INTERNATIONAL JOURNAL OF PHARMACEUTICAL SCIENCES
[ISSN: 0975-4725; CODEN(USA): IJPS00]
Journal Homepage: https://www.ijpsjournal.com

OPEN ACCESS

Review Article

Generative AI and Machine Learning in Drug Discovery

Gaurav Palsamkar*, Aditya Gurav

Govindrao Nikam College of Pharmacy, Mumbai University, 415606-Sawarde, India.

ARTICLE INFO

Published: 20 Nov 2025
Keywords:
Drug Discovery, Artificial Intelligence (AI), Machine Learning (ML), Generative AI (GenAI), De Novo Design, Graph Neural Networks (GNNs)
DOI:
10.5281/zenodo.17658072

ABSTRACT

Finding new drugs takes a lot of effort and money. It often fails and drags on for years. New tools in AI and machine learning fix old problems like spotting targets, tweaking leads, and guessing toxic effects. Key methods now include Graph neural networks, GANs, VAEs, and LLMs. These fit right into drug making steps. They speed up checks, create molecules, and predict how drugs move in the body. Large pharmaceutical and biotech companies rely on AI and ML tools. These tools save time and select stronger drug candidates. Challenges remain with data quality, regulatory guidelines, and model transparency. Ethics and keeping data private need constant care. AI-led drug discovery will grow with clear and multi-type methods. These boost teamwork between people and machines. They build trust and make results reliable. Overall, AI and ML change drug discovery. They bring quicker, safer, and cheaper treatments for health problems worldwide.

INTRODUCTION

❖ The Challenges in Drug Discovery :

Drug discovery is one of the toughest and expensive tasks in modern science. Even with major tech improvements, the drug field still hits many roadblocks. These slow down the quick creation of new medicines. The price tag for finding new drugs stays huge. It costs approximately $2.5 to $2.6 billion to develop a new drug. That makes it the priciest type of research in any field. The high cost comes from the tough process and the high chance of failure in drug work.(1)(2)

Drug development usually takes 10 to 15 years. This spans from the first find to approval for sale. Such a long wait slows down returns on investment. It can also cut the useful time left on patents for the new drug.(1)(2) Just 13% of drugs that start clinical trials gain approval from regulators. This low success rate drives up the total costs of development.(1) Standard drug discovery methods often use trial-and-error steps based on guesswork. These steps have low power to predict

***Corresponding Author:** Gaurav Palsamkar
Address: *Govindrao Nikam College of Pharmacy, Mumbai University, 415606-Sawarde, India.*
Email ✉: gpalsamkar03@gmail.com
Relevant conflicts of interest/financial disclosures: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

outcomes. This shows the clear need for better ways to find drugs.(3) A main problem lies in finding and proving drug targets. This task grows hard due to gaps in our grasp of disease cause and tricky links between proteins.(1)

AI possesses the capacity for transformative change; however, it also presents challenges such as the "black box" issue, wherein the decision-making processes lack transparency, thereby complicating regulatory acceptance.(3) (4)Furthermore, bias present in training datasets can adversely affect the accuracy of AI, necessitating extensive validation to confirm the reliability of AI models across diverse populations.(3)

The ethical dilemmas associated with clinical trials, the equitable distribution of new therapies, and concerns regarding intellectual property complicate the drug discovery process, particularly for diseases that affect small patient populations.(3) Striking a balance between incentivizing innovation and ensuring broad accessibility continues to pose a significant social and ethical challenge.(3)

❖ The Rise Of AI and ML in Biomedicine :

The growth of artificial intelligence (AI) and machine learning (ML) in biomedicine brings big changes. It affects drug discovery and development the most. AI tools, like those that create new ideas, speed up the process. They help find targets, design molecules, and improve them. These steps used to take a lot of time and money. AI learns from old chemical and biological data. It makes new drug options. This speeds up new treatments. (5)

Clear AI methods matter a lot in biomedicine. They show how predictions or choices happen. This lets doctors and researchers see the reasons. Trust counts in health care. So does being responsible. Tools like gradient saliency maps and class activation mapping show key features. They give clear views of tough biomedical data. This helps with better choices in clinics. (6)

ML tools work well in early drug stages. They guess biological effects. They improve lead compounds. They also boost trial results. Models with graph neural networks and transformers do better than the old ways. They are more accurate and quicker. This supports care made for each person. (7) AI also helps repurposing of drugs. It cuts costs by finding fresh uses for old ones. It speeds up screening compounds and spotting toxicity. (8)

Problems still exist. Data can be poor. Rules are hard to meet. Models are tough to explain. Yet AI and ML in biomedicine can change treatment making. It will go faster and cost less. It fits each patient's needs. (9) (5) (6) (7) (8)

❖ Definitions and Scope :

Machine learning, or ML, means computer programs that let systems learn and get better on their own. They spot patterns in big sets of data. In biomedicine, this helps predict traits, sort biology info, and speed up analysis tasks. (6) Deep learning, or DL, is part of ML. It relies on layered fake neural networks. These handle tough, detailed data like gene codes or molecule shapes. This leads to better feature pulls and sharp guesses in areas such as drug trait forecasts and image spotting. (10) (11)

Generative AI, or GenAI, covers smart AI tools that make fresh data like real ones. Models like generative adversarial networks (GANs), variational autoencoders (VAEs), and large language models (LLMs) play a bigger role in biomedicine. They craft new molecule builds, plan

drug options, and explain biology sequences. (12) (10) GANs work by putting two networks against each other, a creator and a checker to boost fake data quality. VAEs turn molecule traits into codes for true-to-life changes. LLMs use deep network setups like transformers. They produce clear biology or chemistry sequences. They also aid in forming new ideas. (10).

De novo design programs can create entirely new drug molecules by predicting what shapes and chemical structures might best fit a biological target. AI helps clinical tests, too. It picks patients with guess models, runs fake trials, and checks mixed biology data. This boosts trial speed and result forecasts. (11) (10)(6) Such tools aim to push exact medicine forward, raise drug discovery rates, and change how research works in clinics. (11)(6)

1. CORE TECHNOLOGIES

❖ Traditional ML and Deep Learning :

'Support Vector Machines' (SVMs) and 'Random Forests' serve as key tools in Quantitative Structure-Activity Relationship (QSAR) modeling and virtual screening for drug discovery. SVMs stand out by spotting active compounds versus inactive ones. They do this through the best dividing line in a high dimensional space of chemical traits. This gives strong, consistent results in arranging and sorting data. (1) 'Random Forests', by contrast, create groups of decision trees. These handle tough chemical data sets well. They manage high changes and linked traits for sorting and prediction tasks in QSAR work.(1) (3)

Deep learning (DL) boosts QSAR modeling even more. Neural networks like 'Auto-Encoders' or deep feed-forward types capture complex curved links in molecule data. (11) DL methods show better detail and accuracy in guessing protein-protein links and quick virtual screens than standard ML tools. (1) For instance, deep neural networks and graph networks predict ADMET traits, absorption, distribution, metabolism, excretion, toxicity and other bio activities with great skill. They spot patterns in molecule graphs or chains. (11)

Today's drug development paths mix classic ML with DL to boost prediction power and clarity. SVMs and Random Forests shine for quick, easy-to-grasp guesses in early checks. Deep learning leads on big data sets with full notes and fine trait pulls needed for lead tweaks and risk checks. (4) (11) (1) (3)

❖ Graph Neural Networks

'Graph neural networks', or GNNs, serve as key tools in today's drug discovery. They model molecules as graphs. Atoms act as nodes, and bonds serve as edges. (13) (14) This method mirrors the real shape and links in molecules. It lets GNNs grasp connections and interactions vital for spotting molecular traits.

