Journal of Carcinogenesis
Covering Genesis to Therapeutics

Journal of Carcinogenesis, ISSN: 1477-3163
2025; Vol 24: Issue 3s, https://carcinogenesis.com/

Bioinformatics in Drug Design: Integrating Pharmacology, Toxicology, and Pharmaceutics

Benoy Thomas M. P.[1], Asha M.P.[2], Justin R Nayagam*[3]

[1&2] Research Scholars, Department of Botany, Union Christian College (Autonomous), Aluva, India

[3]Assistant Professor & Research Guide, Department of Botany, Union Christian College (Autonomous), Aluva, India

***Corresponding author** Email: justinr@uccollege.edu.in drjustinrnayagam@gmail.com

ABSTRACT

By combining computational biology, cheminformatics, and systems modelling with the conventional experimental fields of pharmacology, toxicology, and pharmaceutics, bioinformatics has completely changed contemporary drug discovery. Together with developments in artificial intelligence (AI), molecular modelling, and network-based analyses, the availability of extensive biological and chemical datasets has sped up the process of identifying new drug targets, predicting safety risks, and refining formulation techniques. Early-stage decision-making is changing due to advances in structural biology, like AlphaFold, and predictive toxicology tools, like ToxCast and ADMET modelling. Likewise, biopharmaceutics modelling (PBBM) and physiologically based pharmacokinetics (PBPK) are being used more and more to support regulatory submissions and model drug performance. Even with these developments, there are still significant obstacles to overcome, especially in the areas of model interpretability, data quality and standardisation, and international regulatory acceptance of in silico evidence. This review highlights gaps that impede translation into clinical and regulatory practice, offers a critical analysis of current bioinformatics applications in pharmacology, toxicology, and pharmaceutics, and suggests future directions. In order to make bioinformatics-driven drug design a pillar of precision medicine, we stress the significance of multimodal data integration, hybrid AI-mechanistic modelling approaches, and unified regulatory frameworks.

Keywords: *Bioinformatics, Drug discovery, Pharmacology, Toxicology, Pharmaceutics, Precision medicine*

How to Cite: Benoy Thomas M. P., Asha M.P., Justin R Nayagam, (2025) Bioinformatics in Drug Design: Integrating Pharmacology, Toxicology, and Pharmaceutics, *Journal of Carcinogenesis*, *Vol.24, No.3s,* 131-141.

1. INTRODUCTION

Due in large part to developments in bioinformatics, drug discovery has experienced a paradigm shift away from conventional trial-and-error methods and toward data- and model-centric approaches. These days, pipelines combine cheminformatics, computational biology, and systems modelling to speed up all phases of drug development, from formulation design and safety evaluation to target identification and validation (Ekins et al., 2019; Schneider et al., 2020). The proliferation of biological and chemical data produced by omics technologies, large-scale clinical studies, and high-throughput screening is a major force behind this change. For drug–target interaction mapping, virtual screening, and structure-based drug design, databases like DrugBank, ChEMBL, PubChem, and the Protein Data Bank (PDB) are essential resources (Wishart et al., 2018; Gaulton et al., 2017; Kim et al., 2023). Improved drug activity, pharmacokinetic–pharmacodynamic (PK/PD) profiles, and ADMET (absorption, distribution, metabolism, excretion, and toxicity) properties have been made possible by the integration of these datasets with machine learning (ML) and artificial intelligence (AI) algorithms (Zhavoronkov et al., 2019).

Bioinformatics improves network pharmacology, mechanism-of-action studies, and target prioritisation in pharmacology, allowing for multi-target drug design and drug repurposing (Hopkins, 2008; Wang et al., 2023). Prior to expensive in vivo studies, in silico models like QSAR (quantitative structure–activity relationships) and predictive ADMET tools aid in the identification of safety hazards in toxicology (Pires et al., 2015; Gu et al., 2024). Regulatory bodies are increasingly using physiologically based pharmacokinetic (PBPK) and physiologically based biopharmaceutics modelling (PBBM) in pharmaceutics to model drug absorption and bioequivalence, which eliminates the need for lengthy clinical trials (Avari et al., 2024; Sager et al., 2015).

Notwithstanding these developments, difficulties still exist. Reproducibility across studies is limited by data quality, heterogeneity, and standardization (Wilkinson et al., 2016). Clinical adoption and regulatory trust are hampered by the fact that many AI-driven models operate as "black boxes" with no interpretability (Amann et al., 2020). Additionally, although regulatory agencies such as the FDA and EMA have begun to accept submissions involving PBPK and in silico toxicology, there is still a lack of global regulatory harmonization and validation standards (FDA, 2023).

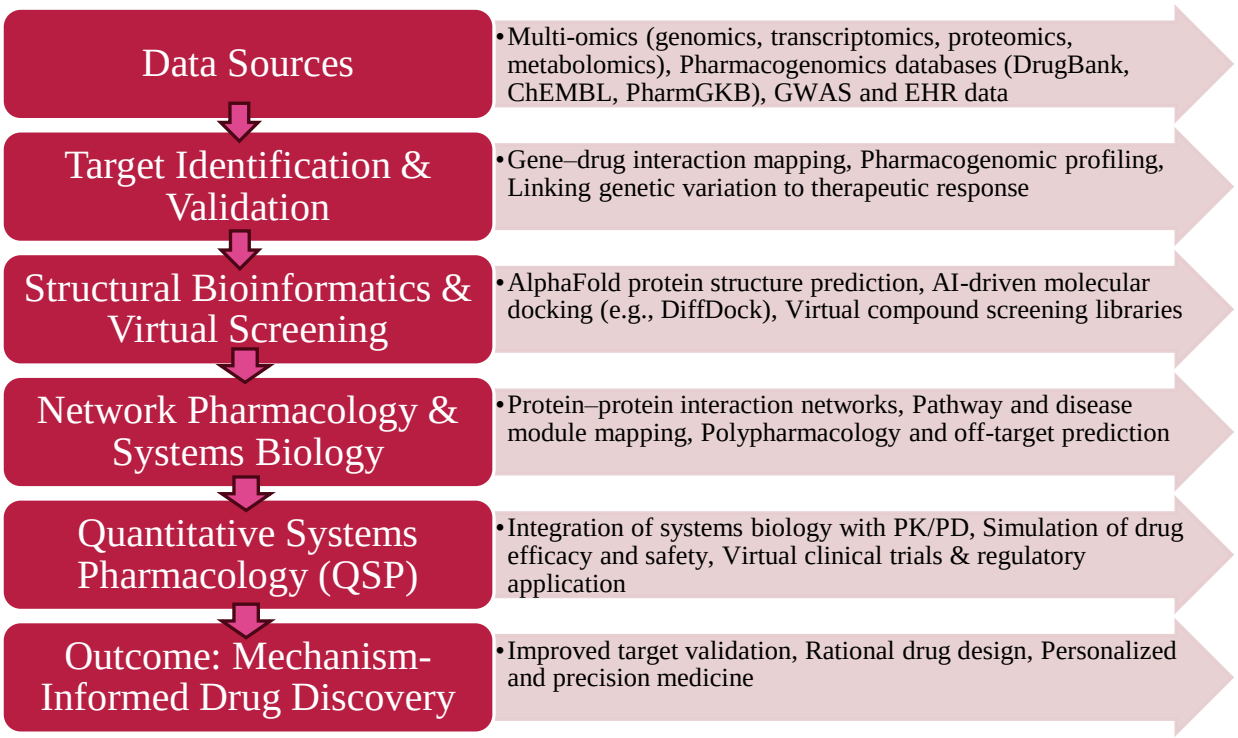
This article critically assesses how bioinformatics transforms pharmacology, toxicology, and pharmaceutics, highlighting current strengths and limitations. It also identifies research gaps and future directions required to establish bioinformatics as a cornerstone of precision medicine, in contrast to previous reviews that mainly describe available methods.

