JAASP http://www.aaspjournal.org
Journal of Asian Association of Schools of Pharmacy 2026; 15: 47–56
doi: https://doi.org/10.62100/jaasp.2026.15107

Review

An AI platform to accelerate drug discovery: integrating target discovery, molecular generation, and simulation

Yasushi Okuno*

Graduate School of Medicine, Kyoto University, Kyoto, Japan

Received May 11, 2026
Revised July 31, 2026
Accepted August 18, 2026

*Corresponding author
Graduate School of Medicine, Kyoto University, Kyoto, Japan
E-mail: okuno.yasushi.4c@kyoto-u.ac.jp

A part of this review was presented as Keynote Lecture at the 12th Asian Association of Schools of Pharmacy Conference in Saitama, Japan, on August 2–3, 2025.

ABSTRACT

Declining research and development productivity is a structural challenge in the pharmaceutical industry, where the discovery, optimization, and clinical evaluation of a single medicine require long timelines, high cost, and repeated decision-making under uncertainty. Artificial intelligence (AI), big data analysis, and computational simulation are expected to accelerate drug discovery, but isolated applications cannot by themselves transform the whole process. This invited narrative review and platform perspective summarizes the concept and current implementation of an integrated Drug Discovery DX Platform (DXPF) that connects disease information and patient omics data to target gene/protein discovery, lead compound generation, and simulation-based evaluation. In the target-discovery component, Bayesian network and graph-based methods infer disease-specific regulatory networks from omics data and support identification of mechanisms and candidate targets. In the drug-design component, graph convolutional neural network-based compound profile prediction is coupled with ChemTS molecular generation, docking, and molecular dynamics simulation to address multi-objective optimization of potency, absorption, distribution, metabolism, excretion (ADME), toxicity, and structural binding behavior. The platform also incorporates integrated databases and federated learning developed through academia-industry collaboration, enabling model improvement while preserving confidential company data. By connecting these modules through a browser-based interface, DXPF aims to shorten an end-to-end computational cycle from analysis-ready patient data to prioritized candidate structures; the approximately one-week objective is an engineering target under favorable input and computing conditions and excludes compound synthesis and biological, pharmacokinetic, toxicological, and clinical validation. Current evidence is mainly module-level, retrospective, or computational, and prospective benchmarking, uncertainty assessment, and experimental validation are required before effects on attrition or regulatory decision-making can be established. Future expansion to biologics, disease-prevention AI, and preclinical and clinical data could broaden the platform's scope, but these extensions remain development goals.

Key words: artificial intelligence, drug discovery digital transformation, molecular generative AI, federated learning, molecular dynamics simulation

1. Introduction

The pharmaceutical industry is facing a persistent decline in research and development (R&D) productivity. Although biomedical knowledge, experimental technologies, and computational resources have advanced, the number of approved medicines generated per unit of R&D expenditure has declined over several decades (Scannell et al., 2012). Recent analyses emphasize that clinical development failure remains common and that inadequate efficacy, unmanageable toxicity, and poor drug-like properties continue to prevent promising candidates from becoming medicines (Sun et al., 2022). Cost estimates vary widely across studies and assumptions, but surveys and systematic reviews consistently show that the capitalized cost of bringing a new drug to market can reach the scale of hundreds of billions of yen or several billion US dollars (DiMasi et al., 2016; OPIR, 2024; Schlander et al., 2021). In Japan, a 2024 survey estimated the capitalized cost for one new drug approval in global development at 141.5 billion yen under a 10% capital cost

assumption (OPIR, 2024).

Against this background, data-driven drug discovery using AI, machine learning, and simulation has attracted intense attention. Deep learning has produced remarkable achievements, including the AlphaGo system for the game of Go (Silver et al., 2016), large language models (Brown et al., 2020), and AlphaFold for protein structure prediction (Jumper et al., 2021). The 2024 Nobel Prize in Chemistry recognized computational protein design and protein structure prediction, further highlighting the scientific impact of AI-based structural biology (Nobel Prize Outreach, 2024). In drug discovery, AI can support target discovery, compound screening, property prediction, de novo design, image analysis, and clinical data interpretation (Jimenez-Luna et al., 2021; Vamathevan et al., 2019).

However, drug discovery is not a single prediction task. A molecule must engage the right biological target, show sufficient potency and selectivity, possess acceptable absorption, distribution, metabolism, and excretion (ADME) and safety profiles, behave appropriately in complex biological systems, and eventually demonstrate clinical efficacy and safety. Therefore, a single AI model, however accurate in one local task, is unlikely to revolutionize the entire process. The critical issue is integration. The outputs of one computational module should become the inputs of the next, allowing researchers to look several steps ahead, eliminate low-value options earlier, and connect previously fragmented decisions. This review describes a Drug Discovery DX Platform (DXPF) developed to meet that need by combining target discovery, molecular design, property prediction, federated data use, and simulation in a unified workflow.

This article is an invited narrative review and platform perspective. It summarizes established methods that underpin DXPF, describes selected implementation examples, and critically delineates current evidence, limitations, and validation needs; it is not a systematic review or a report of prospective end-to-end clinical validation.

2. Concept of the Drug Discovery DX Platform

The central concept of DXPF is to construct a semi-automated computational foundation that links the early discovery process from disease information to target identification and lead compound generation. Conventional drug discovery often involves separate teams for disease-mechanism analysis, target selection, compound screening, medicinal chemistry, ADME evaluation, toxicology, and simulation. Each step can use sophisticated tools, but the discontinuity of data, formats, expertise, and decision criteria produces inefficiencies. DXPF seeks to reduce these discontinuities by workflowing AI and simulation applications so that the result of one step is immediately usable by the next step.