GNNs shine in tasks like forecasting binding strength, toxicity, and solubility. (13)(14) (15) They pull data from nearby nodes step by step. This builds strong views of the graph that show local and broad chemical settings. Take 'GNNSeq', a GNN based on sequences. It shows strong results and wide use in spotting protein-ligand links. It works on big, varied data sets. This aids quick virtual screens and early drug steps. (15)

GNNs also help create new molecules with needed traits, beyond just forecasts. (14) (16) Tools built on GNNs learn to make valid chemical forms tuned for biology or drug body effects. These setups often mix with reward learning or autoencoders. They aid fresh drug builds and lead

tweaks. (14) (16) GNNs also predict drug clashes and build health knowledge graphs. This widens their role in drug pipelines (16) (13)

New geometric GNN designs now handle 3D shapes and forms. This boosts truth in forecasts and creations. (14) (13) In all, GNNs' skill at mapping tough molecular setups makes them key for forecasts, builds, and finds in AI drug work. (14) (13) (16) (15).

❖ Generative AI Models

Generative AI, stands out as a key tool in creating new drugs from scratch. It helps build fresh chemical forms with needed biology and body-processing traits. The top models here are 'Generative Adversarial Networks', called GANs, and 'Variational Auto-Encoders', known as VAEs. Both learn deep patterns in molecules. They also make new compounds that go past known chemical sets.(17,18)

GANs work via a contest setup. A generator makes new molecule ideas. A checker tests if they fit rules. Over many rounds of training, this boosts the quality of designs. (19) In drug hunts, these GANs create SMILES codes or graph forms for molecules. They follow chemistry laws and improve traits like binding strength, fat-liking, and ease of making. (20) Special conditional GANs add molecule details or activity data. This guides the output toward set drug effects. (21)

VAEs differ by using chance-based code-and-decode steps. They turn molecules into a smooth hidden area. This lets easy pulls and blends between. (22) Exploring this area aids tweaks for goals like strength, dissolve rate, and body safety traits. Adding reward learning and stats-based tweaks raises VAEs' skill at making valid, useful chemicals. (23)

New work blends VAE and GAN setups. It mixes VAEs' steady nature with GANs' real-look output. This fixes issues like limited variety and boosts diverse builds. (18) Such blends have shown real wins. They create drug-ready molecules with good fit scores. Some even prove hits in tests for cancer and brain decline paths. GANs and VAEs lead the way in fresh molecule building. They speed up early drug steps with smart, auto chemical ideas. (17)

❖ LLMs and Protein Modeling

Large language models (LLMs) and tools like 'AlphaFold' change biomedical research. They improve data analysis, knowledge pull, and structure forecasts. LLMs learn from huge sets of science texts, medical records, and biology notes. They grasp context. They find key facts. They link scattered biochemistry info for smart drug and protein design.(24) NLP methods let LLMs scan papers for protein links, disease ties, and drug-target pairs. This speeds up idea creation and aids fact-based discovery steps. (25)

New LLMs like 'BioGPT', 'PubMedBERT', and 'Galactica excel' at digging through biomedical texts. They spot process patterns that old search tools missed. (26) These models link biology items on their own. They pull protein-role ties. They help study pathways. Adding LLMs to full biology data flows updates protein banks in real time. It connects sources like Universal Protein Resource, Protein Data Bank, and trial data. These steps boost the grasp of protein actions, fold changes, and mutation effects in various diseases. (27)

AlphaFold pairs well with LLMs. It brings big advances in protein shape prediction. Deep neural nets train on evolution rules, physics basics, and shape math. (28)AlphaFold predicts 3D protein forms from amino acid lists with near-lab precision. It reveals active spots, shape shifts, and

molecular bonds. Linking it to LLM systems ties sequence notes to shape details. This ranks drug targets and explains variant effects at the molecule scale. (29)

LLMs and AlphaFold start a fresh time in AI-aided protein work. Text insights and shape models team up. They sharpen biology knowledge. They forecast mutation outcomes. They quicken drug design aimed at proteins. (24) (28)

❖ **Explainable AI**

Explainable AI, or XAI, builds trust, openness, and responsibility in drug discovery processes. It makes sure predictive models give right results and clear reasons people can grasp for key choices. In pharma work, AI tools like deep learning face heat as black boxes. Scientists can't see how they make choices. XAI fixes this with visual, text, or feature breakdowns. These show why a molecule, protein target, or dose got picked. (30)

XAI, fits into drug development pipelines that use chemical and biological methods. It picks out molecular features that influence predictions of activity. These include functional groups, 3D structures, and simple traits. Methods like SHAP values, LIME approaches, and attention systems reveal the main factors.

These include atom kinds, binding spots, or drug safety traits. They show what sways model results most. Chemists then check these against known drug rules. This cuts risks from false leads. (31)

New work pushes full systems that link prediction to after the fact checks. They let users trace choices through step by step logic. Teams use them in QSAR, virtual hunts, and toxin checks. Here, grasping the model's thinking shapes which compounds get tested and ranked.(32)

For rules and fair AI use, clear views matter a lot. Groups and experts need explained results to back safety and effect claims. XAI helps speed up approvals. It also aids teamwork between data pros and biology experts. This links math guesses to real bio sense. (33)

Fresh ideas mix graph networks with clear add-ons. They uncover ties between molecule parts and drug effects. Such links help grasp structure-function ideas. They steer making of new drugs and spot how they work. This sparks fresh treatment plans. (34)

In all, XAI turns tough model results into useful science facts. It boosts openness, repeat tests, and smart choices. From finding hits to final tweaks in drug work, XAI matches computer tips with lab proof and drug sense. (30) (31)

2. APPLICATIONS ACROSS THE PIPELINE:

❖ **Target Identification & Validation:**

- Predictive Target Discovery: Machine learning models examine large sets of genomic and proteomic data. They spot genes and proteins linked to disease processes. This boosts the search for new targets. (35)
- Pattern Recognition in Omics Data: AI tools find weak links in transcriptomic and proteomic data. These links reveal hidden ties between molecules. They help rank potential targets. (36)
- Generative AI for Novel Target Prediction: Models like GANs and VAEs create fake biological networks. They produce possible biomolecular patterns. This aids in suggesting fresh disease targets. (37)

- Integrated Genomic-Proteomic Target Mapping: Machine learning combines genomic and proteomic traits. It pinpoints key spots in molecular networks that cause disease. This refines the list of target candidates. (38)
- Supervised Learning for Gene-Target Interaction Prediction: Supervised machine learning uses past target-disease info to train. It forecasts new useful targets. This works well for tough illnesses like cancer and brain decay. (39)
- Network-Based Target Selection: Graph neural networks map protein interactions. Advanced machine learning spots key or strong targets in disease paths. (36)
- Domain Adaptation for Cross-Dataset Generalization: AI applies domain shifts to match data from varied omics and patient records. This makes target forecasts more reliable across sources. (11)
- Multimodal AI for Target Validation: Multimodal systems blend genomic, transcriptomic, proteomic, image, and patient data. They offer full views on a target's role in function. (40)
- Clinical Relevance Assessment: AI validation checks a target's real-world use. It predicts trial results or patient reactions before tests start. (40)
- Simulation of Biological Pathways: Generative AI builds mock disease networks. It tests cell reactions to changes. This allows virtual checks of drug target ideas. (37)
- Automated Prioritization of Drug Targets: Machine learning systems score targets using molecular, trait, and patient data. They pick the best paths for treatment. (36)