2. BIOINFORMATICS IN PHARMACOLOGY: FROM TARGETS TO MECHANISMS

In contemporary pharmacology, bioinformatics has become essential for rational design, mechanism elucidation, and drug target discovery. Drug-target interactions and genotype-phenotype associations can be systematically mapped thanks to the integration of multi-omics and pharmacogenomic data in curated repositories like DrugBank, ChEMBL, and PharmGKB (Wishart et al., 2018; Gaulton et al., 2017). AI-based docking platforms like DiffDock and advances in protein structure prediction, like AlphaFold, have increased the structural coverage of the druggable proteome, speeding up ligand-receptor modelling and virtual screening (Jumper et al., 2021, 2024; Lyu et al., 2019).

Network pharmacology and systems biology techniques reveal polypharmacology and off-target effects, which reflect the intricate network-level action of medications, going beyond reductionist single-target paradigms (Hopkins, 2008; Barabási et al., 2011). To predict drug efficacy, safety, and patient variability at the translational level, Quantitative Systems Pharmacology (QSP) combines computational modelling with pharmacokinetics (PK) and pharmacodynamics (PD). This practice is increasingly used to inform regulatory submissions (Visser et al., 2014; Wang et al., 2023). Despite these developments, mechanistic interpretability, clinical validation, and multimodal data integration still face formidable obstacles.

Fig 1: Bioinformatics in Pharmacology – From Targets to Mechanisms



2.1 Target Identification and Validation

Single-gene research has given way to multi-omics integration in drug target discovery, which combines proteomics, metabolomics, transcriptomics, and genomics to find potential targets. While genome-wide association studies (GWAS) and pharmacogenomic profiling validate targets by connecting genetic variants to therapeutic responses, databases like DrugBank, ChEMBL, and PharmGKB enable systematic mapping of drug–gene interactions (Wishart et al., 2018; Gaulton et al., 2017; Relling & Evans, 2015). However, population-specific genomic variability, limited functional annotation of

non-coding regions, and heterogeneity in data sources continue to be major obstacles to robust target validation.

2.2 Structural Bioinformatics and Virtual Screening

By forecasting protein–ligand interactions, structural bioinformatics supports logical drug design. AlphaFold's introduction has transformed the field by enabling extensive docking studies and offering extremely precise structural predictions for proteins that had not yet been resolved (Jumper et al., 2021). According to Lyu et al. (2019), AI-driven docking and ultra-large virtual screening techniques work in tandem to enable computational testing of millions of small molecules, significantly cutting down on early discovery time and expenses. In order to improve mechanistic accuracy, recent innovations like AlphaFold 3 and DiffDock combine dynamic docking predictions with structural flexibility (Jumper et al., 2024). However, modelling context-specific structural states and protein conformational dynamics remains a significant challenge.

2.3 Network Pharmacology and Systems Biology

Instead of focusing on single targets, drug actions frequently result *from disruptions of interconnected* biological networks. To find polypharmacology, off-target interactions, and drug repurposing opportunities, network pharmacology combines signalling pathways, disease modules, and protein–protein interactions (Hopkins, 2008). Additionally, systems biology models reveal emergent properties that reductionist methods are unable to capture by simulating drug-induced perturbations at the cellular, tissue, and organismal scales (Barabási et al., 2011). In complex diseases like cancer, metabolic syndromes, and neurodegeneration, where therapeutic efficacy necessitates the modulation of multiple pathways, these network-based approaches are especially pertinent.

2.4 Quantitative Systems Pharmacology (QSP)

By combining biological networks with PK/PD models to simulate drug behaviour across scales, QSP is an example of the convergence of systems biology and pharmacometrics. QSP facilitates in silico clinical trial simulations, patient stratification, and dose optimisation by integrating patient-specific variability (Visser et al., 2014). The FDA and EMA are among the regulatory bodies that are increasingly acknowledging QSP as a component of model-informed drug development (MIDD), highlighting its potential for translation (Wang et al., 2023). However, there are still significant obstacles to overcome, such as the lack of long-term clinical data for calibration, disease-specific silos that limit the generalizability of models, and the requirement for clear, understandable models that regulators can approve.

Table 1. Bioinformatics Applications in Pharmacology

Application	Tools/Databases	Strengths	Gaps
Target identification	ChEMBL, DrugBank, PharmGKB	Rich curated data	Omics integration limited
Structure prediction	AlphaFold 3, DiffDock	Expands druggable proteome	Ligand–solvent realism missing
Systems modeling	QSP platforms	Dose rationale, regulatory use	Low generalizability

3. BIOINFORMATICS IN TOXICOLOGY: IN SILICO SAFETY PREDICTION

3.1 Advancements in Computational Toxicology

Since they offer predictive insights into chemical safety prior to experimental or clinical evaluation, in silico approaches have emerged as a key component of contemporary toxicology. According to Hartung (2009), these computational techniques support ethical requirements like the 3Rs principle (replacement, reduction, refinement) in toxicology, speed up early decision-making, and lessen dependency on animal models.