In its current small-molecule-oriented implementation, the platform contains two major parts: a target protein-gene discovery part and a drug design part. The target discovery part starts from disease names, patient-derived omics data, or both. It estimates biological networks, detects disease-specific networks, and infers candidate disease mechanisms and drug targets. The drug design part receives the candidate target protein and uses compound profile prediction AI,

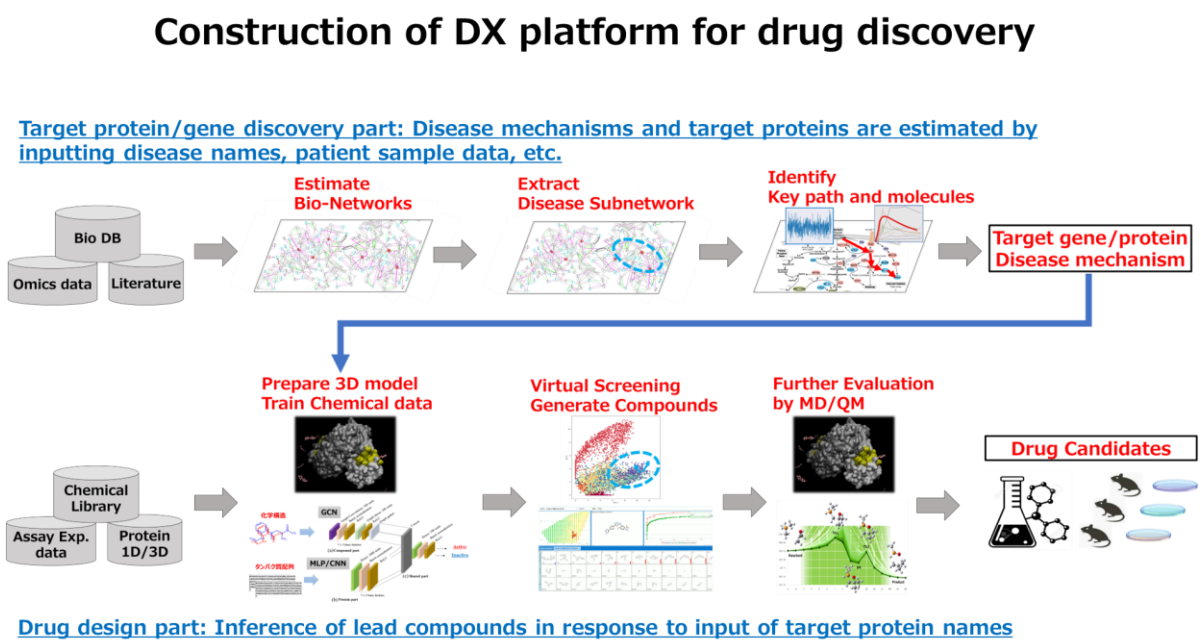

Figure 1. Conceptual workflow of the Drug Discovery DX Platform (DXPF). Disease information and patient-derived omics data are connected to target discovery, compound profile prediction, molecular generation, molecular dynamics simulation, and experimental validation. The diagram is a conceptual architecture: implemented modules and selected demonstrations do not yet constitute a prospectively validated end-to-end workflow, and experimental validation remains outside the automated computational cycle.

Table 1. Representative components of DXPF and their functions. The final column distinguishes available evidence from principal limitations and validation needs.

Component	Main methods	Primary outputs	Role in the platform	Current status and principal validation needs
Target discovery	Bayesian networks, graph convolutional networks (GCNs), omics analysis	Disease-specific networks; candidate genes and proteins	Connects disease data to biological mechanisms and target hypotheses	Retrospective network and subtype analyses are available; causal relevance, human-genetic support, tractability, safety, and perturbational validation remain target- and disease-specific.
Compound profile prediction	GCN, compound-protein interaction (CPI) models, ADME and toxicity prediction	On/off-target activity, ADME, and toxicity scores	Provides rapid multi-endpoint scoring for design and triage	Models cover multiple endpoints; external and assay-stratified validation, calibration, uncertainty, applicability domain, and transfer to new chemical series must be established per endpoint.
Molecular generation	ChemTS, recurrent neural networks (RNNs), MCTS, docking-based rewards	Novel chemical structures optimized for target profiles	Searches chemical space under multiple constraints	Computational demonstrations include CDK2 and DDR1; platform-wide synthesis success, chemical novelty/IP review, experimental hit rate, and property confirmation have not yet been established.
Federated learning	Distributed training and model-parameter sharing	Improved global models without centralizing raw company data	Addresses data scarcity while preserving confidentiality	DAIIA and kMoL demonstrate collaborative infrastructure; endpoint-specific benefit over local models, privacy risk, data heterogeneity, and model drift require continuing evaluation and governance.
Simulation/high-performance computing (HPC)	Docking, molecular dynamics, Fugaku/high-performance computing	Dynamic binding behavior and final candidate ranking	Adds physical validation to AI-generated candidates	Large-scale HPC and selected-target demonstrations are available; results remain sensitive to structures and simulation protocols and require convergence analysis and experimental confirmation.

molecular generative AI, docking, and molecular dynamics (MD) simulation to propose and prioritize candidate compounds. This structure allows the platform to move from patient information to candidate drug structures without leaving the computational workflow.

Most algorithmic classes used in DXPF, including Bayesian networks, graph neural networks, Monte Carlo tree search (MCTS), docking, molecular dynamics (MD) simulation, and federated learning, are established in the literature. Components developed by the author's group and collaborators include kGCN, ChemTS/ChemTSv2, and kMoL. The distinctive contribution proposed for DXPF is therefore platform-level: standardized data exchange and workflow orchestration that link patient-derived omics and target hypotheses to multi-endpoint prediction, molecular generation, structure-based simulation on high-performance computing (HPC) resources, and a browser interface. This review does not claim that each component is novel or superior in isolation; the value of the integrated workflow must be tested by end-to-end benchmarks against conventional and single-module workflows.

