International Journal of Physiology, Pathophysiology and Pharmacotherapy (IJPPP)
Website: https://ijppp.in/
ISSN: 3139-5481
Vol. 1, Issue 1, Jan-Jun, 2026
Page No.: 01-14

Review

AI-Guided Molecular Design for Accelerating Small-Molecule Drug Discovery

Amit Gupta
Head of Department, Rakshpal Bahadur College of Pharmacy, Bareilly

Corresponding Author:
Dr Amit Gupta

Email: amitfeb081@gmail.com

Conflict of interest: *NIL*

DOI: *10.62896/ijppp.1.1.01*

Article History
Received: 15/02/2026
Accepted: 20/03/2026
Published: 11/04/2026

Abstract:
The pervasive productivity crisis in pharmaceutical R&D, encapsulated by Eroom's Law, demands a fundamental transformation in drug discovery. Artificial intelligence emerges as a pivotal solution, not by replacing human expertise, but by forging a synergistic partnership with medicinal chemists. By leveraging vast, curated biological and chemical data, sophisticated molecular representations, and core AI paradigms—including predictive modeling, unsupervised exploration, and generative design—AI is re-engineering the discovery pipeline. It accelerates target identification, enables billion-compound virtual screening, and drives the *de novo* design of optimized lead candidates. Critically, AI-powered predictive ADMET and safety profiling frontload risk assessment, aiming to reduce costly late-stage attrition. This integration of AI across the workflow—from target to candidate—promises to compress timelines, lower costs, and enhance the precision of therapeutic design. Ultimately, the AI-guided paradigm represents a necessary and powerful evolution, positioning the field to finally bend the curve of Eroom's Law and deliver innovative medicines to patients with unprecedented efficiency.

Keywords: Eroom's Law, Artificial Intelligence, Drug Discovery, Molecular Design, Chemical Space, Generative AI

This is an Open Access article that uses a funding model which does not charge readers or their institutions for access and distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0) and the Budapest Open Access Initiative (http://www.budapestopenaccessinitiative.org/read), which permit unrestricted use, distribution, and reproduction in any medium, provided original work is properly credited.

Introduction

The pharmaceutical industry faces a profound and paradoxical challenge known as "Eroom's Law"—Moore's Law spelled backwards—which describes the observation that the cost of developing a new drug approximately doubles every nine years, a stark inverse of the exponential growth in computing power and technological advancement that defines the modern era[1]. This trend signifies a deep-seated productivity crisis, where despite unprecedented investments in research and development, genomics, high-throughput screening, and combinatorial chemistry, the number of new drugs approved per billion US dollars spent has steadily and alarmingly declined since the 1950s[2]. The drivers of Eroom's Law are multifaceted and complex, encompassing the "better than the Beatles" problem, where the low-hanging fruit of obvious therapeutic targets has been picked, forcing research towards increasingly complex and poorly understood diseases like Alzheimer's or many cancers; the "burden of knowledge" effect, where the sheer volume of scientific data and regulatory requirements lengthens training and development timelines; heightened regulatory scrutiny demanding larger and more rigorous clinical trials; and a pervasive risk-aversion within large organizations that prioritizes incremental improvements over radical innovation. This crisis threatens the sustainability of the entire drug discovery ecosystem, making it financially and temporally prohibitive to address unmet medical

International Journal of Physiology, Pathophysiology and Pharmacotherapy (IJPPP)
Website: https://ijppp.in/
ISSN: 3139-5481
Vol. 1, Issue 1, Jan-Jun, 2026
Page No.: 01-14

needs, and calls for nothing short of a paradigm shift in how we invent medicines[3][4].