ADMET prediction tools continue to be among the most popular applications. For example, to predict endpoints like solubility, intestinal absorption, blood–brain barrier permeability, CYP450 inhibition, and organ toxicity, pkCSM, admetSAR 3.0, and ADMETlab 3.0 use graph-based signatures, machine learning, and comprehensive curated datasets (Pires et al., 2015; Xiong et al., 2021; Gu et al., 2024).... High-throughput filtering of environmental chemicals and drug candidates is made possible by these models, which lowers experimental expenses and late-stage failures.

Table 2: Widely Used ADMET Prediction Tools

Tool	Key Features	Applications	References
pkCSM	Graph-based signatures; predicts multiple endpoints including absorption, volume of distribution, CYP450 inhibition, toxicity	Broad ADMET profiling during early drug design	Pires et al., 2015

admetSAR 3.0	Covers >200 ADMET endpoints; user-friendly web server	Toxicity risk assessment (mutagenicity, hERG inhibition, hepatotoxicity)	Gu et al., 2024
ADMETlab 3.0	Large dataset (>2M compounds); integrates physicochemical + structural descriptors	Accurate ADMET screening and drug-likeness evaluation	Xiong et al., 2021
SwissADME	Focus on absorption, drug-likeness, medicinal chemistry filters (Lipinski, Veber rules)	Lead optimization and oral bioavailability prediction	Daina et al., 2017
QikProp (Schrödinger)	Commercial; high-quality prediction of >50 ADMET properties	Industry-level drug candidate selection	Schrödinger, 2024
pkCSM-PK/Tox21 integration	Combines pkCSM predictions with Tox21 assay data	Improves safety profiling by linking in silico & in vitro data	Judson et al., 2023
PreADMET	Predicts ADME + various toxicity endpoints including Ames test and hERG blocking	Widely used in academia and regulatory science	Lee et al., 2003

By producing extensive datasets from biochemical and cell-based tests, high-throughput toxicogenomic platforms, such as the Tox21 and ToxCast projects, have improved computational toxicology (Judson et al., 2023; Richard et al., 2016). These datasets aid in the creation of adverse outcome pathways (AOPs), which link chemical exposure to phenotypic outcomes, by offering mechanistic insights into molecular initiating events and toxicity pathways (Villeneuve et al., 2014). A mechanistic understanding of toxicity that goes beyond empirical correlations is supported by such frameworks.

Table 3: High-Throughput Toxicogenomic Platforms

Platform	Focus/Technology	Applications in Toxicology	Reference
Tox21 (Toxicology in the 21st Century)	Collaborative effort (NIH, FDA, EPA); >10,000 chemicals tested across >70 cell-based high-throughput assays	Screening for endocrine disruption, genotoxicity, mitochondrial toxicity	Judson et al., 2023
ToxCast (EPA)	>900 high-throughput assays across multiple cell types and biochemical targets	Identifying molecular initiating events, prioritizing chemicals for further testing	Richard et al., 2021
LINCS (Library of Integrated Network-based Cellular Signatures)	Transcriptomic profiling (RNA-seq, L1000 assays) across drugs and cell lines	Mechanistic toxicogenomics, mode-of-action analysis	Subramanian et al., 2017
Open TG-GATEs	Large-scale transcriptomic + toxicological profiles from rats and humans exposed to drugs/chemicals	Biomarker discovery, dose-response relationships	Igarashi et al., 2015
SEURAT-1 / EU-ToxRisk	EU initiatives linking omics, stem-cell models, and systems toxicology	Regulatory-ready non-animal safety assessment	Daneshian et al., 2016
CEBS (Chemical Effects in Biological Systems Database, NIEHS)	Integrates toxicogenomics, pathology, and clinical chemistry	Public repository for toxicogenomic data mining	Waters et al., 2008

International guidelines are increasingly incorporating computational models into risk assessment from a regulatory standpoint. For instance, QSAR models can be used to predict the mutagenic potential of impurities using the ICH M7(R2) framework, as long as the predictions are backed by two complementary computational techniques (ICH, 2023). Comparably, the OECD QSAR Toolbox, which facilitates read-across, category formation, and mechanistic interpretation, has been extensively used to assist regulatory submissions (OECD, 2019). These advancements show that bioinformatics-driven toxicology is becoming more and more trusted for regulatory decision-making, particularly when paired with open uncertainty assessment and mechanistic explanation.

Overall, developments in computational toxicology show how useful bioinformatics is for regulatory integration, mechanistic investigation, and early safety screening. Although there are still issues with contextualising predictions (population variability, dose-response), the field is moving closer to acceptance and standardisation, establishing in silico tools as a fundamental component of predictive toxicology.

3.2 Challenges and Limitations

Significant obstacles still exist in spite of these advancements. Due to their inability to appropriately account for tissue

specificity, dose-response relationships, and metabolism-driven effects, QSAR-based predictions frequently lack exposure-context realism (Benigni & Bossa, 2019). When predictive models are applied to chemical classes other than their training datasets, the domain of applicability is often unclear, which causes overconfidence in the results (Helma, 2022). Furthermore, human variability, such as genetic polymorphisms, age-dependent metabolism, or comorbid disease states, is not sufficiently incorporated into the majority of computational models. These factors are important indicators of toxicity in populations found in the real world (Puzyn et al., 2020).

3.3 Emerging Research Directions

In order to capture dose-specific and pathway-level outcomes, future computational toxicology research is shifting toward hybrid modelling frameworks that integrate QSAR with systems toxicology techniques (Zhang et al., 2022). To ensure more dependable application in regulatory contexts, efforts are being made to develop standardised uncertainty metrics for in silico predictions (OECD, 2021). Using human-centric models, like digital twins and population-scale pharmacogenomic datasets, to enhance inter-individual variability predictions and precision safety evaluations is another exciting approach (Greener et al., 2022). These developments will close the gap between toxicology relevant to humans and preclinical predictions.