The design philosophy also recognizes a practical limitation: integrating all stages of drug development at once is not realistic. For this reason, the platform initially focuses on the red-zone of early discovery, namely target gene/protein discovery and lead compound search/design. This focus is important because early decisions determine much of the later risk. Incorrect targets, weak selectivity, or poor ADME and toxicity profiles are difficult to rescue after extensive experimental investment. By combining rapid AI-based exploration with more precise simulation and experimental validation, DXPF is intended to improve the quality of early decisions rather than merely accelerate existing trial-and-error workflows. Predictions at this stage are hypotheses for selection; they are not validated leads or development candidates.

3. Target protein-gene discovery through disease networks

A major challenge in target discovery is that diseases are rarely caused by isolated molecular changes. Genomics, transcriptomics, proteomics, and other omics technologies can measure thousands of variables in patient or disease samples, but conventional analyses often focus on the abundance of individual genes or proteins. Such analyses are essential, but they may miss the network-level regulatory changes that distinguish disease subtypes, severity, prognosis, or therapeutic response. Network medicine has long emphasized that disease modules and molecular interactions can provide a more system-level view of pathophysiology (Barabasi et al., 2011).

In the target discovery component of DXPF, Bayesian network methods are used to estimate gene regulatory networks from transcriptome data. A key feature is the construction of patient-specific or sample-specific networks from a larger estimated network. This approach attempts to

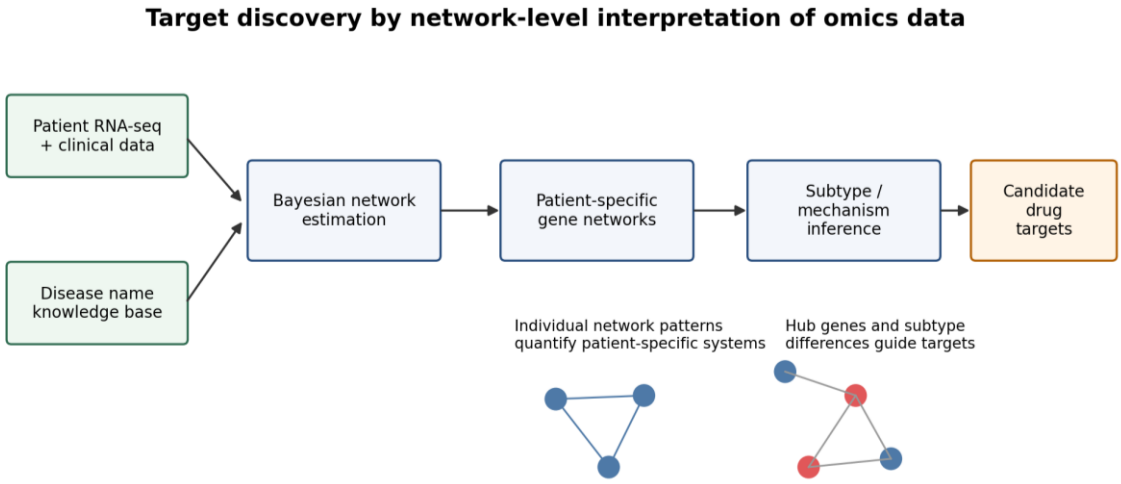

Figure 2. Network-level target discovery. Patient omics data are used to estimate regulatory networks, derive patient-specific networks, detect subtypes or mechanisms, and prioritize candidate target genes/proteins. Network associations generate target hypotheses; they do not by themselves establish causality, druggability, safety, or therapeutic efficacy.

quantify how the regulatory system differs in each patient rather than treating all samples as if they share a single average network. Patient-specific network patterns can then be clustered to reveal subgroups, used to identify hub genes, and interpreted to infer disease mechanisms. A related study demonstrated that patient-specific gene regulatory networks estimated from transcriptome data could identify clinically meaningful cancer subtypes in The Cancer Genome Atlas datasets, including stomach adenocarcinoma (Nakazawa et al., 2021).

Representative platform examples illustrate the rationale. In SARS-CoV-2 infected cells, network analysis was used to separate gene networks associated with severe and mild conditions and to identify genes linked to existing therapeutic agents, including baricitinib and dexamethasone. In gastric adenocarcinoma, RNA sequencing (RNA-seq) data from 365 patients were analyzed to estimate individual patient networks; clustering based on network structures produced three subclasses associated with different survival patterns and hub-gene structures. These examples show how network representations may support both target discovery and patient stratification.

These network outputs should be treated as target hypotheses rather than evidence of causality. Downstream prioritization should combine network centrality and reproducibility with orthogonal evidence, including human-genetic support, replication across cohorts, disease- and tissue-specific expression, causal perturbation data, tractability and druggability, structural information, predicted safety liabilities, and assay feasibility. Human-genetic support has been associated with higher clinical success at the portfolio level (Nelson et al., 2015), but no single criterion is sufficient. Candidate targets therefore require perturbational validation in relevant cells or models and explicit go/no-go criteria before compound generation. This validation framework is a design requirement and has not yet been benchmarked prospectively across DXPF projects.

4. Compound profile prediction by graph-based AI

Once candidate targets have been identified, the platform must evaluate which chemical structures are likely to show useful biological activity and acceptable properties. Traditional cheminformatics methods represent molecules using predefined substructures, descriptors, or fingerprints. These features are powerful but require manual design and can miss aspects of molecular topology. Graph convolutional networks (GCNs) and related message-passing neural networks address this limitation by treating molecules as graphs whose nodes are atoms and whose edges are bonds (Gilmer et al., 2017; Kipf & Welling, 2017).