It is within this landscape of escalating complexity and cost that artificial intelligence emerges not as a replacement for the medicinal chemist, but as an indispensable and transformative complement, offering a suite of tools to augment human intuition, creativity, and expertise. The core promise of AI in medicinal chemistry lies in its ability to navigate the astronomically vast "chemical space"—estimated to contain over 10^{60} plausible drug-like molecules—and the equally complex "biological space" of disease pathophysiology, at a scale and speed utterly beyond human capability[5][6]. AI and machine learning models can predict the binding affinity of a molecule to a protein target, forecast its pharmacokinetic and toxicity profiles, and even generate novel molecular structures with optimized properties, all before a single compound is synthesized in a lab. This capability fundamentally re-orders the discovery workflow, shifting from a linear, trial-and-error process of synthesize-test-analyze to a more efficient, design-focused cycle of in-silico prediction, virtual screening, and prioritized synthesis. Crucially, this augmentation liberates the medicinal chemist from the drudgery of routine data analysis and inefficient screening, allowing them to focus on higher-order strategic thinking: interpreting AI-generated outputs in a biological context, applying their deep knowledge of chemical synthesisability and structure-activity relationships (SAR) to refine suggestions, designing elegant synthetic routes for complex candidates, and ultimately making the nuanced, risk-balanced decisions that are the hallmark of drug discovery[7]. The AI acts as a powerful computational collaborator, proposing possibilities and identifying patterns hidden in multidimensional data, while the chemist provides the critical oversight, creative insight, and practical judgment to translate these possibilities into viable drug candidates. Therefore, the path forward is not an automated pipeline from algorithm to medicine, but a synergistic partnership where AI handles scale, pattern recognition, and brute-force computation, and the human expert provides direction, validation, and inventive genius. This collaborative model represents the most viable solution to bending the curve of Eroom's Law, injecting efficiency, speed, and a new level of intelligent design into the heart of medicinal chemistry, and ultimately accelerating the delivery of transformative therapies to patients[8].

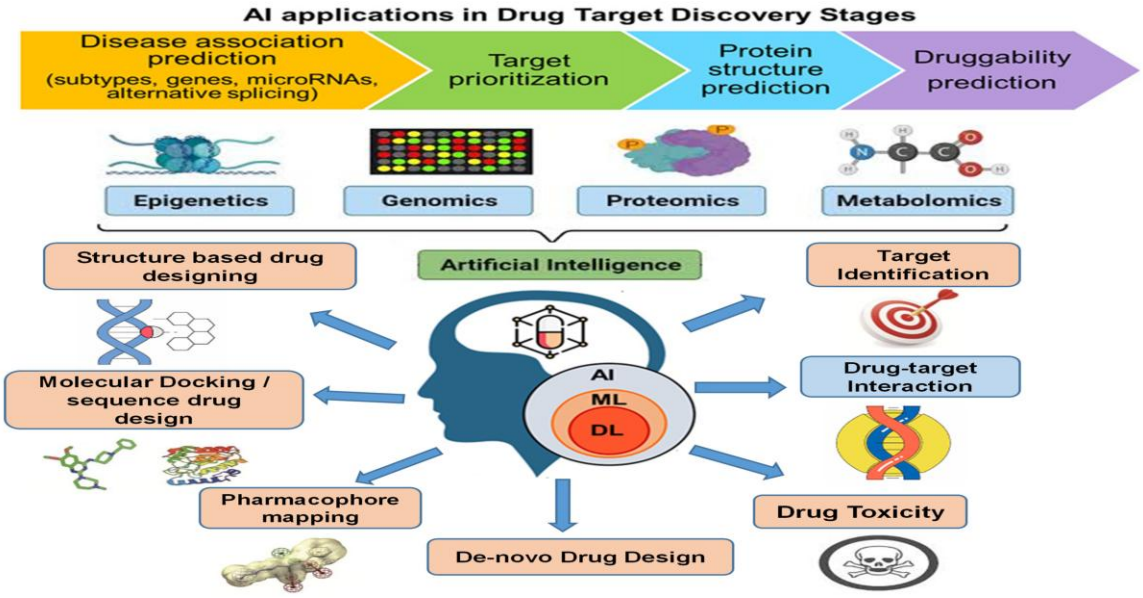

Fig: 1 AI-Augmented Drug Discovery Workflow: From Target to Candidate

Foundational Pillars: Data, Representations, and Core AI Paradigms

The transformative potential of artificial intelligence in bending the curve of Eroom's Law is entirely contingent upon the quality, quantity, and