❖ **Hit Discovery & Virtual Screening:**

- ML-Based Scoring Boost: Machine learning swaps out old docking scores. It learns from big sets of test data on affinities. This predicts how ligands bind to targets better. (41)
- Fast Virtual Screening: ML tools scan millions of compounds quickly. They guess binding strengths and poses with more speed and accuracy. (42)
- Quick Hit Finding: GenAI, GNNs, and ML scores team up. They find strong hits fast with real bio fit and fewer errors. (42)
- AI for New Molecule Creation: Models like GANs and VAEs build fresh chemical shapes for set targets. Then ML checks them for drug traits and effects. (43)
- GNNs for New Hit Making: GNNs in AI generators create molecules shaped for target pockets. (43)
- Learning to Improve Molecules: Reinforcement learning tweaks structures step by step. It balances binding strength, body processing, and low harm. (39)
- Mixing ML with Quantum Data: ML pairs with quantum details to better guess docking spots and energies. It links data methods to physics rules. (44)
- Scores That Work Widely: Deep learning scores train on varied targets. They deliver strong screening results across protein groups. (45)
- GNNs in Docking Tasks: GNNs show molecules as graphs of atoms and bonds. They spot key space and electron traits for binding guesses. (46)

- AI Pipelines for Re-Scoring: Transformer or GNN AI models polish docking outputs. They sharpen affinity guesses and order accuracy. (46)
- GNNs Improve Scores: GNNs added to docking steps predict binding energies and ranks more right than old ways. (47)
- GNNs Refine Docking Results: Docking outputs link with GNN hidden traits. This cuts wrong hits and drops weak poses. (48)
- Better Target Choice: ML-GNN score mixes spot picky binders. They blend chem details with shape limits. (48)
- Maps of Protein-Ligand Touches: GNNs use contact maps to track close interactions at the site. They make predictions easier to grasp. (49)

❖ **De Novo Molecule Generation:**

- VAEs for Property-Guided Molecule Creation: Variational autoencoders build smooth hidden models of molecules. They help create fresh compounds with key traits, such as drug-like features and biological effects. This happens by drawing samples from the hidden space.(50)
- Graph VAEs to Boost Structural Variety: These VAEs rely on graph formats to grasp how atoms connect and their shapes. They yield new compounds that chemists can make and that offer fresh structures for drug hunts. (51)
- GANs Bring New Ideas to Chemical Building: Generative adversarial networks craft novel molecule forms. A builder competes with a checker, which sparks bold, drug-ready molecules outside the range of known data sets. (52,53)
- GANs for Libraries Aimed at Specific Targets: These networks shape compound sets for certain biology goals. They fine-tune outputs with special rewards or loss terms to raise success in computer-based screens. (54)
- Reinforcement Learning for Targeted Molecule Building: RL systems tweak or assemble molecules step by step. They chase top scores tied to traits like strength, ease of making, or mixed goals to fit tough treatment demands. (52,55)
- RL for Balancing Multiple Goals: These RL setups pair with trait forecasters. They adjust molecules at once for strength, safety, and body handling to meet strict needs for drug picks. (55)
- Blending Generative Tools in Work Flows: VAEs, GANs, and RL team up for making, picking, and refining molecule hopefuls. They speed up finds and add fresh chemical ideas. (51,54)
- Quick Scans of Chemical Options: These methods open up huge chemical realms. They spot new compounds that old screen ways would miss. (54)
- Ease of Making and Live Input: New pipelines add rules for buildable compounds. They draw from screen feedback to sharpen outputs for drug fit and growth promise. (53)
- Promise for Custom and Hard Targets: Better data and smarter AI let these models aid tailored drugs. They probe tough health zones and push forward exact treatments. (56)

❖ **Lead Optimization:**

- Graph-Based Scaffold Changes: Generative AI applies graph neural networks to turn molecular structures into graphs. It suggests changes to scaffolds. These boost potency, selectivity, and bioactivity. The AI draws on patterns from chemical data.(57,58)

- Forecasting Bioisosteric Swaps: Models built on graphs propose new bioisosteres. They also suggest tweaks to functional groups. This raises metabolic stability and permeability. Target binding stays strong. (58)

- Balancing Multiple ADMET Traits: Machine learning tools adjust several ADMET features at once. These include solubility, toxicity, and metabolic stability. The process happens as molecules form. It weighs drug strength against safety risks. (59,60)

- Adding Synthetic Ease: Generative models build in rules for real-world making. This ensures molecules work well in tests. They also prove easy to create in labs.(59)

- Active Learning from Lab Results: AI runs loops of design steps. It creates candidates for synthesis. Teams test them in labs. Results feed back to sharpen the model's guesses and designs. (61)

- Modeling Structure-Activity and Structure-Property Links: Graph neural networks map out tricky SAR and SPR ties. They spot odd scaffold shifts. These enhance ADMET traits. Drug activity holds firm.(62)

- Cutting Time in Lead Optimization: AI guides scaffold work and ADMET forecasts. This speeds the shift from hits to clinical leads. It cuts down on random lab trials.(57)

- Better Odds in Drug Creation: Tuned candidates display stronger safety and drug movement in the body. This lowers failure rates in late preclinical work. (57,58)

- Data-Guided Molecular Tweaks: Models use big sets of chemical and drug data. They predict effects of structure changes with strong accuracy. This helps chemists plan smart designs.(60)

- Coming Ties to Clear AI: Gains in explainable AI will clarify model choices in optimization. This builds trust. It spurs use in drug research.(61)

❖ **Predictive Toxicology & ADMET:**

- Deep Learning for Toxicity Prediction: Models based on deep learning, like graph neural networks, use big sets of labeled data to spot liver damage, heart risks, gene harm, and other bad effects. This flags risky compounds early. (63,64)

- Multi-Task Checks for ADMET: Frameworks for multi-task learning test several drug traits at once. They look at how stable drugs stay, how well they pass barriers, and side effects. This speeds up picking leads.(62)

- Molecular Graphs for Better Views: Models use graphs of molecules to show atom links and electron flows. These explain ADMET traits better than old methods. (65)

- AI Generation to Boost Safety: AI tools create and tweak new molecules. They focus on better safety and ADMET by studying toxicity and drug movement data.(66,67)

- Refining with Reinforcement Learning: These learning setups tweak molecules step by step. They reward good ADMET traits and punish toxicity. (68)

- Clear AI for Risk Insights: Methods like attention tools and feature maps point out molecule parts that cause toxicity. This builds trust and understanding. (69,70)
- Aiming Fixes at Molecules: Clear AI spots bad groups in molecules that lead to toxicity or breakdown issues. This guides smart changes to cut risks. (71)
- Fast Checks on Safety: AI screens thousands of similar compounds quickly for toxicity and ADMET. It cuts expensive failures late in tests. (63,65)
- Better Rules Followed: Clear AI risk checks help with approvals. It matches computer forecasts to expert checks. (68,69)
- Smarter Choices Before Trials: Toxicity and ADMET forecasts spot unsafe drugs early. They drop them fast, save resources, and aid patient safety. (64,66)

❖ **Protein Therapeutics & Antibodies**

- AI Tools for Finding New Antibodies: Tools from AION Labs and DenovAI rely on deep learning. They craft new antibodies with better binding strength, easier development, and less risk of immune reactions. This comes from studying sequence, shape, and function data. (72,73)
- Going Past Natural Antibody Sets: Models that generate designs create antibodies with fresh binding parts and bases. This boosts the range of drug options beyond what nature provides. (74)
- Cutting Time and Cost in Development: AI-based design speeds up antibody creation. It beats old lab methods like hybridoma or phage display. (75)
- Spot-On Predictions of Protein Shapes with AlphaFold: AlphaFold figures out 3D protein forms from amino acid lists. It does this with great precision. The results give key shape details for drug protein work. (76)
- Blending Shape Predictions with Design Models: AlphaFold's shape forecasts pair with AI generators. Together, they build proteins that hold up better, bind stronger, and last longer in the body. (77,78)
- Building Proteins from Scratch: AI generators make new enzymes, frames, and drug proteins. They tailor these for exact chemical tasks. The process uses sequence-to-shape maps and reward-based learning. (74)
- Tailored Changes to Antibodies: AI helps tweak antibody parts that vary or stay constant. Changes boost their power, lifespan, and control over immune effects. (73)
- Faster Path for Biologic Drugs: AI trims the wait from finding targets to picking candidates. It cuts down on wide lab tests. (72)
- New Types of Drug Options: AI design aids fresh biologics. These include dual-target antibodies, drug-linked antibodies, and custom protein frames.(75)
- Solid Bases for Protein Tweaks: AlphaFold's exact shapes offer sure starting points. They guide more changes and AI edits.(77)