The kGCN framework was developed as an open-source graph-based deep learning framework for chemical structures. It supports prediction tasks and provides interfaces for both computational specialists and users with limited programming experience (Kojima et al., 2020). In the compound-protein interaction (CPI) setting, the platform can combine a GCN-based molecular representation with a neural representation of protein sequences, such as a convolutional neural network. The model receives a compound-protein pair and predicts whether the compound is likely to bind the protein. This multimodal formulation is consistent with earlier end-to-end CPI learning approaches using molecular graphs and protein sequences (Tsubaki et al., 2019).

DXPF extends activity prediction beyond simple target binding. The compound profile prediction AI is designed to predict on-target activity, off-target liability, ADME properties, and toxicity. The ADME parameters considered in this platform include solubility, fraction unbound in plasma (fu,p) for human, rat, and mouse, Caco-2 permeability (Papp), intrinsic clearance (CLint) for human and mouse,

cytochrome P450 inhibition, volume of distribution at steady state (Vdss), and P-glycoprotein activities. Toxicity endpoints include genetic toxicity, such as the Ames test and in vitro micronucleus test, phototoxicity, mitochondrial toxicity, and phospholipidosis. Treating these endpoints together is essential because real drug design is not the maximization of potency alone but a multi-objective optimization problem.

Predicted endpoints are not interchangeable with measured properties. ADME and toxicity training data can be assay-specific, censored, imbalanced, incomplete, and concentrated in particular chemical scaffolds; performance may therefore degrade for a new series. Each endpoint requires scaffold- and time-aware external validation, calibration, uncertainty estimation, and an applicability-domain assessment (Sheridan, 2012). Conflicting endpoint predictions should be managed with pre-specified multi-objective constraints, confidence thresholds, and traceable decision rules rather than a single opaque aggregate score. Low-confidence or out-of-domain candidates should be flagged for measurement, not automatically advanced or discarded.

5. Molecular generative AI and multi-objective optimization

Molecular generative AI is used to propose new chemical structures that satisfy a desired profile. Whereas large language models generate sequences of words or tokens, molecular generative models generate molecular strings, graphs, or fragments that correspond to chemical structures. DXPF uses ChemTS, a molecular generation approach that combines a recurrent neural network with Monte Carlo tree search (MCTS), a reinforcement learning technique also associated with game-playing AI (Ishida et al., 2023; Yang et al., 2017).

In the platform workflow, ChemTS does not generate molecules blindly. Candidate molecules are evaluated by prediction models for target activity, off-target effects, ADME, and toxicity, and these scores are integrated into a reward function. The reward is then fed back to the generator, which proposes improved candidate structures in subsequent iterations. This loop allows the platform to search chemical space under multiple constraints rather than optimizing only one endpoint. For example, a candidate with high predicted potency but poor permeability, high intrinsic clearance, or high phototoxicity risk should be penalized relative to a balanced candidate.

A representative demonstration used cyclin-dependent kinase 2 (CDK2), a kinase associated with cancer growth, to illustrate this concept. After random seed compounds were generated, the AI repeatedly predicted CPI scores and generated subsequent structures; the best-scoring structure was then selected as a candidate that could bind CDK2. This demonstration is conceptually simple, but the underlying computation searches a large space of possible molecules. DXPF further incorporates structure-based molecular generation in which docking scores, rather than only AI prediction scores, are used as rewards for ChemTS. In this mode, the generator can use three-dimensional information about the target protein and binding pocket to propose molecules optimized for structural interaction.

The final design goal is a drug design AI that can generate candidate compounds while simultaneously considering potency, off-target avoidance, ADME properties, toxicity, and three-dimensional binding activity. This is the core drug-design advantage of an integrated platform: generative AI supplies large numbers of candidates, predictive AI provides rapid multi-endpoint scoring, docking introduces structural

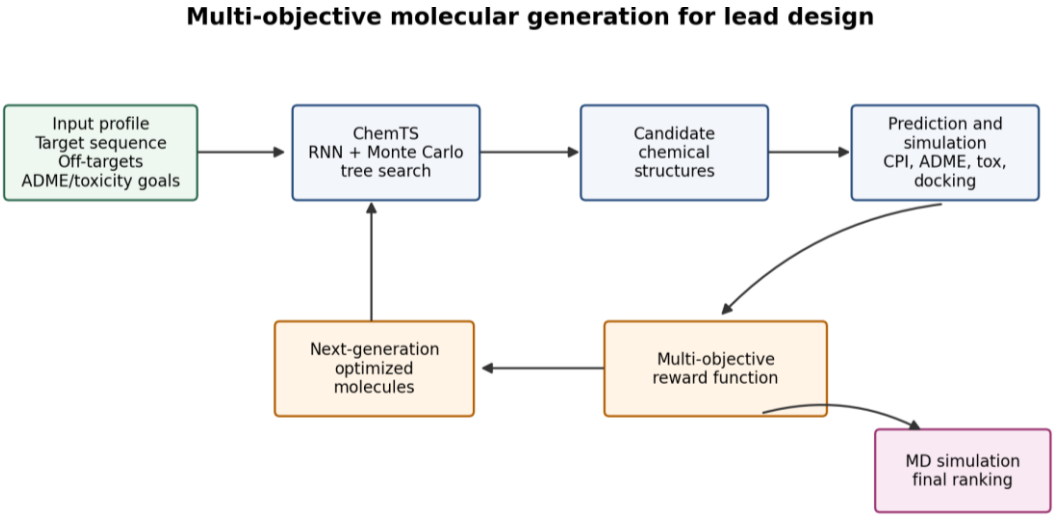

Figure 3. Multi-objective molecular generation. ChemTS proposes candidate structures, predictive AI and docking evaluate target activity, off-target liability, ADME, toxicity, and structural binding, and the reward is fed back to generate improved compounds. The feedback loop has been demonstrated computationally for selected targets; synthetic feasibility, novelty, measured activity, and safety must be evaluated before generated structures are considered leads.

constraints, and MD simulation can be used to evaluate a smaller set of prioritized molecules in greater detail.