accessibility of the data upon which it is trained, making data the fundamental fuel that powers the entire AI-driven discovery engine[9]. This foundational layer encompasses diverse and expansive repositories, most notably public chemical and structural databases like ChEMBL and PubChem, which serve as massive, crowd-sourced vaults of decades of pharmaceutical research. ChEMBL provides meticulously curated bioactivity data—millions of quantitative measurements linking chemical structures to biological targets—while PubChem offers a broader, more comprehensive repository of chemical substances, their structures, properties, and associated bioassays. These resources collectively form the historical record of medicinal chemistry, encoding both its triumphs and its failures, which AI models mine to learn the complex, non-linear relationships between molecular structure and biological effect. The data landscape has been radically expanded and enriched by the AlphaFold revolution, a paradigm shift precipitated by DeepMind's AI system that predicts protein three-dimensional structures from amino acid sequences with unprecedented, often experimental-level accuracy. This has democratized structural biology, providing predicted models for nearly all known proteins and thus illuminating the once-dark terrain of "undruggable" targets, enabling structure-based drug design on a proteome-wide scale. AlphaFold and its successors have not only supplied static snapshots but are also paving the way for understanding dynamic conformational states, thereby offering a more complete picture of the molecular machinery of disease[9]. However, this abundance of data introduces the critical and pervasive challenge of data quality, standardization, and integration. The real-world data residing in these public and proprietary sources is inherently noisy, heterogeneous, and plagued by biases; it is scattered across incompatible formats, suffers from inconsistent annotation, and is riddled with historical errors or reporting biases (such as the over-representation of certain chemical classes like "privileged scaffolds" or the under-reporting of negative results). The adage "garbage in, garbage out" is acutely true for AI, making sophisticated data curation—the processes of cleaning, standardizing, harmonizing, and applying expert-driven rules to flag potential errors—not a mere preliminary step, but a continuous, resource-intensive discipline that is as crucial as algorithm development itself. Effective AI requires the integration of disparate data modalities: chemical structures, bioactivities, genomic profiles, cellular imaging, clinical outcomes, and now protein structures, into a coherent, machine-readable knowledge graph that reflects the multifaceted nature of biological reality. Without this rigorous foundation of curated and integrated data, even the most sophisticated AI models are building on sand, liable to produce misleading predictions, perpetuate historical biases, or fail to generalize beyond their narrow training sets[10].

To render this raw, curated data intelligible to machine learning algorithms, molecules must be translated from chemical entities into mathematical or computational representations—a process that defines the very lens through which the AI "sees" and reasons about chemistry. The choice of representation profoundly influences a model's capabilities, biases, and performance. The most traditional and human-readable format is string-based representation, epitomized by the Simplified Molecular Input Line Entry System (SMILES), which uses a compact string of ASCII characters to denote atoms, bonds, and branching in a molecule, effectively linearizing its two-dimensional graph[11]. While SMILES are incredibly efficient for storage and database searching, their arbitrary syntactic rules (the same molecule can have multiple valid SMILES strings) and sensitivity to small changes (a single character typo can create a valid but completely different molecule) pose significant challenges for stable learning. This led to the development of more robust alternatives like SELFIES (SELF-referencing Embedded Strings), which are inherently constrained by grammatical rules to always generate syntactically valid molecular strings, a crucial advantage for generative AI models. A more natural and powerful paradigm is the graph-based representation, which explicitly models a molecule as a mathematical graph where atoms are nodes and bonds are edges, each adorned with feature vectors describing atomic properties (element, charge, hybridization) and bond characteristics (type, conjugation). This

representation aligns perfectly with the innate topology of a molecule and is natively processed by graph neural networks (GNNs), which learn by passing messages along bonds, effectively capturing the local chemical environment and long-range intramolecular interactions in a way that mirrors a chemist's intuitive understanding of functional groups and connectivity[12][13]. For tasks demanding an understanding of biomolecular interaction, such as predicting binding affinity or modeling protein-ligand docking, three-dimensional and geometric representations become essential. These encode the spatial coordinates of atoms, capturing critical features like molecular shape, volume, surface electrostatic potentials, and conformational flexibility. The rise of geometric deep learning, including techniques like SE(3)-equivariant neural networks, allows models to learn from this 3D data in a way that is invariant to rotations and translations—ensuring that a predicted property does not change simply because the molecule is rotated in virtual space. This hierarchy of representations—from linear strings to 2D graphs to 3D point clouds—provides a multi-faceted toolkit, allowing researchers to match the representation to the task: SMILES for rapid library generation, molecular graphs for predicting physicochemical properties, and 3D representations for understanding precise molecular recognition events at the heart of drug action[14][15].