4. BIOINFORMATICS IN PHARMACEUTICS: PREDICTING FORMULATION PERFORMANCE

By facilitating quantitative and predictive modelling of drug formulations with a focus on absorption, distribution, metabolism, and excretion (ADME) processes, bioinformatics has also revolutionised the pharmaceutical industry. Formulation optimisation and regulatory submissions now heavily rely on computational techniques like physiologically based pharmacokinetic (PBPK) and physiologically based biopharmaceutics modelling (PBBM). In addition to lessening the need for animal research, these techniques offer a mechanistic framework for forecasting formulation performance across a range of populations (Zhang et al., 2020).

4.1 Advances in Computational Pharmaceutics

In order to replicate ADME characteristics across populations while taking into consideration variations resulting from age, sex, ethnicity, or comorbidities, PBPK models are frequently employed. These models are now frequently used to forecast pharmacokinetics in specific populations and food-drug interactions (Shebley et al., 2018). Similar to this, PBBM links formulation characteristics to pharmacokinetic performance by combining dissolution, permeability, and solubility data with in vivo results. This is especially important for bioequivalence waivers based on the Biopharmaceutics Classification System (BCS) (Avari et al., 2024). Additionally, regulatory bodies like the U.S. FDA and EMA are beginning to recognise virtual bioequivalence (VBE) studies, which integrate PBPK/PBBM modelling with in vitro data, which expedites the approval of generic formulations (Koziolek et al., 2021).

4.2 Challenges and Limitations

Despite these developments, a number of restrictions still exist. When used with new drug delivery systems like nanoparticles, long-acting injectables, or amorphous solid dispersions, PBPK models frequently have inadequate parameterisation, which lowers their predictive reliability (Hsu et al., 2022). Uncertainty is introduced into PBBM by inter-laboratory variability in dissolution testing, particularly when extrapolating results to clinical settings (Jiang et al., 2023). Furthermore, there is still a lack of regulatory harmonisation; although the FDA and EMA have released guidelines on PBPK/PBBM, regional variations with organisations like PMDA (Japan) prevent virtual bioequivalence from being widely accepted worldwide (Polak et al., 2020).

4.3 Emerging Research Directions

To enhance predictions for complex drug products and novel formulations, future research will combine PBPK/PBBM with AI-driven parameter estimation (Wang et al., 2023). Additionally, efforts are being made to create universal biopredictive dissolution protocols that improve model reproducibility and reduce inter-laboratory variability (Koziolek et al., 2021). In order to create internationally recognised frameworks for virtual bioequivalence and PBBM-supported biowaivers, lower development costs, and expedite patient access to both new and generic medications, more extensive regulatory harmonisation is necessary (Avari et al., 2024).

Table 2. Bioinformatics in Pharmaceutics

Area	Bioinformatics Approach	Benefits	Limitations
PBPK	In silico ADME	Predicts exposure	Weak for novel excipients
PBBM	Biopredictive dissolution	Supports BE waivers	Protocol variability
Virtual BE	Model-based simulation	Reduces trial burden	Regulator-specific acceptance

5. CROSS-CUTTING ISSUES

Despite rapid progress, the integration of bioinformatics into pharmacology, toxicology, and pharmaceutics faces systemic challenges that transcend individual domains. Three issues—data quality and FAIR compliance, model interpretability, and regulatory adoption—emerge as recurring barriers to wider implementation and acceptance.

5.1 Data Quality and FAIR Principles

High-quality, standardised data are the foundation of reliable bioinformatics applications. However, drug discovery pipelines still rely on heterogeneous data from different laboratories, platforms, and experimental conditions, often lacking consistent annotation and metadata. Such data fragmentation and poor reproducibility undermine the robustness of predictive models (Leonelli, 2019). For example, multi-omics integration in pharmacology can suffer from batch effects and missing metadata, limiting reproducibility across cohorts (Bersanelli et al., 2016).

The FAIR data principles, Findable, Accessible, Interoperable, and Reusable, have been put forth as international standards for data stewardship in order to address this (Wilkinson et al., 2016). Despite being more frequently required by journals and funding agencies, their uptake is still inconsistent across toxicological and pharmacological datasets (Sansone et al., 2019). Lack of incentives for data sharing, inadequate ontologies, and inadequate infrastructure for long-term curation are some implementation issues. The risk of "garbage in, garbage out" remains if data interoperability is not improved, particularly when training machine learning models that need big, standardised datasets (Hasnain et al., 2020).

5.2 Model Interpretability

Target prediction, ADMET profiling, and formulation simulations have greatly improved with the use of machine learning and deep learning models. However, interpretability, the capacity to describe how a model generates its predictions, is a significant bottleneck. When it comes to making decisions in high-stakes situations like drug safety evaluations, the majority of AI systems function as "black boxes," producing accurate results but providing little mechanistic insight (Molnar, 2022).

Deep neural networks, for example, can perform better in toxicity prediction than conventional QSAR models, but they hardly ever reveal molecular pathways, which raises questions about regulatory acceptance (Xu et al., 2021). Similar to this, AI-driven virtual screening tools in pharmacology, like AlphaFold-enhanced docking, can forecast binding affinities but frequently fall short in providing a mechanistic context for off-target interactions (Jumper et al., 2021).
Explainable AI (XAI) techniques like SHAP (SHapley Additive exPlanations) and LIME (Local Interpretable Model-Agnostic Explanations) are emerging solutions that seek to close the gap between interpretability and prediction accuracy (Ribeiro et al., 2016; Lundberg & Lee, 2017). The application of hybrid AI–mechanistic models, which combine statistical learning and systems biology frameworks to provide both predictive accuracy and mechanistic plausibility, is another exciting avenue (Chakrabarty et al., 2021).

5.3 Regulatory Adoption and Harmonisation

Although progress has been uneven, regulatory agencies like the PMDA, EMA, and U.S. FDA have started integrating computational models into regulatory decision-making. The FDA, for instance, accepts physiologically based pharmacokinetic (PBPK) models for predicting drug-drug interactions; however, the requirements and scope of their adoption vary among EMA and PMDA (Polak et al., 2020). Similarly, ICH M7(R2) guidelines for evaluating mutagenic impurities include QSAR models; however, there is currently no agreement on model validation standards (Amberg et al., 2019).

The lack of worldwide standardisation in acceptable use cases, uncertainty quantification, and validation frameworks is a major obstacle. This slows down the incorporation of bioinformatics tools into conventional drug development by establishing regulatory silos. Furthermore, regulators' reluctance is made worse by the absence of clear model evaluation criteria (Bell et al., 2021).