Generated structures are proposals, not automatically viable compounds. Candidates should be filtered for valence and reactivity, chemical stability, structural alerts, synthetic accessibility, retrosynthetic-route plausibility, starting-material availability, chemical novelty, and intellectual-property exposure, followed by medicinal-chemistry review. High generative-model scores do not necessarily imply synthesizability (Gao & Coley, 2020). The current DXPF demonstrations do not establish a platform-wide synthesis success rate or experimental hit rate. Prospective evaluation should report the fractions of valid, unique, novel, and route-feasible molecules; synthesis success; biochemical and cellular hit rates; and measured property profiles.

6. Federated learning and the data ecosystem

The performance of machine learning models depends strongly on the quality, quantity, and diversity of training data. In drug discovery, this creates a structural challenge. Public data are valuable but incomplete, while proprietary company data are often larger, more relevant to practical assays, and confidential. Japanese pharmaceutical companies also face data-volume limitations compared with larger international firms, making cross-company collaboration attractive but difficult.

To address this issue, the DAIIA project (Data and AI-driven Drug Innovation Alliance for Improvement of Value), supported by the Japan Agency for Medical Research and Development (AMED), promoted industry-academia collaboration for next-generation drug discovery AI. The project aimed to build an AI-based drug discovery support foundation and an integrated drug discovery AI platform using public data and participating company data for on/off-target activity, pharmacokinetics, and toxicity (AMED, 2025). A key technology was federated learning, in which each company trains models locally and exchanges only model parameters, not raw proprietary data (McMahan et al., 2017).

Federated learning is particularly suitable for drug discovery because it allows model improvement while limiting the risk that competitors access individual companies' data. In this framework, more than 32 million data points were gathered across compound-protein interactions, ADME, and toxicity, and 17 Japanese pharmaceutical companies participated in the DAIIA framework. Similar international efforts, such as MELLODDY, have shown that large-scale cross-company federated learning can improve quantitative structure-activity relationship (QSAR) modeling while preserving confidentiality (Heyndrickx et al., 2023; Oldenhof et al., 2023).

The kMoL library is an open-source machine and federated learning library developed for drug discovery pipelines. It includes machine learning models, Bayesian optimization, explainability functions, and federated learning mechanisms, and is available on GitHub (Cozac et al., 2025). In the context of DXPF, such tools provide a practical implementation layer for building, evaluating, and improving compound profile prediction AI while maintaining data security. kMoL has been evaluated in local and distributed benchmark experiments, but these software-level results should not be interpreted as evidence that the complete DXPF workflow improves project outcomes.

Federated learning reduces the need to centralize raw data,

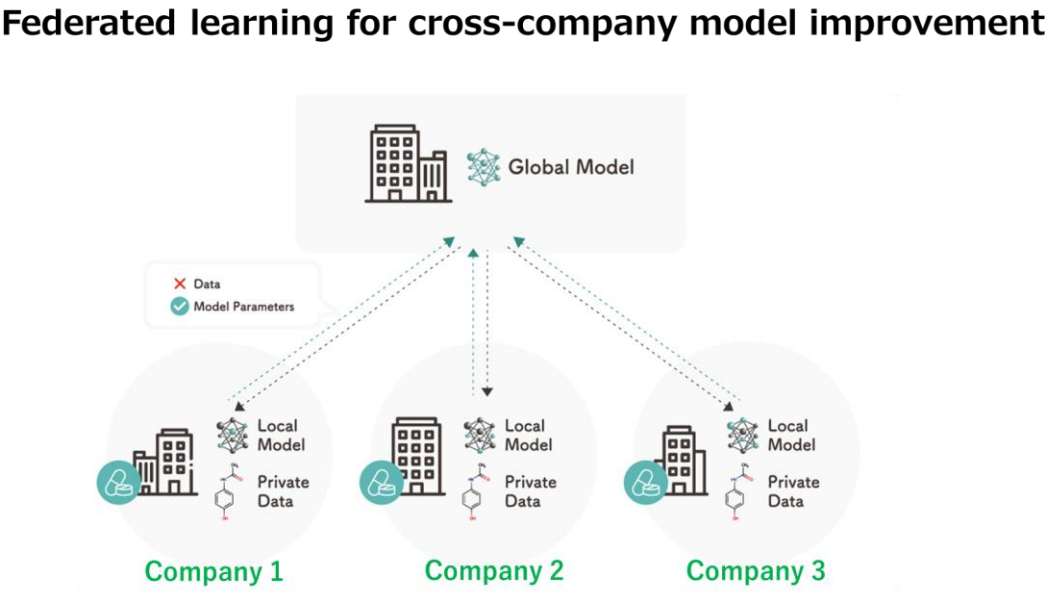

Figure 4. Federated learning model for drug discovery. Individual companies keep proprietary data inside their organizations, while model parameters are exchanged and aggregated to improve a global model. Parameter exchange limits raw-data movement but does not eliminate privacy, heterogeneity, or generalizability risks; performance must be evaluated for each endpoint and deployment setting.

but it does not automatically resolve assay and label heterogeneity, class imbalance, data leakage, privacy attacks, uneven partner contributions, or model drift. Evaluation should therefore be endpoint-specific and compare local-only, pooled where legally and operationally permissible, and federated models on held-out cross-company and temporal test sets. Governance should also define access, auditability, privacy testing, model updates, and responsibility for erroneous predictions.

7. Simulation-based validation using high-performance computing

AI and docking methods can evaluate many candidate compounds quickly, but their speed comes with limitations in physical detail. Rigid-body docking often simplifies protein flexibility, solvent effects, and dynamic interactions. Molecular dynamics simulation addresses these limitations by simulating atomic motion of proteins, compounds, and solvent over time. MD is computationally expensive, but it can provide detailed information about binding stability, protein conformational changes, and ligand behavior in a dynamic environment (Hollingsworth & Dror, 2018; Salo-Ahen et al., 2021).

DXPF therefore uses a staged strategy. Low-cost AI and docking calculations are used to generate and score large numbers of compounds, after which MD simulations are applied to a smaller set of promising candidates. This combination uses each method where it is strongest: AI and docking provide throughput, while MD provides higher-fidelity evaluation of prioritized candidates.

MD outputs should be interpreted as ranking evidence rather than definitive proof of binding or efficacy. Results can depend on the starting protein structure, missing residues, protonation and tautomer states, force field, water model, boundary conditions, simulation length, replicate number, sampling adequacy, and the metric used for ranking. A credible workflow should document these choices, test convergence and replicate consistency, perform sensitivity analyses when appropriate, and confirm prioritized interactions with orthogonal binding and functional experiments (Hollingsworth & Dror, 2018; Salo-Ahen et al., 2021).

A representative example is the simulation-driven search for COVID-19 therapeutic candidates using the Fugaku supercomputer. In this project, 2,128 existing drugs were evaluated by MD simulation against SARS-CoV-2-related proteins, including the main protease, and several promising compounds were identified (Okuno, 2022; RIKEN R-CCS, 2022). Niclosamide was highlighted as a promising candidate in simulations and has also shown cellular activity and has been investigated in multiple clinical trials. This example illustrates how high-performance computing can extend virtual screening beyond static docking by incorporating dynamic molecular behavior. This was a large-scale simulation demonstration rather than prospective end-to-end DXPF validation, and computational prioritization alone does not establish cellular activity or clinical efficacy.

8. Implementation status and future perspectives

The current DXPF implementation is being developed as a workflow system that links many individual applications. In the current implementation, the target discovery part consists of 14 modules and 20 core applications, while the drug design part consists of 6 modules and 35 core applications. Connecting these 55 applications requires common data specifications, provenance tracking, failure handling, and reproducible workflow execution. This operational orchestration, rather than a new algorithm in every module, is the intended distinction between DXPF and a collection of unrelated tools.

A browser-based interface is being developed so that users who are not AI or computational science specialists can operate optimized workflows. In a demonstration using discoidin domain receptor 1 (DDR1) kinase, the user entered the target gene name, obtained three-dimensional protein structures from the Protein Data Bank, selected a structure, automatically estimated the binding pocket, and used ChemTS to generate candidate molecules. The system generated 215 candidate structures and displayed the top-ranking compounds and predicted complex structures. This demonstration establishes browser-based workflow execution and candidate ranking, but the 215 structures were computational proposals; no platform-wide synthesis success or experimental hit rate is inferred. The interface may lower technical barriers, but non-specialists still require endpoint-specific guidance, uncertainty flags, provenance, and explicit indications for experimental review.

Current timing figures are approximate engineering estimates from ongoing demonstration workflows, not median turnaround times from a standardized multi-project benchmark. Under conditions in which quality-controlled patient omics data, required reference data, pretrained models, an experimentally or structurally defined target, and adequate computing resources are already available, target-gene inference has taken approximately two days, compound generation and multi-endpoint scoring approximately three days, and higher-cost MD simulation approximately three days. These stages may vary or overlap, and the approximately one-week goal is an aspirational target for computational hypothesis generation and prioritization only. It excludes sample generation and quality control, target confirmation, compound synthesis, formulation, biochemical and cellular assays, pharmacokinetics, toxicology, manufacturing, clinical testing, and candidate nomination. Prospective reporting should therefore include input readiness, hardware, workflow version, wall-clock time, failure rate, the number of candidates entering and leaving each stage, and experimental confirmation.

8.1. Current evidence, limitations, and validation priorities

Evidence for DXPF currently combines peer-reviewed module-level methods, retrospective biological analyses, software and infrastructure demonstrations, and selected computational case studies. A prospective platform-level study comparing the complete workflow with conventional or isolated-module baselines across multiple projects has not yet been reported; consequently, no general claim about improved experimental hit rate, reduced late-stage attrition, or reproducible one-week turnaround is warranted. Table 2 summarizes the quantitative information available in this review and the corresponding validation boundary. Priority benchmarks should be pre-specified and include enrichment or ranking performance, calibration and uncertainty, scaffold- and time-split external validation, computational cost, synthesis success, biochemical and cellular hit rates, false-positive and failure analyses, and downstream pharmacokinetic and safety outcomes.

For routine use, the browser layer should implement role-based workflow templates, data-quality checks, model and workflow version provenance, applicability-domain and uncertainty flags, intermediate-result visualization, and mandatory review gates. Users should be shown when an output is out of domain, which evidence supports a rank, and which experimental test is required next. These safeguards are design priorities; their effectiveness with non-specialist users remains to be evaluated through usability studies and error analyses.

8.2. Relationship to downstream development and regulatory evidence

Better early target selection and earlier rejection of candidates with predicted off-target, ADME, toxicity, or structural liabilities could reduce avoidable downstream work, but this causal pathway has not yet been demonstrated for DXPF. Computational prioritization may inform selection of compounds for synthesis and testing; it does not itself establish safety, efficacy, manufacturability, investigational new drug (IND) readiness, or regulatory acceptability. If an AI model is later used to generate evidence for regulatory decision-making, its context of use, model risk, data representativeness, performance criteria, validation, traceability, and change control would require documentation commensurate with that use (U.S. Food and Drug Administration, 2025). Thus, the realistic near-term contribution of DXPF is better-documented hypothesis generation and candidate triage, followed by conventional experimental, nonclinical, clinical, and regulatory evidence generation.