Operating on these engineered representations is a diverse arsenal of core artificial intelligence and machine learning techniques, each offering distinct strategic advantages for different phases of the drug discovery pipeline. Supervised learning, the workhorse of predictive modeling, involves training algorithms on labeled datasets where the output (e.g., solubility, potency, toxicity) is known. Classical, interpretable models like Random Forests (ensembles of decision trees) and Support Vector Machines (SVMs) have long been used for quantitative structure-activity relationship (QSAR) modeling, offering robust performance on smaller, well-curated datasets and providing some insight into feature importance. Their modern, more powerful successors are deep neural networks, which, with their multiple layered architectures, can automatically learn hierarchical feature representations from raw input data (like graphs or 3D structures) and model extremely complex, non-linear relationships, often achieving state-of-the-art accuracy in property prediction tasks crucial for triaging virtual compounds[16][17]. In contrast, unsupervised learning operates on unlabeled data, seeking to discover hidden patterns, intrinsic structures, or natural groupings. Techniques like principal component analysis (PCA), t-distributed stochastic neighbor embedding (t-SNE), and clustering algorithms (e.g., k-means) are invaluable for exploring the vastness of chemical space, visualizing molecular datasets, identifying dense regions of desirable properties, or flagging outliers. They help medicinal chemists understand the landscape of their corporate or public databases, assess the diversity of compound libraries, and de-risk projects by ensuring novel compounds are not aberrant outliers in property space. Perhaps the most dynamic frontier is reinforcement learning (RL), which frames molecular design as a sequential decision-making game. In this paradigm, an AI "agent" learns to navigate chemical space by taking actions (e.g., adding a methyl group, changing a heterocycle) on a molecular "environment," receiving rewards or penalties based on how well the modified molecule meets a multi-objective goal (e.g., high potency, good solubility, low toxicity)[18][19]. Through iterative trial and error, often simulated in-silico across millions of steps, the agent learns an optimal policy for generating molecules that maximize this composite reward function. This makes RL exceptionally powerful for the *de novo* design of novel molecular entities, enabling the exploration of regions of chemical space not explicitly covered by training data, thereby fostering true innovation beyond the confines of known chemical series. Together, this portfolio of techniques—supervised learning for precise prediction, unsupervised learning for intelligent exploration, and reinforcement learning for goal-directed creation—forms a complementary toolkit. In practice, they are often used in concert: unsupervised methods can cluster data to define meaningful subsets for supervised model training; supervised models can act as the fast, virtual scorers that provide the reward signal for a reinforcement learning agent; and RL can generate novel candidates

International Journal of Physiology, Pathophysiology and Pharmacotherapy (IJPPP)
Website: https://ijppp.in/
ISSN: 3139-5481
Vol. 1, Issue 1, Jan-Jun, 2026
Page No.: 01-14

that are then validated by more accurate supervised predictors. This integrated AI workflow, built upon a foundation of curated data and sophisticated molecular representations, creates a powerful, iterative engine for hypothesis generation and testing, capable of accelerating the search for viable lead compounds from years to months or weeks, and offering a concrete, technologically mature pathway to overcoming the productivity crisis encapsulated by Eroom's Law[20].