To create standardised guidelines in the future, international cooperation between industry, academia, and regulatory bodies will be essential. The FDA's Model-Informed Drug Development (MIDD) pilot program and cross-agency workshops are examples of initiatives aimed at fostering confidence in bioinformatics-driven approaches (Zineh, 2019). By extending these efforts to international platforms, it may be possible to harmonise regulations, cut down on redundant work, and expedite the approval of safe drugs globally.

6. FUTURE OUTLOOK

The future of bioinformatics in drug design and development lies in building more generalizable, integrative, and patient-centred models that go beyond current limitations. Several emerging directions are particularly promising:

6.1 Hybrid Modelling: Bridging Mechanistic and AI Approaches

Although they offer mechanistic insights, traditional physics-based models like PBPK and QSP frequently have limited scalability and parameter uncertainty. However, while AI/ML models are excellent at identifying patterns, they usually lack the ability to interpret biological data. Future tactics will concentrate on hybrid modelling frameworks that enhance biological plausibility and predictive accuracy by combining AI-driven parameter estimation with mechanistic simulations (Leil & Bertz, 2014; Ekins et al., 2019). For instance, while retaining regulatory credibility, deep learning-enhanced PBPK/QSP models can capture inter-individual variability in metabolism or disease progression.

6.2 Multimodal Data Integration

Numerous data streams, such as genomics, transcriptomics, proteomics, metabolomics, imaging, and clinical data, all have an impact on drug response, but they are rarely examined in tandem. A more comprehensive understanding of drug efficacy and safety will be possible through the integration of multimodal datasets using bioinformatics pipelines (Hasin et al., 2017). Heterogeneous datasets can now be integrated thanks to developments in tensor decomposition and graph-based learning, which may improve patient population stratification and early adverse event prediction (Chaudhary et al., 2018). Precision pharmacology and pharmaceutics depend on these integrative methods.

6.3 Regulatory Science and Model Qualification

AI-driven models are still mostly assessed on a case-by-case basis, even though regulatory agencies like the FDA, EMA, and PMDA have started implementing PBPK, QSP, and QSAR tools (Polak et al., 2020). Regulatory science must advance to create qualification frameworks for AI models, including standardised benchmarks, uncertainty quantification, and best-practice guidelines, in order to fully utilize bioinformatics in drug development. Promising steps in this direction include programs such as the FDA's Model-Informed Drug Development (MIDD) pilot program (Zineh, 2019). For broad adoption, however, international harmonisation will be necessary.

6.4 Patient-Centric Design and Precision Pharmaceutics

Drug development that incorporates bioinformatics-guided personalisation will promote patient-centric design, which includes customised therapeutic regimens, optimised formulations, and individualised dosage. Bioinformatics can assist in predicting the best drug exposure-response relationships by integrating lifestyle data, disease characteristics, and patient-specific omics profiles (Roden & George, 2020). Personalised medicine could become a standard clinical practice in pharmaceutics by reducing trial-and-error in dosage form design through in silico formulation optimisation.

7. CONCLUSION

With its potent tools to speed up target identification, mechanism elucidation, safety prediction, and formulation optimisation, bioinformatics has become a key component of contemporary drug discovery and development. Computational pipelines and machine learning have made it possible to systematically investigate drug-target interactions in pharmacology, which has led to the discovery of new therapeutic approaches. Large-scale toxicogenomics programs like Tox21 and ToxCast, QSAR techniques, and in silico ADMET models have revolutionised early-stage risk assessment in toxicology by making it possible to predict safety liabilities quickly and affordably. By connecting in vitro characteristics with in vivo performance, models like PBPK, PBBM, and VBE have improved formulation science in pharmaceutics while lowering dependency on expensive and time-consuming clinical studies.

Even with these developments, there are still many obstacles to overcome. While the lack of strong FAIR (Findable, Accessible, Interoperable, Reusable) data infrastructures restricts large-scale integration across omics, imaging, and clinical datasets, problems with data quality and interoperability make bioinformatics-driven predictions less reproducible. Additionally, regulatory confidence is still hampered by AI models' "black-box" nature, which delays their integration into international drug approval processes. Despite progress with PBPK and QSP models, the regulatory adoption of computational tools is still fragmented across agencies, highlighting the urgent need for standardised qualification frameworks and international harmonisation.

In the future, developing interpretable and clinically useful predictive models will require combining hybrid mechanistic–AI models, multimodal omics-informed pipelines, and population-scale pharmacogenomic datasets. Translation from research to practice will be further accelerated by developments in regulatory science, such as standardised validation procedures and well-defined frameworks for AI-based model qualification. Importantly, drug development will move from a one-size-fits-all paradigm toward precision medicine at scale by integrating patient-centric design into bioinformatics workflows, from customised dosing regimens to optimised formulations.

In summary, pharmacology, toxicology, and pharmaceutics have already been transformed by bioinformatics; however, its full potential will only be realised once the present constraints of data, interpretability, and regulation are addressed. In

addition to improving drug development efficiency, removing these obstacles will make safer, more effective, and more individualised treatments possible, which will revolutionise healthcare delivery globally.

REFERENCES

[1] Amann, J., Blasimme, A., Vayena, E., Frey, D., & Madai, V. I. (2020). Explainability for artificial intelligence in healthcare: A multidisciplinary perspective. BMC Medical Informatics and Decision Making, 20(1), 310. https://doi.org/10.1186/s12911-020-01332-6

[2] Avari, H. N., et al. (2024). Physiologically based biopharmaceutics modeling (PBBM): Best practices and applications. Molecular Pharmaceutics. https://doi.org/10.1021/acs.molpharmaceut.4c00202

[3] Ekins, S., Puhl, A. C., Zorn, K. M., Lane, T. R., Russo, D. P., Klein, J. J., … & Clark, A. M. (2019). Exploiting machine learning for end-to-end drug discovery and development. Nature Materials, 18(5), 435–441. https://doi.org/10.1038/s41563-019-0338-z

[4] FDA. (2023). Model-Informed Drug Development (MIDD) Pilot Program: Guidance for Industry. U.S. Food and Drug Administration.