8.3. Near-term and longer-term development

Development priorities can be separated by horizon. Near-term work comprises stable integration of the current small-molecule modules, standardized benchmarking, uncertainty reporting, user guidance, and links to experimental feedback. Medium-term development could extend validated workflows to antibodies, peptides, oligonucleotides, and other modalities, which require modality-specific representations and assays rather than simple reuse of small-molecule models. Longer-term possibilities include connecting disease-risk or prevention models and preclinical and clinical data to create learning loops across discovery and development. These extensions are not current platform capabilities and should be evaluated against modality-specific and regulatory requirements. They are aligned with national discussions on strengthening Japan's drug discovery capability and building a sustainable innovation ecosystem (Cabinet Secretariat, 2024).

Table 2. Representative DXPF-related demonstrations, quantitative information reported in this review, and present validation boundaries.

Representative example	Input and workflow	Quantitative output reported	Present interpretation and next validation
Gastric adenocarcinoma network analysis	RNA sequencing from 365 patients; patient-specific regulatory networks and clustering	Three network-based subclasses with different survival patterns and hub-gene structures	Retrospective association supports subtype discovery; cohort replication, target causality, druggability, and perturbational validation remain required.
DDR1 browser workflow	Target name, Protein Data Bank structures, automated pocket estimation, ChemTS generation, and ranking	215 candidate structures generated	Demonstrates workflow execution and ranking; candidates remain computational until synthesis, assay testing, and comparison with baselines.
DAIIA federated-learning ecosystem	Distributed training across participating Japanese pharmaceutical companies without centralizing raw data	More than 32 million compound-protein interaction, ADME, and toxicity data points; 17 companies in the collaboration	Demonstrates data scale and infrastructure; endpoint-specific gains over local models, external generalizability, privacy risk, and drift require reporting.
Fugaku COVID-19 simulation	Molecular-dynamics evaluation of existing drugs against SARS-CoV-2-related proteins	2,128 drugs screened; candidates including niclosamide prioritized	Demonstrates large-scale simulation; it is not prospective end-to-end new-molecule validation and does not prove cellular or clinical efficacy.

The broader future of AI in research also includes large language models, multimodal biomedical foundation models, and laboratory robotics. Scaling laws and emergent behavior in large language models suggest that general-purpose AI systems may increasingly support hypothesis generation, literature interpretation, experiment planning, and manuscript preparation (Kaplan et al., 2020; Wei et al., 2022). When connected to robotics and automated laboratories, such systems could create iterative loops in which AI proposes hypotheses, robots perform experiments, AI analyzes results, and the cycle repeats. For drug discovery, the most productive near-term direction is not to replace human researchers but to connect human expertise with AI, simulation, and automation in a transparent and experimentally validated workflow. Within DXPF, these remain longer-term research directions and will require separate technical, usability, reproducibility, safety, and governance evaluation.

9. Conclusion

Drug discovery requires coordinated decisions across biology, chemistry, pharmacology, toxicology, computation, and clinical science. The Drug Discovery DX Platform reviewed here is designed to address this coordination problem by connecting target discovery, graph-based compound profile prediction, molecular generative AI, federated learning, and high-performance simulation into a single workflow. The platform aims to move from disease and patient information to target proteins and candidate compound structures rapidly, while maintaining the scientific rigor needed for experimental validation. At present, the evidence base is predominantly module-level, retrospective, and computational; platform-level gains over conventional workflows remain to be demonstrated prospectively.

The central proposition is that AI may become more useful in drug discovery when individual applications are connected. A target discovery model, a CPI predictor, a molecular generator, a docking engine, an ADME and toxicity model, and an MD simulator each solve only part of the problem. When their outputs and inputs are linked, however, they can support multi-step reasoning, early risk reduction, and rational prioritization. Continued development of DXPF and similar platforms could improve early discovery only if prospective multi-project benchmarks show reproducible gains over conventional or single-module workflows and prioritized candidates survive synthesis, experimental, translational, and regulatory evaluation. Transparent validation, uncertainty reporting, and analysis of failures are therefore as important as speed.

Conflict of Interest

The author declares no conflict of interest.

References

AMED. Industry-academia collaboration for next-generation AI drug discovery (DAIIA), 2025. Available from: https://www.amed.go.jp/program/list/11/02/001_02-04.html

Barabasi AL, Gulbahce N and Loscalzo J. Network medicine: a network-based approach to human disease. Nat Rev Genet. 2011; 12: 56-68.

Brown TB, Mann B, Ryder N, Subbiah M, Kaplan J, Dhariwal P, et al. Language models are few-shot learners. Adv Neural Inf Process Syst. 2020; 33: 1877-1901.

Cabinet Secretariat. Council of the Concept for Early Prevalence of the Novel Drugs to Patients by Improving Drug Discovery Capabilities. Interim Report: Establish a top-level drug discovery site that contributes to the health of people worldwide, 2024. Available from: https://www.cas.go.jp/jp/seisaku/souyakuryoku/pdf/interim_report.pdf

Cozac R, Hasic H, Choong JJ, Richard V, Beheshti L, Froehlich C, et al. kMoL: an open-source machine and federated learning library for drug discovery. J Cheminform. 2025; 17: 22.

DiMasi JA, Grabowski HG and Hansen RW. Innovation in the pharmaceutical industry: new estimates of R&D costs. J Health Econ. 2016; 47: 20-33.