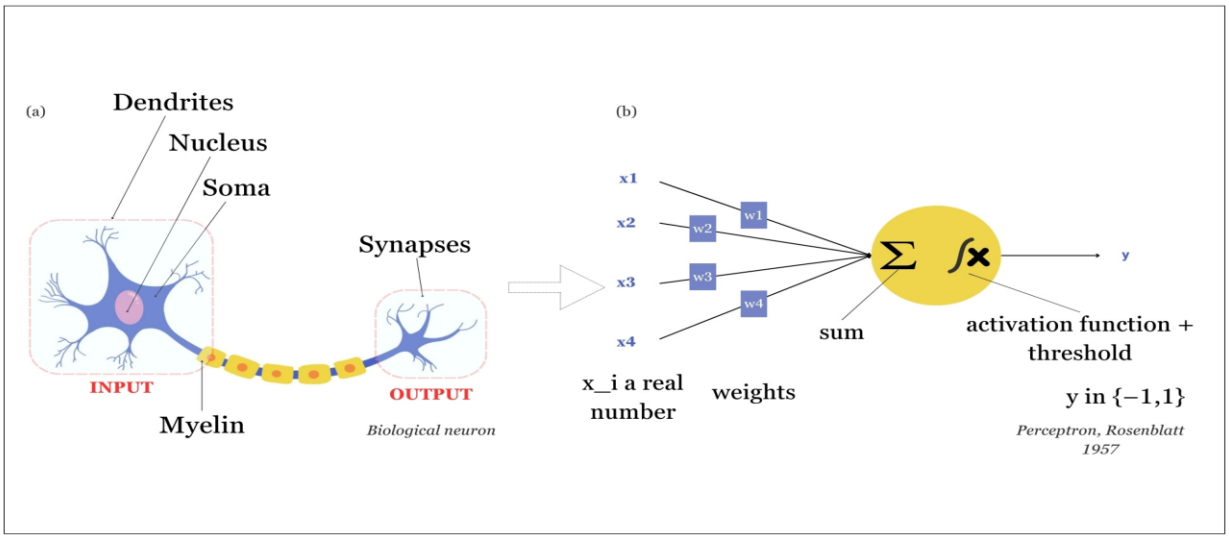

Fig:2 Artificial adaptation of a biological neuron.

4. Translational Applications Across the Drug Discovery Pipeline

The integration of artificial intelligence into the pharmaceutical research and development pipeline is heralding a paradigm shift, moving from a traditionally slow, serendipity-driven, and high-attrition process toward a precision-driven, hypothesis-rich, and accelerated model[21]. This transformation is most profoundly felt across three critical pillars: the initial discovery of targets and compounds, the intricate optimization of leads, and the crucial prediction of a molecule's behavior within the human body. Together, these AI-powered advances are coalescing into a new operating system for drug discovery, promising to deliver better medicines to patients faster and at a lower cost[22].

The journey of a new drug begins with the identification of a biological target—a protein or pathway implicated in a disease—and the subsequent hunt for a molecule that can modulate it. AI is dramatically accelerating this early-stage discovery. In target identification and validation, the challenge has shifted from data scarcity to data deluge. The advent of multi-omics—genomics, transcriptomics, proteomics, and metabolomics—generates vast, complex datasets that are often intractable for human analysis alone. AI algorithms, particularly machine learning and deep learning models, excel at mining these layers of biological information to uncover subtle, non-linear patterns. They can sift through genetic associations, gene expression profiles from diseased tissues, and protein-protein interaction networks to nominate novel, previously unexplored therapeutic targets. Furthermore, AI can validate these targets by predicting their "druggability," their role in disease pathways, and their potential safety profile based on biological similarity to known targets. This data-driven approach moves beyond established, often crowded targets to reveal entirely new biological mechanisms for therapeutic intervention[23].

Once a target is secured, the search for an initial "hit" compound begins. Virtual screening and hit identification represent one of the most mature applications of AI in the industry. Instead of physically testing hundreds of thousands of compounds in a lab—an expensive and time-consuming process known as high-throughput screening—AI can prioritize libraries containing millions, even billions, of virtual molecules. Using structures derived from techniques like cryo-electron microscopy or AI-predicted protein folding models

like AlphaFold, AI models simulate how likely each molecule is to bind to the target's active site. These models are trained on vast databases of known chemical structures and binding affinities, learning the complex physical and chemical rules that govern molecular interaction. This *in silico* triage can narrow a billion-molecule library down to a few hundred of the most promising candidates for physical testing, increasing the hit rate from often less than 1% to a significantly higher figure, thereby saving immense resources and time[24].

After identifying a hit, the arduous process of lead optimization begins, where chemists iteratively modify the molecule to improve its potency, selectivity, and drug-like properties. This phase, which traditionally takes three to six years, is being revolutionized by AI's ability to learn and propose[25][26]. The efficiency gains are starkly illustrated by case studies such as that of Insilico Medicine, which leveraged its generative AI platform to go from a novel AI-identified target for idiopathic pulmonary fibrosis to a fully designed, novel molecule, and then to a preclinical candidate in under 18 months—a timeline that defies conventional benchmarks. This acceleration is powered by generative chemistry models. Unlike virtual screening, which filters existing compounds, generative AI *designs* new ones from scratch. Given a set of desired parameters—strong binding to the target, avoidance of certain toxicophores, optimal solubility—the AI generates novel molecular structures that meet these criteria. It operates in an iterative cycle: propose molecules, predict their properties, receive feedback, and learn to propose better ones. This transforms lead optimization from a trial-and-error art into a precision engineering discipline[27].