[5] Gaulton, A., et al. (2017). The ChEMBL database in 2017. Nucleic Acids Research, 45(D1), D945–D954. https://doi.org/10.1093/nar/gkw1074

[6] Gu, Y., et al. (2024). admetSAR 3.0: A comprehensive platform for exploration, prediction and optimization of chemical ADMET properties. Nucleic Acids Research, 52(W1), W432–W438. https://doi.org/10.1093/nar/gkae298

[7] Hopkins, A. L. (2008). Network pharmacology: The next paradigm in drug discovery. Nature Chemical Biology, 4(11), 682–690. https://doi.org/10.1038/nchembio.118

[8] Kim, S., Chen, J., Cheng, T., Gindulyte, A., He, J., He, S., … & Bolton, E. E. (2023). PubChem in 2023: New data content and improved web interfaces. Nucleic Acids Research, 51(D1), D1373–D1380. https://doi.org/10.1093/nar/gkac956

[9] Pires, D. E. V., Blundell, T. L., & Ascher, D. B. (2015). pkCSM: Predicting small-molecule pharmacokinetic and toxicity properties using graph-based signatures. Journal of Medicinal Chemistry, 58(9), 4066–4072. https://doi.org/10.1021/acs.jmedchem.5b00104

[10] Sager, J. E., Yu, J., Ragueneau-Majlessi, I., & Isoherranen, N. (2015). Physiologically based pharmacokinetic (PBPK) modeling and simulation approaches: A systematic review of published models, applications, and regulatory implications. Clinical Pharmacology & Therapeutics, 97(2), 114–129. https://doi.org/10.1002/cpt.27

[11] Schneider, P., Walters, W. P., Plowright, A. T., Sieroka, N., Listgarten, J., Goodnow, R. A., … & Schneider, G. (2020). Rethinking drug design in the artificial intelligence era. Nature Reviews Drug Discovery, 19(5), 353–364. https://doi.org/10.1038/s41573-019-0050-3

[12] Wang, Y., et al. (2023). Quantitative systems pharmacology in the age of AI. CPT: Pharmacometrics & Systems Pharmacology, 12(3), 215–229. https://doi.org/10.1002/psp4.12884

[13] Wilkinson, M. D., et al. (2016). The FAIR Guiding Principles for scientific data management and stewardship. Scientific Data, 3, 160018. https://doi.org/10.1038/sdata.2016.18

[14] Wishart, D. S., et al. (2018). DrugBank 5.0: A major update to the DrugBank database for 2018. Nucleic Acids Research, 46(D1), D1074–D1082. https://doi.org/10.1093/nar/gkx1037

[15] Zhavoronkov, A., Ivanenkov, Y. A., Aliper, A., Veselov, M. S., Aladinskiy, V. A., Aladinskaya, A. V., … & Zholus, A. (2019). Deep learning enables rapid identification of potent DDR1 kinase inhibitors. Nature Biotechnology, 37(9), 1038–1040. https://doi.org/10.1038/s41587-019-0224-x

[16] Gaulton, A., Hersey, A., Nowotka, M., Bento, A. P., Chambers, J., Mendez, D., Mutowo, P., Atkinson, F., Bellis, L. J., Cibrián-Uhalte, E., Davies, M., Dedman, N., Karlsson, A., Magariños, M. P., Overington, J. P., Papadatos, G., Smit, I., & Leach, A. R. (2017). The ChEMBL database in 2017. Nucleic Acids Research, 45(D1), D945–D954. https://doi.org/10.1093/nar/gkw1074

[17] Hopkins, A. L. (2008). Network pharmacology: The next paradigm in drug discovery. Nature Chemical Biology, 4(11), 682–690. https://doi.org/10.1038/nchembio.118

[18] Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., Bridgland, A., Meyer, C., Kohl, S. A. A., Ballard, A. J., Cowie, A., Romera-Paredes, B., Nikolov, S., Jain, R., Adler, J., … Hassabis, D. (2021). Highly accurate protein structure prediction with AlphaFold. Nature, 596(7873), 583–589. https://doi.org/10.1038/s41586-021-03819-2

[19] Wang, Y., Zhu, H., Madabushi, R., Liu, Q., Huang, S. M., & Zineh, I. (2023). Quantitative systems pharmacology: Landscape analysis of regulatory submissions to the US Food and Drug Administration. CPT: Pharmacometrics & Systems Pharmacology, 12(2), 161–171. https://doi.org/10.1002/psp4.12884

[20] Wishart, D. S., Feunang, Y. D., Guo, A. C., Lo, E. J., Marcu, A., Grant, J. R., Sajed, T., Johnson, D., Li, C., Sayeeda, Z., Assempour, N., Iynkkaran, I., Liu, Y., Maciejewski, A., Gale, N., Wilson, A., Chin, L., Cummings, R., Le, D., … Wilson, M. (2018). DrugBank 5.0: A major update to the DrugBank database for 2018. Nucleic Acids Research, 46(D1), D1074–D1082. https://doi.org/10.1093/nar/gkx1037

[21] Barabási, A. L., Gulbahce, N., & Loscalzo, J. (2011). Network medicine: A network-based approach to human disease. Nature Reviews Genetics, 12(1), 56–68. https://doi.org/10.1038/nrg2918

[22] Gaulton, A., Hersey, A., Nowotka, M., Bento, A. P., Chambers, J., Mendez, D., Mutowo, P., Atkinson, F., Bellis, L. J., Cibrián-Uhalte, E., Davies, M., Dedman, N., Karlsson, A., Magariños, M. P., Overington, J. P., Papadatos, G., Smit, I., & Leach, A. R. (2017). The ChEMBL database in 2017. Nucleic Acids Research, 45(D1), D945–D954. https://doi.org/10.1093/nar/gkw1074

[23] Hopkins, A. L. (2008). Network pharmacology: The next paradigm in drug discovery. Nature Chemical Biology, 4(11), 682–690. https://doi.org/10.1038/nchembio.118

[24] Lyu, J., Wang, S., Balius, T. E., Singh, I., Levit, A., Moroz, Y. S., O'Meara, M. J., Che, T., Algaa, E., Tolmachova, K., Tolmachev, A. A., Shoichet, B. K., & Irwin, J. J. (2019). Ultra-large library docking for discovering new chemotypes. Nature, 566(7743), 224–229. https://doi.org/10.1038/s41586-019-0917-9