Gao W and Coley CW. The synthesizability of molecules proposed by generative models. J Chem Inf Model. 2020; 60: 5714-5723.

Gilmer J, Schoenholz SS, Riley PF, Vinyals O and Dahl GE. Neural message passing for quantum chemistry. Proc 34th Int Conf Mach Learn. 2017; 70: 1263-1272.

Heyndrickx W, Mervin LH, Morawietz T, Sturm N, Friedrich L, Zalewski A, et al. MELLODDY: cross-pharma federated learning at unprecedented scale unlocks benefits in QSAR without compromising proprietary information. J Chem Inf Model. 2023; 63: 2331-2344.

Hollingsworth SA and Dror RO. Molecular dynamics simulation for all. Neuron. 2018; 99: 1129-1143.

Ishida S, Aasawat T, Sumita M, Katouda M, Yoshizawa T, Yoshizoe K, et al. ChemTSv2: functional molecular design using de novo molecule generator. WIREs Comput Mol Sci. 2023; 13: e1680.

Jimenez-Luna J, Grisoni F and Schneider G. Artificial intelligence in drug discovery: recent advances and future perspectives. Expert Opin Drug Discov. 2021; 16: 949-959.

Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, et al. Highly accurate protein structure prediction with AlphaFold. Nature. 2021; 596: 583-589.

Kaplan J, McCandlish S, Henighan T, Brown TB, Chess B, Child R, et al. Scaling laws for neural language models. arXiv:2001.08361. 2020.

Kipf TN and Welling M. Semi-supervised classification with graph convolutional networks. International Conference on Learning Representations, 2017.

Kojima R, Ishida S, Ohta M, Iwata H, Honma T and Okuno Y. kGCN: a graph-based deep learning framework for chemical structures. J Cheminform. 2020; 12: 32.

McMahan HB, Moore E, Ramage D, Hampson S and Aguera y Arcas B. Communication-efficient learning of deep networks from decentralized data. Proc 20th Int Conf Artif Intell Stat. 2017; 54: 1273-1282.

Nakazawa MA, Tamada Y, Tanaka Y, Ikeguchi M, Higashihara K and Okuno Y. Novel cancer subtyping method based on patient-specific gene regulatory network. Sci Rep. 2021; 11: 23653.

Nelson MR, Tipney H, Painter JL, Shen J, Nicoletti P, Shen Y, et al. The support of human genetic evidence for approved drug indications. Nat Genet. 2015; 47: 856-860.

Nobel Prize Outreach. The Nobel Prize in Chemistry 2024, 2024. Available from: https://www.nobelprize.org/prizes/chemistry/2024/summary/

Office of Pharmaceutical Industry Research, Japan Pharmaceutical Manufacturers Association. The Actual Condition of Pharmaceutical R&D: R&D duration, success probability, and R&D costs based on questionnaire surveys, 2024. Available from: https://en.jpma.or.jp/opir/research/rs_082/article_082.html

Okuno Y. Drug screening for COVID-19 using supercomputer "Fugaku". Folia Pharmacol Jpn. 2022; 157: 111-114.

Oldenhof M, van Leeuwen J, Nauta M, Fehr J, Friedrich L, Mervin L, et al. Industry-scale orchestrated federated learning for drug discovery.

Proc AAAI Conf Artif Intell. 2023; 37: 15576-15584.

RIKEN Center for Computational Science. Using simulations to search existing drugs for possible COVID-19 treatments. HPC- and AI-driven Drug Development Platform Division, 2022. Available from: https://fugaku100kei.jp/assets/data/e_hpcimag_01.pdf

Salo-Ahen OMH, Alanko I, Bhadane R, Bonvin AMJJ, Honorato RV, Hossain S, et al. Molecular dynamics simulations in drug discovery and pharmaceutical development. Processes. 2021; 9: 71.

Scannell JW, Blanckley A, Boldon H and Warrington B. Diagnosing the decline in pharmaceutical R&D efficiency. Nat Rev Drug Discov. 2012; 11: 191-200.

Schlander M, Hernandez-Villafuerte K, Cheng CY, Mestre-Ferrandiz J and Baumann M. How much does it cost to research and develop a new drug? A systematic review and assessment. Pharmacoeconomics. 2021; 39: 1243-1269.

Sheridan RP. Three useful dimensions for domain applicability in QSAR models using random forest. J Chem Inf Model. 2012; 52: 814-823.

Silver D, Huang A, Maddison CJ, Guez A, Sifre L, van den Driessche G, et al. Mastering the game of Go with deep neural networks and tree search. Nature. 2016; 529: 484-489.

Sun D, Gao W, Hu H and Zhou S. Why 90% of clinical drug development fails and how to improve it? Acta Pharm Sin B. 2022; 12: 3049-3062.

Tsubaki M, Tomii K and Sese J. Compound-protein interaction prediction with end-to-end learning of neural networks for graphs and sequences. Bioinformatics. 2019; 35: 309-318.

U.S. 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 and Other Interested Parties, 2025. Available from: https://www.fda.gov/regulatory-information/search-fda-guidance-documents/considerations-use-artificial-intelligence-support-regulatory-decision-making-drug-and-biological

Vamathevan J, Clark D, Czodrowski P, Dunham I, Ferran E, Lee G, et al. Applications of machine learning in drug discovery and development. Nat Rev Drug Discov. 2019; 18: 463-477.

Wei J, Tay Y, Bommasani R, Raffel C, Zoph B, Borgeaud S, et al. Emergent abilities of large language models. Trans Mach Learn Res. 2022.

Yang X, Zhang J, Yoshizoe K, Terayama K and Tsuda K. ChemTS: an efficient Python library for de novo molecular generation. Sci Technol Adv Mater. 2017; 18: 972-976.