3. INDUSTRIAL LANDSCAPE & CASE STUDIES:

Generative AI and machine learning have quickly changed drug discovery in the industry. They alter the way drug firms spot, build, and refine potential drugs. These tools blend computer models, auto processes, and data review to speed up a process

that used to take a long time and cost a lot. Drug companies around the world now use AI-based steps to gain edges in picking targets, shaping molecules, and early checks.(79,80)

❖ **Evolving Industrial Landscape**:

Drug and biotech firms have added generative AI tools into their research work to fix main problems in old drug finding methods. These setups can scan huge chemical areas, guess how molecules act, and make compounds with good drug effects. (79) Top groups like Insilico Medicine, Exscientia, Recursion Pharmaceuticals, and BenevolentAI apply AI creation models to speed up drug line building and choices. (80)

Big drug firms such as Pfizer, AstraZeneca, and Novartis have formed key ties with AI start-ups. They use machine learning to create compounds, predict harm, and model how structures link to actions. For example, team efforts with cloud AI systems now allow quick making and fixing of drug options in areas like cancer, brain diseases, and body function issues. (2)

The use of creation models—like variational autoencoders (VAEs), generative adversarial networks (GANs), and diffusion models—lets scientists make new compounds with needed physical traits. This cuts down on lab tests. Such methods allow joint goal fixing to match strength, choice, and safety in drug building. (81)

❖ **Impactful Case Studies in the Industry**:

- A strong example comes from Insilico Medicine's AI-discovered drug for fibrosis, INS018_055(Investigational drug name for Rentosertib, a small-molecule inhibitor designed using generative AI to treat idiopathic pulmonary fibrosis). It has advanced to Phase II trials. This proves AI generative models can develop drugs ready for clinical tests in months, not years.(82)
- Exscientia has pushed forward with AI in drug building. They made DSP-1181 for obsessive-compulsive issues with reward learning and active learning tools. These AI-built drugs started early clinic checks with much shorter find times. (83)
- BenevolentAI's system uses text processing and graph machine learning to find new links between diseases and targets. It helps pick usable targets. This method helped shift old drugs for COVID-19 use. It shows AI's range in quick drug finding. (84)
- Recursion Pharmaceuticals maps cell reactions on a large scale with image phenomics and machine learning. Their setup joins auto microscopes and creation models. This has led to new paths for rare gene diseases and swelling issues. (85)

❖ **Industry Advancement and Future Directions:**

AI mixed with lab auto tools, cloud systems, and fast screens has set a new way for drug finding in the field. Firms now take mixed paths that join generative AI with robots and chip labs. These allow self-run cycles of design, build, and test to fix molecules right away. (86)

Looking ahead, the focus shifts to transparent AI that builds trust and follows rules in design processes. Shared learning platforms will enable secure, private collaboration between pharma partners. As AI in drug discovery expands, it will transform drug development into faster paths to safer, more effective treatments. (2,79–81,85,86)

4. BENEFITS, CHALLENGES & LIMITATIONS:

❖ Benefits:

- Finding and Checking Targets: Machine learning combines various data types from biology. It reveals new targets and checks links between drugs and diseases. This aids tough or rare conditions most. [87], (87)
- Better Prediction Power: Deep learning tools forecast how drugs bind, their body effects, and risks with more detail. This cuts losses in late drug stages. (88)
- Clear Insights from Explainable AI: Tools that explain AI show how results form. They build faith among chemists and rule makers.(89)
- Lower Costs and Faster Times: AI cuts need for pricey lab tests and mass screens. It shortens drug timelines and expenses a lot. (90)
- Balancing Many Goals: AI tunes several drug traits at once. It matches strength, safety, and body handling well. (89)
- Aid for Custom Treatments: AI tools shape therapies for single patients. They blend clinic notes and gene data. (91)

❖ Challenges:

- Data Quality and Access: AI needs big, mixed, clean data sets. Short, slanted, or locked data harms model reach and truth. (92)
- Repeatability and Set Rules: No standard ways to check models or make molecules block repeats across tools and steps. (93)
- Problems with Clarity: Most deep models hide their work. This slows uptake since reasons for outputs stay vague. (88)
- Fitting into Current Steps: Adding AI to drug flows demands skill training, tech setup, and team shifts. These pose real blocks. (3)
- High Compute Needs: Deep models demand lots of power and space. This strains small groups most. (94)
- Privacy and Moral Issues with Data: Sharing key clinic and gene info for AI training sparks privacy and rule worries. Strong rules must guide it. (94)
- Rule-Making Barriers: AI-created drugs face unclear approval routes. Safety, potency, and patent reviews spark problems.(87)
- Fitting Too Close to Data and Wide Use: Models may cling to train data. This harms work on new chemicals or biology setups.(88)
- Spreading Slants: Data slants carry over or grow in AI. This may miss small groups or odd traits.(92)

❖ Limitation:

- Blocks from Size and Reach: Steep costs and skill needs curb wide use. Small labs often miss full gains.(94)
- Unclear Rights to Ideas: AI-generated molecules rise patent laws. This creates doubt about who owns machine-generated molecules.(87)
- Need for Real Tests: AI guesses still demand full lab checks. This stresses mixed human-AI approaches.(88)

- Right Use of AI: AI choices in drugs must fit moral rules. They prevent harm or unfair care access.(3)

- Handling Open Data: Strict guidelines and patient consent, ensure proper use of clinical data in AI drug development.(94)

5. FUTURE DIRECTIONS:

❖ Future Directions of Generative AI and Machine Learning in Drug Discovery

Generative AI and machine learning change drug discovery. They boost the pace and success of new treatments. Main paths ahead cover clear AI and cause-based AI. They include mixed AI systems that blend biology data types. Better teamwork between people and AI comes with special training. AI also runs the full drug discovery process. This includes new types of treatments.(95,96)

❖ Explainability and Causal AI

AI tools in drug discovery act like hidden boxes. Clear AI keeps decisions open and simple to follow It fits needs in drug work. Cause-based AI goes further. It spots real cause and effect links, not just links. This adds firm science and better fit for tough biology. For one, mixing clear AI with cause tools raises trust in poison forecasts and target finds. It aids rules follow-through.(95–97)

❖ Multimodal AI Pipelines Integrating Omics, Imaging, and Literature

Next steps in AI drug work stress blending varied data. This covers gene sets, cell messages, protein info (all gene data), body scans, and paper digs. Mixed AI setups that check these together reveal deep biology ties. One-type data misses these. Take gene data plus cell scans and paper facts. They sharpen trait details, check targets, and rank picks. This speeds full system grasp. It helps craft better compounds with molecule views, cell acts, and patient factors.(98–100)

❖ Stronger Human–AI Collaboration and Training Programs

Human skills and knowledge remain vital in drug production. It aids idea starts, process grasps, and right choices. Ahead lies AI built for smooth work with scientists. It gives clear ideas, hands-on tweaks, and reply loops. With these tech steps, set training raises AI skills in workers. This builds good match between people and AI. It lifts choice quality. It speeds drug flows by using both sides' strong points.(101,102)

6. CONCLUSION:

Generative AI and machine learning change every step in the drug discovery process. They cover target identification and validation. They also include hit discovery, lead optimization, predictive toxicology, and clinical trial design. These tools help search huge chemical areas. They speed up creating new molecules. They predict drug properties with better accuracy. In this way, they fix many problems in old drug discovery ways. Early wins show AI can cut development time. It lowers costs too. It raises chances of success in trials. This comes from better, data-based choices.

Still, big hurdles exist. Data quality, access, and standards hold back AI models. Models lack broad use and strength. The hidden inner workings of deep learning create doubt. Regulators and doctors find it hard to trust. Clear AI methods must improve. Ethics matter for data privacy, openness, and fair access to treatments. These need constant focus. Adding AI to current pharma work requires teamwork across fields. Staff need tech skills and mindset changes.

The future looks bright with ongoing growth in these technologies. Clear AI, cause-effect analysis, mixed data types, and human-AI teamwork will help beat limits. Better explanations and rules will build trust in AI finds. In the end, steady new ideas and joint work from industry, schools, and regulators will unlock AI's full power. This means safer, quicker, cheaper drugs to meet global health gaps.

7. Abbreviation:

AI: Artificial Intelligence

ML: Machine Learning

DL: Deep Learning

GenAI: Generative Artificial Intelligence

GANs: Generative Adversarial Networks

VAEs: Variational Autoencoders

LLMs: Large Language Models

QSAR: Quantitative Structure-Activity Relationship

SVMs: Support Vector Machines

RF: Random Forest

GNNs: Graph Neural Networks

XAI: Explainable Artificial Intelligence

ADMET: Absorption, Distribution, Metabolism, Excretion, Toxicity

RL: Reinforcement Learning

SAR: Structure-Activity Relationship

SPR: Structure-Property Relationship

NLP: Natural Language Processing

HTS: High Throughput Screening

MLP: Multi-Layer Perceptron

CNNs: Convolutional Neural Networks

RNNs: Recurrent Neural Networks

DNNs: Deep Neural Networks

LSTM: Long Short-Term Memory

BiLSTM: Bidirectional LSTM

MCC: Matthews' Correlation Coefficient

MHC-I: Major Histocompatibility Complex Class I

CPPs: Cell-Penetrating Peptides

AMPs: Antimicrobial Peptides

AUC: Area Under the Curve

AUROC: Area Under the Receiver Operating Characteristic Curve

BERT: Bidirectional Encoder Representations from Transformers

CPANNs: Counter-Propagation Artificial Neural Networks

ECM: Extracellular Matrix

REFERENCES

1. Dara S, Dhamercherla S, Jadav SS, Babu CM, Ahsan MJ. Machine Learning in Drug Discovery: A Review. Artif Intell Rev [Internet]. 2021 Mar 1 [cited 2025 Oct 5];55(3):1947. Available from:

https://pmc.ncbi.nlm.nih.gov/articles/PMC8356896/

2. Ferreira FJN, Carneiro AS. AI-Driven Drug Discovery: A Comprehensive Review. ACS Omega [Internet]. 2025 Jun 17 [cited 2025 Oct 5];10(23):23889. Available from: https://pmc.ncbi.nlm.nih.gov/articles/PMC12177741/
3. Blanco-González A, Cabezón A, Seco-González A, Conde-Torres D, Antelo-Riveiro P, Piñeiro Á, et al. The Role of AI in Drug Discovery: Challenges, Opportunities, and Strategies. Pharmaceuticals [Internet]. 2023 Jun 1 [cited 2025 Oct 5];16(6):891. Available from: https://pmc.ncbi.nlm.nih.gov/articles/PMC10302890/
4. Vora LK, Gholap AD, Jetha K, Thakur RRS, Solanki HK, Chavda VP. Artificial Intelligence in Pharmaceutical Technology and Drug Delivery Design. Pharmaceutics [Internet]. 2023 Jul 1 [cited 2025 Oct 5];15(7):1916. Available from: https://pmc.ncbi.nlm.nih.gov/articles/PMC10385763/
5. Gangwal A, Ansari A, Ahmad I, Azad AK, Kumarasamy V, Subramaniyan V, et al. Generative artificial intelligence in drug discovery: basic framework, recent advances, challenges, and opportunities. Front Pharmacol. 2024 Feb 7;15:1331062.
6. Qadri YA, Shaikh S, Ahmad K, Choi I, Kim SW, Vasilakos A V. Explainable Artificial Intelligence: A Perspective on Drug Discovery. Pharmaceutics [Internet]. 2025 Sep 1 [cited 2025 Oct 7];17(9). Available from: https://pubmed.ncbi.nlm.nih.gov/41012456/
7. Ocana A, Pandiella A, Privat C, Bravo I, Luengo-Oroz M, Amir E, et al. Integrating artificial intelligence in drug discovery and early drug development: a transformative approach. Biomark Res [Internet]. 2025 Dec 1 [cited 2025 Oct 7];13(1). Available from: https://pubmed.ncbi.nlm.nih.gov/40087789/
8. Paul D, Sanap G, Shenoy S, Kalyane D, Kalia K, Tekade RK. Artificial intelligence in drug discovery and development. Drug Discov Today [Internet]. 2021 Jan 1 [cited 2025 Oct 7];26(1):80–93. Available from: https://pubmed.ncbi.nlm.nih.gov/33099022/
9. Visan AI, Negut I. Integrating Artificial Intelligence for Drug Discovery in the Context of Revolutionizing Drug Delivery. Life (Basel) [Internet]. 2024 Feb 1 [cited 2025 Oct 7];14(2). Available from: https://pubmed.ncbi.nlm.nih.gov/38398742/
10. Liu X, Zhang J, Wang X, Teng M, Wang G, Zhou X. Application of artificial intelligence large language models in drug target discovery. Front Pharmacol [Internet]. 2025 [cited 2025 Oct 7];16. Available from: https://pubmed.ncbi.nlm.nih.gov/40697660/
11. Niazi SK. Artificial Intelligence in Small-Molecule Drug Discovery: A Critical Review of Methods, Applications, and Real-World Outcomes. Pharmaceuticals (Basel) [Internet]. 2025 Sep 1 [cited 2025 Oct 7];18(9). Available from: https://pubmed.ncbi.nlm.nih.gov/41011141/
12. Tang X, Dai H, Knight E, Wu F, Li Y, Li T, et al. A survey of generative AI for de novo drug design: new frontiers in molecule and protein generation. Brief Bioinform [Internet]. 2024 Jul 1 [cited 2025 Oct 7];25(4). Available from: https://pubmed.ncbi.nlm.nih.gov/39007594/
13. Yao R, Shen Z, Xu X, Ling G, Xiang R, Song T, et al. Knowledge mapping of graph neural networks for drug discovery: a bibliometric and visualized analysis. Front Pharmacol [Internet]. 2024 [cited 2025 Oct 8];15:1393415. Available from: https://pmc.ncbi.nlm.nih.gov/articles/PMC11116974/

14. Zhang O, Lin H, Zhang X, Wang X, Wu Z, Ye Q, et al. Graph Neural Networks in Modern AI-Aided Drug Discovery. Chem Rev [Internet]. 2025 Sep 17 [cited 2025 Oct 8]; Available from: https://pubmed.ncbi.nlm.nih.gov/40959983/
15. Dandibhotla S, Samudrala M, Kaneriya A, Dakshanamurthy S. GNNSeq: A Sequence-Based Graph Neural Network for Predicting Protein–Ligand Binding Affinity. Pharmaceuticals [Internet]. 2025 Mar 1 [cited 2025 Oct 8];18(3):329. Available from: https://pmc.ncbi.nlm.nih.gov/articles/PMC11945123/
16. Abbas K, Hao C, Yong X, Hasan MK, Islam S, Rahman AHA. Graph neural network-based drug-drug interaction prediction. Sci Rep [Internet]. 2025 Dec 1 [cited 2025 Oct 8];15(1):30340. Available from: https://pmc.ncbi.nlm.nih.gov/articles/PMC12365209/
17. Chakraborty C, Bhattacharya M, Pal S, Islam MdA. Generative AI in drug discovery and development: the next revolution of drug discovery and development would be directed by generative AI. Ann Med Surg (Lond) [Internet]. 2024 Oct [cited 2025 Oct 8];86(10):6340–3. Available from: https://pubmed.ncbi.nlm.nih.gov/39359753/
18. Poddiakov I, Umerenkov D, Shulcheva I, Golovina V, Borisova V, Pozdnyakova-Filatova I, et al. An iterative strategy to design 4-1BB agonist nanobodies de novo with generative AI models. Sci Rep [Internet]. 2025 Dec 1 [cited 2025 Oct 8];15(1). Available from: https://pubmed.ncbi.nlm.nih.gov/40659740/
19. Mouchlis VD, Afantitis A, Serra A, Fratello M, Papadiamantis AG, Aidinis V, et al. Advances in de Novo Drug Design: From Conventional to Machine Learning Methods. Int J Mol Sci [Internet]. 2021 Feb 2 [cited 2025 Oct 8];22(4):1–22. Available from: https://pubmed.ncbi.nlm.nih.gov/33562347/
20. Kumar P, Mahor A, Tomar R. Recent Development, Applications, and Patents of Artificial Intelligence in Drug Design and Development. Curr Drug Discov Technol [Internet]. 2025 Feb 11 [cited 2025 Oct 8];22(4). Available from: https://pubmed.ncbi.nlm.nih.gov/39931986/
21. Vidiyala N, Sunkishala P, Parupathi P, Nyavanandi D. The Role of Artificial Intelligence in Drug Discovery and Pharmaceutical Development: A Paradigm Shift in the History of Pharmaceutical Industries. AAPS PharmSciTech [Internet]. 2025 Jun 1 [cited 2025 Oct 8];26(5). Available from: https://pubmed.ncbi.nlm.nih.gov/40360908/
22. Bordukova M, Makarov N, Rodriguez-Esteban R, Schmich F, Menden MP. Generative artificial intelligence empowers digital twins in drug discovery and clinical trials. Expert Opin Drug Discov [Internet]. 2024 [cited 2025 Oct 8];19(1):33–42. Available from: https://pubmed.ncbi.nlm.nih.gov/37887266/
23. Khater T, Alkhatib SA, AlShehhi A, Pitsalidis C, Pappa AM, Ngo ST, et al. Generative artificial intelligence based models optimization towards molecule design enhancement. J Cheminform [Internet]. 2025 Aug 4 [cited 2025 Oct 8];17(1). Available from: https://pubmed.ncbi.nlm.nih.gov/40759950/
24. Lu J, Choi K, Eremeev M, Gobburu J, Goswami S, Liu Q, et al. Large Language Models and Their Applications in Drug Discovery and Development: A Primer. Clin Transl Sci [Internet]. 2025 Apr 1 [cited 2025 Oct 8];18(4). Available from: https://pubmed.ncbi.nlm.nih.gov/40208836/

25. Chakraborty C, Bhattacharya M, Pal S, Chatterjee S, Das A, Lee SS. AI-enabled language models (LMs) to large language models (LLMs) and multimodal large language models (MLLMs) in drug discovery and development. J Adv Res [Internet]. 2025 [cited 2025 Oct 8]; Available from: https://pubmed.ncbi.nlm.nih.gov/39952319/
26. Liu X, Zhang J, Wang X, Teng M, Wang G, Zhou X. Application of artificial intelligence large language models in drug target discovery. Front Pharmacol [Internet]. 2025 [cited 2025 Oct 8];16:1597351. Available from: https://pmc.ncbi.nlm.nih.gov/articles/PMC12279696/
27. Varadi M, Bertoni D, Magana P, Paramval U, Pidruchna I, Radhakrishnan M, et al. AlphaFold Protein Structure Database in 2024: providing structure coverage for over 214 million protein sequences. Nucleic Acids Res [Internet]. 2024 Jan 5 [cited 2025 Oct 8];52(D1):D368–75. Available from: https://pubmed.ncbi.nlm.nih.gov/37933859/
28. Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, et al. Highly accurate protein structure prediction with AlphaFold. Nature [Internet]. 2021 Aug 26 [cited 2025 Oct 8];596(7873):583–9. Available from: https://pubmed.ncbi.nlm.nih.gov/34265844/
29. Ali S, Chourasia P, Patterson M. When Protein Structure Embedding Meets Large Language Models. Genes (Basel) [Internet]. 2023 Jan 1 [cited 2025 Oct 8];15(1):25. Available from: https://pmc.ncbi.nlm.nih.gov/articles/PMC10815811/
30. Ding Q, Yao R, Bai Y, Da L, Wang Y, Xiang R, et al. Explainable Artificial Intelligence in the Field of Drug Research. Drug Des Devel Ther [Internet]. 2025 [cited 2025 Oct 8];19:4501–16. Available from: https://pubmed.ncbi.nlm.nih.gov/40458811/
31. Kırboğa KK, Abbasi S, Küçüksille EU. Explainability and white box in drug discovery. Chem Biol Drug Des [Internet]. 2023 Jul 1 [cited 2025 Oct 8];102(1):217–33. Available from: https://pubmed.ncbi.nlm.nih.gov/37105727/
32. Sturm H, Teufel J, Isfeld KA, Friederich P, Davis RL. Mitigating Molecular Aggregation in Drug Discovery With Predictive Insights From Explainable AI. Angew Chem Int Ed Engl [Internet]. 2025 Jul 14 [cited 2025 Oct 8];64(29):e202503259. Available from: https://pmc.ncbi.nlm.nih.gov/articles/PMC12258689/
33. Wong F, Zheng EJ, Valeri JA, Donghia NM, Anahtar MN, Omori S, et al. Discovery of a structural class of antibiotics with explainable deep learning. Nature [Internet]. 2024 Feb 1 [cited 2025 Oct 8];626(7997):177–85. Available from: https://pubmed.ncbi.nlm.nih.gov/38123686/
34. Kim H, Kim E, Lee I, Bae B, Park M, Nam H. Artificial Intelligence in Drug Discovery: A Comprehensive Review of Data-driven and Machine Learning Approaches. Biotechnol Bioprocess Eng [Internet]. 2020 Dec 1 [cited 2025 Oct 8];25(6):895–930. Available from: https://pubmed.ncbi.nlm.nih.gov/33437151/
35. Najm M, Azencott CA, Playe B, Stoven V. Drug Target Identification with Machine Learning: How to Choose Negative Examples. Int J Mol Sci [Internet]. 2021 May 2 [cited 2025 Oct 9];22(10). Available from: https://pubmed.ncbi.nlm.nih.gov/34066072/
36. Wenteler A, Cabrera CP, Wei W, Neduva V, Barnes MR. AI approaches for the discovery and validation of drug targets. Cambridge prisms Precision medicine [Internet]. 2024 [cited 2025 Oct 9];2. Available from: https://pubmed.ncbi.nlm.nih.gov/39258224/

37. Lin M, Guo J, Gu Z, Tang W, Tao H, You S, et al. Machine learning and multi-omics integration: advancing cardiovascular translational research and clinical practice. J Transl Med [Internet]. 2025 Dec 1 [cited 2025 Oct 9];23(1):388. Available from: https://pmc.ncbi.nlm.nih.gov/articles/PMC11966820/
38. Du P, Fan R, Zhang N, Wu C, Zhang Y. Advances in Integrated Multi-omics Analysis for Drug-Target Identification. Biomolecules [Internet]. 2024 Jun 1 [cited 2025 Oct 9];14(6). Available from: https://pubmed.ncbi.nlm.nih.gov/38927095/
39. Bian Y, Kwon JJ, Liu C, Margiotta E, Shekhar M, Gould AE. Target-driven machine learning-enabled virtual screening (TAME-VS) platform for early-stage hit identification. Front Mol Biosci [Internet]. 2023 [cited 2025 Oct 9];10. Available from: https://pubmed.ncbi.nlm.nih.gov/36994428/
40. Hao Y, Cheng C, Li J, Li H, Di X, Zeng X, et al. Multimodal Integration in Health Care: Development With Applications in Disease Management. J Med Internet Res [Internet]. 2025 Aug 21 [cited 2025 Oct 9];27. Available from: https://pubmed.ncbi.nlm.nih.gov/40840463/
41. Graff DE, Shakhnovich EI, Coley CW. Accelerating high-throughput virtual screening through molecular pool-based active learning. Chem Sci [Internet]. 2021 Jun 14 [cited 2025 Oct 10];12(22):7866. Available from: https://pmc.ncbi.nlm.nih.gov/articles/PMC8188596/
42. Alves LA, Ferreira NC da S, Maricato V, Alberto AVP, Dias EA, Jose Aguiar Coelho N. Graph Neural Networks as a Potential Tool in Improving Virtual Screening Programs. Front Chem [Internet]. 2022 Jan 20 [cited 2025 Oct 10];9. Available from: https://pubmed.ncbi.nlm.nih.gov/35127645/
43. Hall BW, Tummino TA, Tang K, Irwin JJ, Shoichet BK. A database for large-scale docking and experimental results. bioRxiv [Internet]. 2025 Feb 27 [cited 2025 Oct 10];2025.02.25.639879. Available from: https://pmc.ncbi.nlm.nih.gov/articles/PMC11888352/
44. Kumar N, Acharya V. Advances in machine intelligence-driven virtual screening approaches for big-data. Med Res Rev [Internet]. 2024 May 1 [cited 2025 Oct 10];44(3):939–74. Available from: https://pubmed.ncbi.nlm.nih.gov/38129992/
45. Blanes-Mira C, Fernández-Aguado P, de Andrés-López J, Fernández-Carvajal A, Ferrer-Montiel A, Fernández-Ballester G. Comprehensive Survey of Consensus Docking for High-Throughput Virtual Screening. Molecules [Internet]. 2022 Jan 1 [cited 2025 Oct 10];28(1). Available from: https://pubmed.ncbi.nlm.nih.gov/36615367/
46. Liu Y, Moretti R, Wang Y, Dong H, Yan B, Bodenheimer B, et al. Advancements in Ligand-Based Virtual Screening through the Synergistic Integration of Graph Neural Networks and Expert-Crafted Descriptors. J Chem Inf Model [Internet]. 2025 May 26 [cited 2025 Oct 10];65(10):4898–905. Available from: https://pubmed.ncbi.nlm.nih.gov/40365985/
47. Thaingtamtanha T, Preto J, Gentile F. Pose ensemble graph neural networks to improve docking performances. Chem Sci [Internet]. 2025 [cited 2025 Oct 10]; Available from: https://pubmed.ncbi.nlm.nih.gov/41030805/
48. Han Y, He F, Chen Y, Qin W, Yu H, Xu D. Quality Assessment of Protein Docking Models Based on Graph Neural Network. Frontiers in Bioinformatics [Internet]. 2021 [cited 2025 Oct 10];1:693211. Available from:

https://pmc.ncbi.nlm.nih.gov/articles/PMC9581034/

49. Buccheri R, Rescifina A. High-Throughput, High-Quality: Benchmarking GNINA and AutoDock Vina for Precision Virtual Screening Workflow. Molecules [Internet]. 2025 Aug 1 [cited 2025 Oct 10];30(16). Available from: https://pubmed.ncbi.nlm.nih.gov/40871514/
50. Baskin II. The power of deep learning to ligand-based novel drug discovery. Expert Opin Drug Discov [Internet]. 2020 Jul 2 [cited 2025 Oct 11];15(7):755–64. Available from: https://pubmed.ncbi.nlm.nih.gov/32228116/
51. Lin E, Lin CH, Lane HY. De Novo Peptide and Protein Design Using Generative Adversarial Networks: An Update. J Chem Inf Model [Internet]. 2022 Feb 28 [cited 2025 Oct 18];62(4):761–74. Available from: https://pubmed.ncbi.nlm.nih.gov/35128926/
52. Martinelli DD. Generative machine learning for de novo drug discovery: A systematic review. Comput Biol Med [Internet]. 2022 Jun 1 [cited 2025 Oct 18];145. Available from: https://pubmed.ncbi.nlm.nih.gov/35339849/
53. Palazzesi F, Pozzan A. Deep Learning Applied to Ligand-Based De Novo Drug Design. Methods Mol Biol [Internet]. 2022 [cited 2025 Oct 18];2390:273–99. Available from: https://pubmed.ncbi.nlm.nih.gov/34731474/
54. Atz K, Cotos L, Isert C, Håkansson M, Focht D, Hilleke M, et al. Author Correction: Prospective de novo drug design with deep interactome learning. Nat Commun [Internet]. 2025 Dec 1 [cited 2025 Oct 19];16(1). Available from: https://pubmed.ncbi.nlm.nih.gov/39824793/
55. Lin E, Lin CH, Lane HY. Relevant Applications of Generative Adversarial Networks in Drug Design and Discovery: Molecular De Novo Design, Dimensionality Reduction, and De Novo Peptide and Protein Design. Molecules [Internet]. 2020 Jul 1 [cited 2025 Oct 19];25(14). Available from: https://pubmed.ncbi.nlm.nih.gov/32708785/
56. Zeng X, Wang F, Luo Y, Kang S gu, Tang J, Lightstone FC, et al. Deep generative molecular design reshapes drug discovery. Cell Rep Med [Internet]. 2022 Dec 20 [cited 2025 Oct 19];3(12). Available from: https://pubmed.ncbi.nlm.nih.gov/36306797/
57. Gaudelet T, Day B, Jamasb AR, Soman J, Regep C, Liu G, et al. Utilizing graph machine learning within drug discovery and development. Brief Bioinform [Internet]. 2021 Nov 1 [cited 2025 Oct 19];22(6). Available from: https://pubmed.ncbi.nlm.nih.gov/34013350/
58. Vanhaelen Q, Lin YC, Zhavoronkov A. The Advent of Generative Chemistry. ACS Med Chem Lett [Internet]. 2020 Aug 13 [cited 2025 Oct 19];11(8):1496–505. Available from: https://pubmed.ncbi.nlm.nih.gov/32832015/
59. Lim J, Hwang SY, Moon S, Kim S, Kim WY. Scaffold-based molecular design with a graph generative model. Chem Sci [Internet]. 2019 [cited 2025 Oct 19];11(4):1153–64. Available from: https://pubmed.ncbi.nlm.nih.gov/34084372/
60. Wu F, Zhou Y, Li L, Shen X, Chen G, Wang X, et al. Computational Approaches in Preclinical Studies on Drug Discovery and Development. Front Chem [Internet]. 2020 Sep 11 [cited 2025 Oct 19];8. Available from: https://pubmed.ncbi.nlm.nih.gov/33062633/
61. Chen Z, Min MR, Parthasarathy S, Ning X. A Deep Generative Model for Molecule Optimization via One Fragment Modification. Nat Mach Intell [Internet]. 2021 Dec 1 [cited 2025 Oct 19];3(12):1040–9. Available from: https://pubmed.ncbi.nlm.nih.gov/35187404/
62. Yi JC, Yang ZY, Zhao WT, Yang ZJ, Zhang XC, Wu CK, et al. ChemMORT: an automatic ADMET optimization platform using deep

learning and multi-objective particle swarm optimization. Brief Bioinform [Internet]. 2024 Mar 1 [cited 2025 Oct 19];25(2). Available from: https://pubmed.ncbi.nlm.nih.gov/38385872/

63. Venkataraman M, Rao GC, Madavareddi JK, Maddi SR. Leveraging machine learning models in evaluating ADMET properties for drug discovery and development. ADMET DMPK [Internet]. 2025 [cited 2025 Oct 20];13(3). Available from: https://pubmed.ncbi.nlm.nih.gov/40585410/
64. Du BX, Xu Y, Yiu SM, Yu H, Shi JY. ADMET property prediction via multi-task graph learning under adaptive auxiliary task selection. iScience [Internet]. 2023 Nov 17 [cited 2025 Oct 20];26(11). Available from: https://pubmed.ncbi.nlm.nih.gov/38026198/
65. Seal S, Mahale M, García-Ortegón M, Joshi CK, Hosseini-Gerami L, Beatson A, et al. Machine Learning for Toxicity Prediction Using Chemical Structures: Pillars for Success in the Real World. Chem Res Toxicol [Internet]. 2025 May 19 [cited 2025 Oct 20];38(5):759–807. Available from: https://pubs.acs.org/doi/pdf/10.1021/acs.chemrestox.5c00033?ref=article_openPDF
66. Abdelwahab AA, Elattar MA, Fawzi SA. Advancing ADMET prediction for major CYP450 isoforms: graph-based models, limitations, and future directions. Biomed Eng Online [Internet]. 2025 Dec 1 [cited 2025 Oct 20];24(1). Available from: https://pubmed.ncbi.nlm.nih.gov/40702475/
67. Guo W, Liu J, Dong F, Song M, Li Z, Khan MKH, et al. Review of machine learning and deep learning models for toxicity prediction. Exp Biol Med [Internet]. 2023 Nov 1 [cited 2025 Oct 20];248(21):1952–73. Available from: https://pubmed.ncbi.nlm.nih.gov/38057999/
68. Swanson K, Walther P, Leitz J, Mukherjee S, Wu JC, Shivnaraine R V., et al. ADMET-AI: a machine learning ADMET platform for evaluation of large-scale chemical libraries. Bioinformatics [Internet]. 2024 Jul 1 [cited 2025 Oct 20];40(7). Available from: https://pubmed.ncbi.nlm.nih.gov/38913862/
69. Zhang J, Li H, Zhang Y, Huang J, Ren L, Zhang C, et al. Computational toxicology in drug discovery: applications of artificial intelligence in ADMET and toxicity prediction. Brief Bioinform [Internet]. 2025 Aug 31 [cited 2025 Oct 20];26(5). Available from: https://pubmed.ncbi.nlm.nih.gov/41052279/
70. Mostafa F, Chen M. Computational models for predicting liver toxicity in the deep learning era. Frontiers in toxicology [Internet]. 2024 [cited 2025 Oct 20];5. Available from: https://pubmed.ncbi.nlm.nih.gov/38312894/
71. Romano JD, Hao Y, Moore JH. Improving QSAR Modeling for Predictive Toxicology using Publicly Aggregated Semantic Graph Data and Graph Neural Networks. Pacific Symposium on Biocomputing. 2022;187–98.
72. Matsunaga R, Tsumoto K. Accelerating antibody discovery and optimization with high-throughput experimentation and machine learning. J Biomed Sci [Internet]. 2025 Dec 1 [cited 2025 Oct 20];32(1). Available from: https://pubmed.ncbi.nlm.nih.gov/40346589/
73. Chungyoun MF, Gray JJ. AI Models for Protein Design are Driving Antibody Engineering. Curr Opin Biomed Eng [Internet]. 2023 Dec 1 [cited 2025 Oct 20];28. Available from: https://pubmed.ncbi.nlm.nih.gov/37484815/
74. Bielska W, Jaszczyszyn I, Dudzic P, Janusz B, Chomicz D, Wrobel S, et al. Applying computational protein design to therapeutic antibody discovery - current state and perspectives. Front Immunol [Internet]. 2025

[cited 2025 Oct 20];16. Available from: https://pubmed.ncbi.nlm.nih.gov/40475769/

75. Jin R, Zhou R, Zhang D. Recent advances in antibody optimization based on deep learning methods. J Zhejiang Univ Sci B [Internet]. 2025 May 1 [cited 2025 Oct 20];26(5):409–20. Available from: https://pubmed.ncbi.nlm.nih.gov/40436639/
76. Almeida DS, Almeida M V., Sampaio J V., Gaieta EM, Costa AHS, Rabelo FFA, et al. AbSet: A Standardized Data Set of Antibody Structures for Machine Learning Applications. J Chem Inf Model [Internet]. 2025 May 26 [cited 2025 Oct 20];65(10):4767–74. Available from: https://pubmed.ncbi.nlm.nih.gov/40349368/
77. Son A, Park J, Kim W, Lee W, Yoon Y, Ji J, et al. Integrating Computational Design and Experimental Approaches for Next-Generation Biologics. Biomolecules [Internet]. 2024 Sep 1 [cited 2025 Oct 20];14(9). Available from: https://pubmed.ncbi.nlm.nih.gov/39334841/
78. Waight AB, Prihoda D, Shrestha R, Metcalf K, Bailly M, Ancona M, et al. A machine learning strategy for the identification of key in silico descriptors and prediction models for IgG monoclonal antibody developability properties. MAbs [Internet]. 2023 [cited 2025 Oct 20];15(1). Available from: https://pubmed.ncbi.nlm.nih.gov/37610144/
79. Serrano DR, Luciano FC, Anaya BJ, Ongoren B, Kara A, Molina G, et al. Artificial Intelligence (AI) Applications in Drug Discovery and Drug Delivery: Revolutionizing Personalized Medicine. Pharmaceutics [Internet]. 2024 Oct 1 [cited 2025 Oct 20];16(10). Available from: https://pubmed.ncbi.nlm.nih.gov/39458657/
80. Kokudeva M, Vichev M, Naseva E, Miteva DG, Velikova T. Artificial intelligence as a tool in drug discovery and development. World J Exp Med [Internet]. 2024 Sep 20 [cited 2025 Oct 20];14(3). Available from: https://pubmed.ncbi.nlm.nih.gov/39312699/
81. Carini C, Seyhan AA. Tribulations and future opportunities for artificial intelligence in precision medicine. J Transl Med [Internet]. 2024 Dec 1 [cited 2025 Oct 20];22(1). Available from: https://pubmed.ncbi.nlm.nih.gov/38702711/
82. Dermawan D, Alotaiq N. From Lab to Clinic: How Artificial Intelligence (AI) Is Reshaping Drug Discovery Timelines and Industry Outcomes. Pharmaceuticals (Basel) [Internet]. 2025 Jul 1 [cited 2025 Oct 20];18(7). Available from: https://pubmed.ncbi.nlm.nih.gov/40732273/
83. Yoshiura H, Kawata Y, Sengoku S. Open Innovation and Regulatory Challenges in New Modality Development: The Pivotal Role of Contract Development and Manufacturing Organisations in Advancing Antibody Drugs. Ther Innov Regul Sci [Internet]. 2025 Jan 1 [cited 2025 Oct 20];59(1):63–70. Available from: https://pubmed.ncbi.nlm.nih.gov/39466520/
84. Cova T, Vitorino C, Ferreira M, Nunes S, Rondon-Villarreal P, Pais A. Artificial Intelligence and Quantum Computing as the Next Pharma Disruptors. Methods Mol Biol [Internet]. 2022 [cited 2025 Oct 20];2390:321–47. Available from: https://pubmed.ncbi.nlm.nih.gov/34731476/
85. Zheng J, Wang Y, Liang Q, Cui L, Wang L. The Application of Machine Learning on Antibody Discovery and Optimization. Molecules [Internet]. 2024 Dec 1 [cited 2025 Oct 20];29(24). Available from: https://pubmed.ncbi.nlm.nih.gov/39770013/
86. van den Broek RL, Patel S, van Westen GJP, Jespers W, Sherman W. In Search of Beautiful Molecules: A Perspective on Generative Modeling for Drug Design. J Chem Inf Model

[Internet]. 2025 Sep 22 [cited 2025 Oct 20];65(18):9383–97. Available from: https://pubmed.ncbi.nlm.nih.gov/40893044/

87. Askr H, Elgeldawi E, Aboul Ella H, Elshaier YAMM, Gomaa MM, