al. Chemprop: a machine learning package for chemical property prediction. *J Chem Inf Model* (2024) 64(1):9–17. doi: 10.1021/acs.jcim.3c01250

21. Abramson J, Adler J, Dunger J, Evans R, Green T, Pritzel A, et al. Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature* (2024) 630(8016): 493–500. doi: 10.1038/s41586-024-07487-w

22. Callaway E. AI protein-prediction tool AlphaFold3 is now more open. *Nature* (2024) 635:531–2. doi: 10.1038/d41586-024-03708-4

23. Open Source Initiative. *The Open Source AI Definition, Version 1.0*. Santa Clara (CA): OSI (2024). Available online at: https://opensource.org/ai/open-source-ai-definition (Accessed November 2, 2026).

24. Daina A, Michielin O, Zoete V. SwissADME: a free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules. *Sci Rep* (2017) 7:42717. doi: 10.1038/srep42717

25. Banerjee P, Kemmler E, Dunkel M, Preissner R. ProTox 3.0: a webserver for the prediction of toxicity of chemicals. *Nucleic Acids Res* (2024) 52(W1):W513–W520. doi: 10.1093/nar/gkae303

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. Swanson K, Walther P, Leitz J, Mukherjee S, Wu JC, Shivnaraine RV, et al. ADMET-AI: a machine learning ADMET platform for evaluation of large-scale chemical libraries. *Bioinformatics* (2024) 40(7):btae416. doi: 10.1093/bioinformatics/btae416

28. Ahdritz G, Bouatta N, Floristean C, Kadyan S, Xia Q, Gerecke W, et al. OpenFold: retraining AlphaFold2 yields new insights into its learning mechanisms and capacity for generalization. *Nat Methods* (2024) 21(8):1514–24. doi: 10.1038/s41592-024-02272-z

29. Lippert J, Burghaus R, Edginton A, Frechen S, Karlsson M, Kovar A, et al. Open systems pharmacology community: an open access, open source, open science approach to modeling and simulation in pharmaceutical sciences. *CPT Pharmacometrics Syst Pharmacol* (2019) 8(12):878–82. doi: 10.1002/psp4.12473

30. International Electrotechnical Commission. *IEC 62304:2006/AMD1:2015. Medical Device Software: Software Life Cycle Processes*. Geneva: IEC (2015).

31. International Society for Pharmaceutical Engineering. *GAMP 5: A Risk-based Approach to Compliant GxP Computerized Systems*. 2nd ed. Tampa (FL): ISPE (2022).

32. Chen X, Zhang Y, Lu C, Ma W, Guan J, Gong C, et al. Protenix: advancing structure prediction through a comprehensive AlphaFold3 reproduction. *bioRxiv [Preprint]* (2025). doi: 10.1101/2025.01.08.631967

33. Huang K, Fu T, Gao W, Zhao Y, Roohani Y, Leskovec J, et al. Artificial intelligence foundation for therapeutic science (Therapeutics Data Commons). *Nat Chem Biol* (2022) 18(10):1033–6. doi: 10.1038/s41589-022-01131-2

34. Chandrasekhar V, Rajan K, Kanakam SRS, Sharma N, Weißenborn V, Schaub J, et al. COCONUT 2.0: a comprehensive overhaul and curation of the collection of open natural products database. *Nucleic Acids Res* (2025) 53(D1):D634–D643. doi: 10.1093/nar/gkae1063

35. Rutz A, Sorokina M, Galgonek J, Mietchen D, Willighagen E, Gaudry A, et al. The LOTUS initiative for open knowledge management in natural products research. *eLife* (2022) 11:e70780. doi: 10.7554/eLife.70780

36. Corso G, Stärk H, Jing B, Barzilay R, Jaakkola T. DiffDock: diffusion steps, twists, and turns for molecular docking. In: *International Conference on Learning Representations (ICLR)* (2023). arXiv:2210.01776.

37. Adasme MF, Linnemann KL, Bolz SN, Kaiser F, Salentin S, Haupt VJ, et al. PLIP 2021: expanding the scope of the protein-ligand interaction profiler to DNA and RNA. *Nucleic Acids Res* (2021) 49(W1):W530–W534. doi: 10.1093/nar/gkab294

38. Tu Z, Choure SJ, Fong MH, Roh J, Levin I, Yu K, et al. ASKCOS: open-source, data-driven synthesis planning. *Acc Chem Res* (2025) 58(11):1764–75. doi: 10.1021/acs.accounts.5c00155

39. Schyman P, Liu R, Desai V, Wallqvist A. vNN web server for ADMET predictions. *Front Pharmacol* (2017) 8:889. doi: 10.3389/fphar.2017.00889

40. Buniello A, Suveges D, Cruz-Castillo C, Bernal Llinares M, Cornu H, Lopez I, et al. Open targets platform: facilitating therapeutic hypotheses building in drug discovery. *Nucleic Acids Res* (2025) 53(D1):D1467–D1475. doi: 10.1093/nar/gkae1128

41. Whirl-Carrillo M, McDonagh EM, Hebert JM, Gong L, Sangkuhl K, Thorn CF, et al. Pharmacogenomics knowledge for personalized medicine. *Clin Pharmacol Ther* (2012) 92(4):414–7. doi: 10.1038/clpt.2012.96

42. Pires DEV, Blundell TL, Ascher DB. pkCSM: predicting small-molecule pharmacokinetic and toxicity properties using graph-based signatures. *J Med Chem* (2015) 58(9):4066–72. doi: 10.1021/acs.jmedchem.5b00104

43. Wang W, Wang N, Wu Y, Zhao L, Ye Z, Chen X, et al. An integrated AI-PBPK platform for predicting drug *in vivo* fate and tissue distribution in human and inter-species extrapolation. *Clin Pharmacol & Ther* (2025) 118(4):865–75. doi: 10.1002/cpt.3732

44. Jia S, Wang N, Yang R, Zhao X, Ding Y, Ouyang D. FormulationDE: an updated artificial intelligence system for drug-excipient compatibility prediction. *AAPS Open* (2026) 12:20. doi: 10.1186/s41120-026-00156-4

45. Dong J, Wu Z, Xu H, Ouyang D. FormulationAI: a novel web-based platform for drug formulation design driven by artificial intelligence. *Brief Bioinform* (2023) 25(1): bbad419. doi: 10.1093/bib/bbad419

46. Nobel Committee for Chemistry. *The Nobel Prize in Chemistry 2024: Press Release*. Stockholm: Royal Swedish Academy of Sciences (2024). Available online at: https://

www.nobelprize.org/prizes/chemistry/2024/press-release/ (Accessed November 2, 2026).

47. Chai Discovery. Chai-1: decoding the molecular interactions of life. *bioRxiv [Preprint]* (2024). doi: 10.1101/2024.10.10.615955

48. Passaro S, Corso G, Wohlwend J, Reveiz M, Thaler S, Somnath VR, et al. Boltz-2: towards accurate and efficient binding affinity prediction. *bioRxiv [Preprint]* (2025). doi: 10.1101/2025.06.14.659707

49. Zhavoronkov A, Ivanenkov YA, Aliper A, Veselov MS, Aladinskiy VA, Aladinskaya AV, et al. Deep learning enables rapid identification of potent DDR1 kinase inhibitors. *Nat Biotechnol* (2019) 37(9):1038–40. doi: 10.1038/s41587-019-0224-x

50. Buttenschoen M, Morris GM, Deane CM. PoseBusters: AI-based docking methods fail to generate physically valid poses or generalise to novel sequences. *Chem Sci* (2024) 15(9):3130–9. doi: 10.1039/D3SC04185A

51. Capecchi A, Reymond JL. Classifying natural products from plants, fungi or bacteria using the COCONUT database and machine learning. *J Cheminform* (2021) 13(1):82. doi: 10.1186/s13321-021-00559-3

52. Obrezanova O. Artificial intelligence for compound pharmacokinetics prediction. *Curr Opin Struct Biol* (2023) 79:102546. doi: 10.1016/j.sbi.2023.102546

53. van der Lee M, Swen JJ. Artificial intelligence in pharmacology research and practice. *Clin Transl Sci* (2023) 16(1):31–6. doi: 10.1111/cts.13431

54. Takeda Pharmaceutical Company. *Takeda's Zasocitinib Landmark Phase 3 Plaque Psoriasis Data Show Promise to Deliver Clear Skin in a once-daily Pill, Catalyzing a New Era of Treatment*. Tokyo, Japan: Takeda Pharmaceutical Company (2025). Available online at: https://www.businesswire.com/news/home/20251218728397/en/Takedas-Zasocitinib-Landmark-Phase-3-Plaque-Psoriasis-Data-Show-Promise-to-Deliver-Clear-Skin-in-a-Once-Daily-Pill-Catalyzing-a-New-Era-of-Treatment (Accessed November 2, 2026).

55. Takeda's zasocitinib outperforms deucravacitinib in PsO. *The Pharma Letter* (2026). Available online at: https://www.thepharmaletter.com/biotech-news/takeda-s-zasocitinib-outperforms-deucravacitinib-in-pso (Accessed November 2, 2026).

56. Niazi SK. A critical review of the FDA's draft guidance on artificial intelligence in drug and biological product regulation. *J Chem* (2026) 2026:5202999. doi: 10.1155/joch/5202999

57. US Food and Drug Administration. Health Canada, medicines and healthcare products regulatory agency. Good machine learning practice for medical device development: guiding principles (2026). Available online at: https://www.canada.ca/en/health-canada/services/drugs-health-products/medical-devices/good-machine-learning-practice-medical-device-development.html (Accessed November 2, 2026).

58. Health Canada. *Pre-Market Guidance for Machine learning-enabled Medical Devices*. Ottawa: Health Canada (2026). Available online at: https://www.canada.ca/en/health-canada/services/drugs-health-products/medical-devices/application-information/guidance-documents/pre-market-guidance-machine-learning-enabled-medical-devices.html (Accessed November 2, 2026).

59. Nagar A, Gobburu J, Chakravarty A. Artificial intelligence in pharmacovigilance: advancing drug safety monitoring and regulatory integration. *Ther Adv Drug Saf* (2025) 16:20420986251361435. doi: 10.1177/20420986251361435

60. Al-Azzawi F, Mahmoud I, Haguinet F, Bate A, Sessa M. Developing an artificial intelligence-guided signal detection in the FDA adverse event reporting system (FAERS): a proof-of-concept study using galcanezumab and simulated data. *Drug Saf* (2023) 46(8):743–51. doi: 10.1007/s40264-023-01317-0

61. Yamamoto H, Kayanuma G, Nagashima T, Toda C, Nagayasu K, Kaneko S. Early detection of adverse drug reaction signals by association rule mining using large-scale administrative claims data. *Drug Saf* (2023) 46(4):371–89. doi: 10.1007/s40264-023-01278-4

62. Mishra HP, Gupta R. Leveraging generative AI for drug safety and pharmacovigilance. *Curr Rev Clin Exp Pharmacol* (2025) 20(2):89–97. doi: 10.2174/0127724328311400240823062829

63. Algarvio RC, Conceicao J, Rodrigues PP, Ribeiro I, Ferreira-da-Silva R. Artificial intelligence in pharmacovigilance: a narrative review and practical experience with an expert-defined Bayesian network tool. *Int J Clin Pharm* (2025) 47(4):932–44. doi: 10.1007/s11096-025-01975-3

64. Lin Z, Akin H, Rao R, Hie B, Zhu Z, Lu W, et al. Evolutionary-scale prediction of atomic-level protein structure with a language model. *Science* (2023) 379:1123–30. doi: 10.1126/science.ade2574

65. Watson JL, Juergens D, Bennett NR, Trippe BL, Yim J, Eisenach HE, et al. *De novo* design of protein structure and function with RFdiffusion. *Nature* (2023) 620(7976):1089–100. doi: 10.1038/s41586-023-06415-8

66. Dauparas J, Anishchenko I, Bennett N, Bai H, Ragotte RJ, Milles LF, et al. Robust deep learning-based protein sequence design using ProteinMPNN. *Science* (2022) 378:49–56. doi: 10.1126/science.add2187

67. International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use. Q13 continuous manufacturing of drug substances and drug products. In: *ICH Harmonised Guideline, Step 4 Version. Geneva: ICH* (2022).

68. US Food and Drug Administration, Center for Drug Evaluation and Research. *Discussion Paper: Artificial Intelligence in Drug Manufacturing. Silver Spring (MD): FDA* (2023). Docket FDA-2023-N-0487.

69. Seal S, Trapotsi MA, Mahale M, Subramanian V, Greene N, Spjuth O, et al. PKSmart: an open-source computational model to predict intravenous pharmacokinetics of small molecules. *J Cheminform* (2025) 17(1):147. doi: 10.1186/s13321-025-01066-5

70. Choi S, Han S, Lee SJ, Lim B, Bae SH, Han S, et al. DallphinAtoM: physiologically based pharmacokinetics software predicting human PK parameters based on physicochemical properties, *in vitro* and animal *in vivo* data. *Comput Methods Programs Biomed* (2022) 216:106662. doi: 10.1016/j.cmpb.2022.106662

71. Wu Z, Wang N, Ye Z, Xu H, Chan G, Ouyang D. FormulationBCS: a machine learning platform based on diverse molecular representations for biopharmaceutical classification system (BCS) class prediction. *Mol Pharm* (2025) 22(1):330–42. doi: 10.1021/acs.molpharmaceut.4c00946

72. Ghayoor A, Gilzad Kohan H. Revolutionizing pharmacokinetics: the dawn of AI-powered analysis. *J Pharm Pharm Sci* (2024) 27:12671. doi: 10.3389/jpps.2024.12671

73. Mehvar R. An evaluation of artificial intelligence chatbots as alternatives to specialized software in teaching Bayesian pharmacokinetic analysis. *Am J Pharm Educ* (2026) 90:102003. doi: 10.1016/j.ajpe.2026.102003

74. Wishart DS, Tian S, Allen D, Oler E, Peters H, Lui VW, et al. BioTransformer 3.0: a web server for accurately predicting metabolic transformation products. *Nucleic Acids Res* (2022) 50(W1):W115–W123. doi: 10.1093/nar/gkac313

75. Amidon GL, Lennernäs H, Shah VP, Crison JR. A theoretical basis for a biopharmaceutic drug classification: the correlation of *in vitro* drug product dissolution and *in vivo* bioavailability. *Pharm Res* (1995) 12(3):413–20. doi: 10.1023/A:1016212804288

76. International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use. M9 biopharmaceutics classification System-based biowaivers. In: *ICH Harmonised Guideline, Step 4 Version. Geneva: ICH* (2019).

77. Kollipara S, Martins FS, Jereb R, Krajcar D, Ahmed T. Advancing virtual bioequivalence for orally administered drug products: methodology, real-world applications and future outlook. *Pharmaceuticals (Basel)* (2024) 17(7):876. doi: 10.3390/ph17070876

78. Fisher CK, Smith AM, Walsh JR, Simon AJ, Edgar C, Jack CR, et al. Coalition against major diseases. Machine learning for comprehensive forecasting of Alzheimer's disease progression. *Sci Rep* (2019) 9(1):13622. doi: 10.1038/s41598-019-49656-2

79. European Medicines Agency. Committee for medicinal products for human use. In: *Qualification Opinion for Prognostic Covariate Adjustment (PROCOVA). Amsterdam: EMA* (2022). Available online at: https://www.ema.europa.eu/en/documents/regulatory-procedural-guideline/qualification-opinion-prognostic-covariate-adjustment-procovatm_en.pdf (Accessed November 2, 2026).

80. Wilkinson MD, Dumontier M, Aalbersberg IJ, Appleton G, Axton M, Baak A, et al. The FAIR guiding principles for scientific data management and stewardship. *Sci Data* (2016) 3:160018. doi: 10.1038/sdata.2016.18

81. Rieke N, Hancox J, Li W, Milletari F, Roth HR, Albarqouni S, et al. The future of digital health with federated learning. *NPJ Digit Med* (2020) 3:119. doi: 10.1038/s41746-020-00323-1

82. Stirbu V, Granlund T, Helén J, Mikkonen T. Extending SOUP to ML models when designing certified medical systems. In: *Software Engineering in Healthcare (SEH) Workshop at ICSE 2021* (2021). p. 32–5. Preprint: arXiv:2103.09510.

83. R Consortium R Submissions Working Group. R-based submission pilot 1 to the FDA eCTD package submitted (2021). Available online at: https://github.com/RConsortium/submissions-pilot1-to-fda (Accessed November 2, 2026).

84. Frechen S, Solodenko J, Wendl T, Dallmann A, Ince I, Lehr T, et al. A generic framework for the physiologically-based pharmacokinetic platform qualification of PK-Sim and its application to predicting cytochrome P450 3A4-mediated drug-drug interactions. *CPT Pharmacometrics Syst Pharmacol* (2021) 10(6):633–44. doi: 10.1002/psp4.12636

85. Fidler M, Wilkins JJ, Hooijmaijers R, Post TM, Schoemaker R, Trame MN, et al. Nonlinear mixed-effects model development and simulation using nlmixr and related R open-source packages. *CPT Pharmacometrics Syst Pharmacol* (2019) 8(9):621–33. doi: 10.1002/psp4.12445

86. Schoemaker R, Fidler M, Laveille C, Wilkins J, Hooijmaijers R, Post T, et al. Performance of the SAEM and FOCEI algorithms in the open-source, nonlinear mixed effect modeling tool nlmixr. *CPT Pharmacometrics Syst Pharmacol* (2019) 8(12):923–30. doi: 10.1002/psp4.12471

87. Romero Obon M, Rouaz El Hajoui K, Sancho-Ochoa V, Vargas R, Pérez-Lozano P, Suñé-Pou M, et al. Human-in-the-loop AI use in ongoing process verification in the pharmaceutical industry. *Information* (2025) 16(12):1082. doi: 10.3390/info16121082

88. Gregory PAM, Austin Z. Responsible adoption of artificial intelligence (AI) in pharmacy practice: perspectives of regulators in Canada and the United States. *Pharmacy (Basel)* (2025) 13(6):152. doi: 10.3390/pharmacy13060152

89. Siadati SA, Ebrahimzadeh MA, Dadras A. Human-in-the-loop artificial intelligence in pharmaceutical industry: enhancing efficiency and maintaining oversight. *Scientiae Radices* (2026) 5(2):102–15. doi: 10.58332/scirad2026v5i2a03

90. European Parliament and Council of the European Union. Regulation (EU) 2024/1689 of 13 June 2024 (Artificial Intelligence Act). *Int Leg Mater* (2025) 64 (5), 1234–1381. doi: 10.1017/ilm.2024.46

91. European Medicines Agency. *Artificial Intelligence*. Amsterdam: EMA. Available online at: https://www.ema.europa.eu/node/244999 (Accessed November 2, 2026).

92. Prieto Garcia L, Lundahl A, Ahlstrom C, Vildhede A, Lennernas H, Sjogren E. Does the choice of applied physiologically-based pharmacokinetics platform matter? A case study on simvastatin disposition and drug-drug interaction. *CPT Pharmacometrics Syst Pharmacol* (2022) 11(9):1194–209. doi: 10.1002/psp4.12837

93. Hou X, Zhao Y, Wang S, Wang H. Model context protocol (MCP): landscape, security threats, and future research directions. *arXiv [Preprint]* (2025). doi: 10.48550/arXiv.2503.23278

94. Lu C, Lu C, Lange RT, Yamada Y, Hu S, Foerster J, et al. Towards end-to-end automation of AI research. *Nature* (2026) 651:914–9. doi: 10.1038/s41586-026-10265-5

95. Butlin P, Long R, Bayne T, Bengio Y, Birch J, Chalmers D, et al. Identifying indicators of consciousness in AI systems. *Trends Cogn Sci* (2025) 30:488–501. doi: 10.1016/j.tics.2025.10.011

96. Seth AK. Conscious artificial intelligence and biological naturalism. *Behav Brain Sci* (2025) 49:e315. doi: 10.1017/S0140525X25000032

97. Davies NM, Dua K, Löbenberg R. Educating pharmacists for an artificial intelligence augmented profession: a narrative review of curricular and accreditation imperatives. *Res Social Adm Pharm* (2026) 23:980–91. doi: 10.1016/j.sapharm.2026.07.008

98. Darwin C. *On the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life*. London: John Murray (1859).