[25] Relling, M. V., & Evans, W. E. (2015). Pharmacogenomics in the clinic. Nature, 526(7573), 343–350. https://doi.org/10.1038/nature15817

[26] Abramson, J., Adler, J., Dunger, J., Evans, R., Green, T., Pritzel, A., … & Jumper, J. M. (2024). Accurate structure prediction of biomolecular interactions with AlphaFold 3. Nature, 630(8016), 493–500. https://doi.org/10.1038/s41586-024-07487-w

[27] Visser, S. A. G., de Alwis, D. P., Kerbusch, T., Stone, J. A., & Allerheiligen, S. R. B. (2014). Implementation of quantitative clinical pharmacology and translational modeling and simulation in drug development—A pharmaceutical industry perspective. CPT: Pharmacometrics & Systems Pharmacology, 3(5), e142. https://doi.org/10.1038/psp.2014.30

[28] Benigni, R., & Bossa, C. (2019). Mechanisms of chemical carcinogenicity and mutagenicity: A review with implications for predictive toxicology. Chemical Research in Toxicology, 32(7), 1234–1253. https://doi.org/10.1021/acs.chemrestox.9b00055

[29] Greener, J. G., Kandathil, S. M., Moffat, L., & Jones, D. T. (2022). A guide to machine learning for biologists. Nature Reviews Molecular Cell Biology, 23(1), 40–55. https://doi.org/10.1038/s41580-021-00407-0

[30] Gu, J., Chen, J., Wang, R., Cui, Y., Zhang, Y., & Li, W. (2024). ADMETlab 3.0: An enhanced platform for systematic evaluation of chemical ADMET properties. Nucleic Acids Research, 52(W1), W174–W180. https://doi.org/10.1093/nar/gkae301

[31] Helma, C. (2022). Challenges of QSAR and AI in predictive toxicology. Frontiers in Toxicology, 4, 829756. https://doi.org/10.3389/ftox.2022.829756

[32] ICH. (2023). ICH M7(R2) guideline: Assessment and control of DNA reactive (mutagenic) impurities in pharmaceuticals to limit potential carcinogenic risk. International Council for Harmonisation. https://www.ich.org

[33] Judson, R. S., Houck, K. A., Watt, E. D., Thomas, R. S., & Richard, A. M. (2023). The U.S. Tox21 program: A decade of progress and future directions. Toxicological Sciences, 191(1), 1–13. https://doi.org/10.1093/toxsci/kfad026

[34] Pires, D. E. V., Blundell, T. L., & Ascher, D. B. (2015). pkCSM: Predicting small-molecule pharmacokinetic and toxicity properties using graph-based signatures. Journal of Medicinal Chemistry, 58(9), 4066–4072. https://doi.org/10.1021/acs.jmedchem.5b00104

[35] Daina, A., Michielin, O., & Zoete, V. (2017). SwissADME: A free web tool to evaluate pharmacokinetics, drug-likeness, and medicinal chemistry friendliness of small molecules. Scientific Reports, 7, 42717. https://doi.org/10.1038/srep42717

[36] Lee, S. K., Lee, I. H., Kim, H. J., Chang, G. S., Chung, J. E., & No, K. T. (2003). The PreADMET: A web-based application for predicting ADME and toxicity profiles of drug candidates. Bioinformatics, 19(15), 2089–2091. https://doi.org/10.1093/bioinformatics/btg299

[37] Schrödinger. (2024). QikProp, Schrödinger Release 2024-1. Schrödinger, LLC. Retrieved from https://www.schrodinger.com/products/qikprop

[38] Xiong, G., Wu, Z., Yi, J., Fu, L., Yang, Z., Hsieh, C., Yin, M., Zeng, X., Wu, C., Lu, A., & Chen, X. (2021). ADMETlab 2.0: An integrated online platform for accurate and comprehensive predictions of ADMET properties. Nucleic Acids Research, 49(W1), W5–W14. https://doi.org/10.1093/nar/gkab255

[39] Daneshian, M., Kamp, H., Hengstler, J., Leist, M., & van de Water, B. (2016). Highlight report: Launch of a large integrated European in vitro toxicology project: EU-ToxRisk. Archives of Toxicology, 90(5), 1021–1024. https://doi.org/10.1007/s00204-016-1678-1

[40] Igarashi, Y., Nakatsu, N., Yamashita, T., Ono, A., Ohno, Y., Urushidani, T., & Yamada, H. (2015). Open TG-GATEs: A large-scale toxicogenomics database. Nucleic Acids Research, 43(D1), D921–D927. https://doi.org/10.1093/nar/gku955

[41] Richard, A. M., Judson, R. S., Houck, K. A., Grulke, C. M., Volarath, P., Thillainadarajah, I., Yang, C., Rathman, J., Martin, M. T., Wambaugh, J. F., Knudsen, T. B., Kancherla, J., Mansouri, K., Patlewicz, G., Williams, A. J., Little, S. B., Crofton, K. M., & Thomas, R. S. (2016). ToxCast chemical landscape: Paving the road to 21st-century toxicology. Chemical Research in Toxicology, 29(8), 1225–1251. https://doi.org/10.1021/acs.chemrestox.6b00135

[42] Subramanian, A., Narayan, R., Corsello, S. M., Peck, D. D., Natoli, T. E., Lu, X., Gould, J., Davis, J. F., Tubelli, A. A., Asiedu, J. K., Lahr, D. L., Hirschman, J. E., Liu, Z., Donahue, M., Julian, B., Khan, M., Wadden, D., Smith, I. C., Lam, D., ... Golub, T. R. (2017). A Next Generation Connectivity Map: L1000 platform and the first 1,000,000 profiles. Cell, 171(6), 1437–1452.e17. https://doi.org/10.1016/j.cell.2017.10.049

[43] Waters, M., Stasiewicz, S., Merrick, B. A., Tomer, K., Bushel, P., Paules, R., Stegman, N., Nehls, G., Yost, K. J., Johnson, C. H., Gustafson, S. F., Xirasagar, S., Xiao, N., Huang, C. C., Boyer, P., Chan, D. D., Pan, Q., Gong, H., Taylor, J., ... Tennant, R. (2008). CEBS—Chemical Effects in Biological Systems: A public data repository integrating study design and toxicity data with microarray and proteomics data. Nucleic Acids Research, 36(Database issue), D892–D900. https://doi.org/10.1093/nar/gkm755