This precision extends beyond chemical properties to patient stratification, enabling the design of precision therapeutics. AI can analyze clinical and molecular data from patient cohorts to identify distinct disease subpopulations. For instance, in oncology, AI can decipher which specific neoantigens a patient's tumor presents or what unique immune microenvironment it possesses. Armed with this knowledge, generative AI can then be tasked with designing highly specific immunomodulators, bispecific antibodies, or other therapies tailored to that precise biological context. This moves the industry from a one-drug-fits-all model toward a future where therapies are designed *de novo* for defined patient groups, dramatically increasing the likelihood of clinical success for those individuals. Perhaps the most impactful application of AI lies in its potential to reduce the staggering failure rate in drug development, where historically over 90% of candidates that enter clinical trials never reach approval, often due to unforeseen human toxicity or poor pharmacokinetics. Predictive ADMET (Absorption, Distribution, Metabolism, Excretion, and Toxicity) and safety profiling address this core challenge. By building sophisticated models trained on historical experimental and clinical data, AI can now provide remarkably accurate in silico predictions of pharmacokinetics and toxicity[28]. These models can forecast whether a molecule will be absorbed in the gut, how it will be metabolized by liver enzymes, its potential to inhibit vital cardiac ion channels (a common cause of failure), or its risk of causing liver injury[29][30]. Platforms can screen for dozens of toxicity endpoints simultaneously, far beyond the capacity of early experimental assays. This capability is fundamentally reducing late-stage attrition by de-risking candidates early. A molecule predicted to have a high probability of cardiac toxicity or poor metabolic stability can be deprioritized or redesigned before a single dollar is spent on costly animal studies or human trials. This shifts the "fail-fast" paradigm to the earliest, cheapest stages of research. It allows research teams to funnel resources only into molecules with a superior AI-predicted safety and pharmacokinetic profile, thereby increasing the robustness of candidates entering the clinic. The ultimate goal is a holistic, AI-driven profile for each molecule—encompassing its target potency, selectivity, and comprehensive ADMET characteristics—generated almost concurrently with its initial design.

Conclusion:

The integration of artificial intelligence into molecular design represents a definitive paradigm shift, offering a robust and multifaceted response to the existential challenge of Eroom's Law. This

transformation is not merely incremental but foundational, fundamentally rewiring the drug discovery process from a linear, high-attrition endeavor into a dynamic, intelligent, and iterative design cycle. By establishing a collaborative synergy between human expertise and machine intelligence, the field leverages the best of both worlds: the chemist's intuitive knowledge of synthesis, mechanism, and strategic judgment, combined with the AI's ability to navigate astronomical chemical and biological spaces, recognize hidden patterns, and generate novel hypotheses at unprecedented speed. This new operating system, built on pillars of curated data, advanced molecular representations, and a portfolio of AI techniques, delivers concrete efficiencies across the entire pipeline. It enables the rapid identification and validation of novel targets, the intelligent mining or *de novo* creation of optimized lead compounds, and the critical early de-risking of candidates through predictive ADMET profiling. The result is a virtuous cycle of accelerated discovery, where timelines are compressed, costs are contained, and the probability of clinical success is enhanced by design. Ultimately, AI-guided molecular design transcends a simple technological upgrade; it embodies a necessary evolution of the scientific method in medicinal chemistry. By empowering researchers to make more informed, data-driven decisions earlier than ever before, this partnership holds the promise of bending the stubborn curve of Eroom's Law. The path forward is one of continuous refinement and deepening collaboration, steering the industry toward a future where innovative, safe, and effective therapies reach patients with greater efficiency and precision.