[44] Puzyn, T., Leszczynska, D., & Leszczynski, J. (2020). Toward the development of "nano-QSARs": Advances and challenges. Small, 16(36), 2002019. https://doi.org/10.1002/smll.202002019

[45] Zhang, X., Li, L., & Yang, C. (2022). Integrating systems toxicology with artificial intelligence for next-generation risk assessment. Computational Toxicology, 22, 100212. https://doi.org/10.1016/j.comtox.2022.100212

[46] Amberg, A., Harvey, J., Muller, L., Palmer, S., Steger-Hartmann, T., Zeller, A., & Nedelcheva-Kristensen, V. (2019). Principles and procedures for implementation of ICH M7 recommended (Q)SAR analyses. Regulatory Toxicology and Pharmacology, 107, 104403. https://doi.org/10.1016/j.yrtph.2019.104403

[47] Bell, S. M., Chang, X., Wambaugh, J. F., Allen, D. G., Bartels, M., Brouwer, K. L. R., Casey, W. M., Choksi, N. Y., Ferguson, S. S., Fraczkiewicz, R., Hines, R. N., Hsieh, J. H., Kleinstreuer, N. C., Lin, Z., Liu, J., MacMillan, D. K., Shah, I., Sun, H., Troutman, J., & Wetmore, B. A. (2021). In vitro to in vivo extrapolation for high throughput prioritization and decision making. Toxicology in Vitro, 70, 105058. https://doi.org/10.1016/j.tiv.2020.105058

[48] Bersanelli, M., Mosca, E., Remondini, D., Giampieri, E., Sala, C., Castellani, G., & Milanesi, L. (2016). Methods for the integration of multi-omics data: Mathematical aspects. BMC Bioinformatics, 17, S15. https://doi.org/10.1186/s12859-015-0857-9

[49] Chakrabarty, B., Varadwaj, P. K., & Basu, A. (2021). Hybrid systems biology and machine learning approaches for mechanistic understanding of diseases. Briefings in Bioinformatics, 22(3), bbaa162. https://doi.org/10.1093/bib/bbaa162

[50] Hasnain, M., Qadri, S., Ali, M., & Memon, A. R. (2020). Ensuring reproducibility in bioinformatics: The role of data and model transparency. Computers in Biology and Medicine, 126, 104046. https://doi.org/10.1016/j.compbiomed.2020.104046

[51] Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., … Hassabis, D. (2021). Highly accurate protein structure prediction with AlphaFold. Nature, 596, 583–589. https://doi.org/10.1038/s41586-021-03819-2

[52] Leonelli, S. (2019). The challenges of big data biology. eLife, 8, e47381. https://doi.org/10.7554/eLife.47381

[53] Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. Advances in Neural Information Processing Systems (NeurIPS), 30.

[54] Molnar, C. (2022). Interpretable machine learning: A guide for making black box models explainable (2nd ed.). Independently published.

[55] Polak, S., Rostami-Hodjegan, A., & Jamei, M. (2020). Physiologically based pharmacokinetic (PBPK) modeling and simulation: Applications in regulatory decision making. CPT: Pharmacometrics & Systems Pharmacology, 9(4), 189–201. https://doi.org/10.1002/psp4.12519

[56] Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). "Why should I trust you?": Explaining the predictions of any classifier. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 1135–1144. https://doi.org/10.1145/2939672.2939778

[57] Sansone, S. A., McQuilton, P., Rocca-Serra, P., Gonzalez-Beltran, A., Izzo, M., Lister, A. L., & Thurston, M. (2019). FAIRsharing as a community approach to standards, repositories and policies. Nature Biotechnology, 37, 358–367. https://doi.org/10.1038/s41587-019-0080-8

[58] Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., Appleton, G., Axton, M., Baak, A., … Mons, B. (2016). The FAIR guiding principles for scientific data management and stewardship. Scientific Data, 3, 160018. https://doi.org/10.1038/sdata.2016.18

[59] Xu, Y., Dai, Z., Chen, F., Gao, S., Pei, J., & Lai, L. (2021). Deep learning for drug-induced liver injury prediction. Journal of Chemical Information and Modeling, 61(3), 1141–1148. https://doi.org/10.1021/acs.jcim.0c01395

[60] Zineh, I. (2019). Quantitative systems pharmacology and model-informed drug development: A regulatory perspective. CPT: Pharmacometrics & Systems Pharmacology, 8(6), 336–339. https://doi.org/10.1002/psp4.12426

[61] Chaudhary, K., Poirion, O. B., Lu, L., & Garmire, L. X. (2018). Deep learning–based multi-omics integration robustly predicts survival in liver cancer. Clinical Cancer Research, 24(6), 1248–1259. https://doi.org/10.1158/1078-0432.CCR-17-0853

[62] Ekins, S., Puhl, A. C., Zorn, K. M., Lane, T. R., Russo, D. P., Klein, J. J., ... & Clark, A. M. (2019). Exploiting machine learning for end-to-end drug discovery and development. Nature Materials, 18(5), 435–441. https://doi.org/10.1038/s41563-019-0338-z

[63] Hasin, Y., Seldin, M., & Lusis, A. (2017). Multi-omics approaches to disease. Genome Biology, 18(1), 83. https://doi.org/10.1186/s13059-017-1215-1

[64] Leil, T. A., & Bertz, R. (2014). Quantitative systems pharmacology can reduce attrition and improve productivity in pharmaceutical research and development. Frontiers in Pharmacology, 5, 247. https://doi.org/10.3389/fphar.2014.00247

[65] Polak, S., Romero, K., Berg, A., Patel, N., Jamei, M., & Hermann, D. (2020). Model-informed drug development and regulatory science: Report of the ISoP special interest group. Journal of Pharmacokinetics and Pharmacodynamics, 47(2), 119–125. https://doi.org/10.1007/s10928-020-09676-2

[66] Roden, D. M., & George, A. L. (2020). The genetic basis of variability in drug responses. Nature Reviews Drug Discovery, 19(3), 141–152. https://doi.org/10.1038/s41573-019-0050-8

[67] Zineh, I. (2019). Quantitative systems pharmacology: A regulatory perspective on translation. CPT: Pharmacometrics & Systems Pharmacology, 8(7), 336–339. https://doi.org/10.1002/psp4.12405