References

1. Almustafa, E., Assaf, A., & Allahham, M. (2023). Implementation of artificial intelligence for financial process innovation of commercial banks. *Revista Gestão Social Ambiental, 17*, e04119. https://doi.org/10.24857/rgsa.v17n6-013
2. Bailleux, C., Gal, J., Chamorey, E., Mograbi, B., & Milano, G. (2024). Artificial Intelligence and Anticancer Drug Development—Keep a Cool Head. *Pharmaceutics, 16*(2), 211. https://doi.org/10.3390/pharmaceutics16020211
3. Bray, F., Laversanne, M., Sung, H., Ferlay, J., Siegel, R. L., Soerjomataram, I., & Jemal, A. (2024). Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: A Cancer Journal for Clinicians, 74*(3), 229–263. https://doi.org/10.3322/caac.21834
4. Cui, W., Aouidate, A., Wang, S., Yu, Q., Li, Y., & Yuan, S. (2020). Discovering Anti-Cancer Drugs via Computational Methods. *Frontiers in Pharmacology, 11*, 733. https://doi.org/10.3389/fphar.2020.00733
5. Frasca, M., La-Torre, D., Repetto, M., De Nicolo, V., Pravettoni, G., & Cutica, I. (2024). Artificial intelligence applications to genomic data in cancer research: A review of recent trends and emerging areas. *Discover Analytics, 2*(1), 10. https://doi.org/10.1007/s44257-024-00015-2
6. Fui-Hoon Nah, F., Zheng, R., Cai, J., Siau, K., & Chen, L. (2023). Generative AI and ChatGPT: Applications, challenges, and AI-human collaboration. *Journal of Information Technology Case and Application Research, 25*(4), 277–304. https://doi.org/10.1080/15228053.2023.2271265
7. Garg, P., Mohanty, A., Ramisetty, S., Kulkarni, P., Horne, D., Pisick, E., Salgia, R., & Singhal, S. S. (2023). Artificial intelligence and allied subsets in early detection and preclusion of gynecological cancers. *Biochimica et Biophysica Acta (BBA) - Reviews on Cancer, 1878*(6), 189026. https://doi.org/10.1016/j.bbcan.2023.189026
8. Ghimire, A., Tayara, H., Xuan, Z., & Chong, K. T. (2022). CSatDTA: Prediction of Drug-Target Binding Affinity Using Convolution Model with Self-Attention. *International*

Journal of Molecular Sciences, 23(15), 8453. https://doi.org/10.3390/ijms23158453

9. Google Trend on Keyword "Artificial Intelligence". (n.d.). Retrieved May 22, 2024, from https://trends.google.fr/trends/explore?date=today%205-y&q=%2Fm%2F0mkz&hl=fr
10. Holzinger, A., Keiblinger, K., Holub, P., Zatloukal, K., & Müller, H. (2023). AI for life: Trends in artificial intelligence for biotechnology. *New Biotechnology, 74*, 16–24. https://doi.org/10.1016/j.nbt.2023.02.001
11. Iwaloye, O., Ottu, P. O., Olawale, F., Babalola, O. O., Elekofehinti, O. O., Kikiowo, B., Adegboyega, A. E., Ogbonna, H. N., Adeboboye, C. F., Folorunso, I. M., et al. (2023). Computer-aided drug design in anti-cancer drug discovery: What have we learnt and what is the way forward? *Informatics in Medicine Unlocked, 41*, 101332. https://doi.org/10.1016/j.imu.2023.101332
12. Khan, S. U., Fatima, K., Aisha, S., & Malik, F. (2024). Unveiling the mechanisms and challenges of cancer drug resistance. *Cell Communication and Signaling, 22*(1), 109. https://doi.org/10.1186/s12964-024-01504-1
13. Khojasteh, H., Pirgazi, J., & Ghanbari Sorkhi, A. (2023). Improving prediction of drug-target interactions based on fusing multiple features with data balancing and feature selection techniques. *PLOS ONE, 18*(7), e0288173. https://doi.org/10.1371/journal.pone.0288173
14. Kim, J., Park, S., Min, D., & Kim, W. (2021). Comprehensive Survey of Recent Drug Discovery Using Deep Learning. *International Journal of Molecular Sciences, 22*(18), 9983. https://doi.org/10.3390/ijms22189983
15. Koutroumpa, N. M., Papavasileiou, K. D., Papadiamantis, A. G., Melagraki, G., & Afantitis, A. (2023). A systematic review of deep learning methodologies used in the drug discovery process with emphasis on in vivo validation. *International Journal of Molecular Sciences, 24*(7), 6573. https://doi.org/10.3390/ijms24076573
16. Krantz, A. (1998). Protein-site targeting: Diversification of the drug discovery process. *Nature Biotechnology, 16*(13), 1294. https://doi.org/10.1038/4303
17. Lee, M. (2023). Recent Advances in Deep Learning for Protein-Protein Interaction Analysis: A Comprehensive Review. *Molecules, 28*(19), 5169. https://doi.org/10.3390/molecules28195169
18. Li, Y., Zhang, L., Wang, Y., Zou, J., Yang, R., Luo, X., Wu, C., Yang, W., Tian, C., Xu, H., et al. (2022). Generative deep learning enables the discovery of a potent and selective RIPK1 inhibitor. *Nature Communications, 13*(1), 6891. https://doi.org/10.1038/s41467-022-34692-w
19. Madhukar, N. S., Khade, P. K., Huang, L., Gayvert, K., Galletti, G., Stogniew, M., Allen, J. E., Giannakakou, P., & Elemento, O. (2019). A Bayesian machine learning approach for drug target identification using diverse data types. *Nature Communications, 10*(1), 5221. https://doi.org/10.1038/s41467-019-12928-6
20. Niazi, S. K., & Mariam, Z. (2023). Computer-Aided Drug Design and Drug Discovery: A Prospective Analysis. *Pharmaceuticals, 17*(1), 22. https://doi.org/10.3390/ph17010022
21. Olayan, R. S., Ashoor, H., & Bajic, V. B. (2018). DDR: Efficient computational method to predict drug-target interactions using graph mining and machine learning approaches. *Bioinformatics, 34*(7), 1164–1173. https://doi.org/10.1093/bioinformatics/btx731
22. Peng, J., Wang, Y., Guan, J., Li, J., Han, R., Hao, J., Wei, Z., & Shang, X. (2021). An end-to-end heterogeneous graph representation learning-based framework for

drug-target interaction prediction. *Briefings in Bioinformatics, 22*(5), bbaa430. https://doi.org/10.1093/bib/bbaa430

23. Polishchuk, P., Madzhidov, T., & Varnek, A. (2013). Estimation of the size of drug-like chemical space based on GDB-17 data. *Journal of Computer-Aided Molecular Design, 27*(8), 675–679. https://doi.org/10.1007/s10822-013-9672-4
24. Sebastian, A. M., & Peter, D. (2022). Artificial Intelligence in Cancer Research: Trends, Challenges and Future Directions. *Life, 12*(12), 1991. https://doi.org/10.3390/life12121991
25. Shao, K., Zhang, Y., Wen, Y., Zhang, Z., He, S., & Bo, X. (2022). DTI-HETA: Prediction of drug-target interactions based on GCN and GAT on heterogeneous graph. *Briefings in Bioinformatics, 23*(1), bbac109. https://doi.org/10.1093/bib/bbac109
26. Sreelakshmi, K., Vishwakarma, P. K., Rao, S. G., Maqbool, A., Samal, D., Saini, R., & Thakur, G. (2024). Biotechnology and Genetic Engineering using AI: A Review. *International Journal of Intelligent Systems and Applications in Engineering, 12*(3s), 350–364. Retrieved May 24, 2024, from https://ijisae.org/index.php/IJISAE/article/view/4456
27. Visan, A. I., & Negut, I. (2024). Integrating Artificial Intelligence for Drug Discovery in the Context of Revolutionizing Drug Delivery. *Life, 14*(2), 233. https://doi.org/10.3390/life14020233
28. Wang, L., Song, Y., Wang, H., Zhang, X., Wang, M., He, J., Li, S., Zhang, L., Li, K., & Cao, L. (2023). Advances of Artificial Intelligence in Anti-Cancer Drug Design: A Review of the Past Decade. *Pharmaceuticals, 16*(2), 253. https://doi.org/10.3390/ph16020253
29. Wei, L., Niraula, D., Gates, E. D. H., Fu, J., Luo, Y., Nyflot, M. J., Bowen, S. R., El-Naqa, I. M., & Cui, S. (2023). Artificial intelligence (AI) and machine learning (ML) in precision oncology: A review on enhancing discoverability through multiomics integration. *The British Journal of Radiology, 96*(1145), 20230211. https://doi.org/10.1259/bjr.20230211
30. Xu, Y., Liu, X., Cao, X., Huang, C., Liu, E., Qian, S., Liu, X., Wu, Y., Dong, F., Qiu, C.-W., et al. (2021). Artificial intelligence: A powerful paradigm for scientific research. *The Innovation, 2*(4), 100179. https://doi.org/10.1016/j.xinn.2021.100179
31. You, Y., Lai, X., & Pan, Y. (2022). Artificial intelligence in cancer target identification and drug discovery. *Signal Transduction and Targeted Therapy, 7*(1), 156. https://doi.org/10.1038/s41392-022-00994-0

*****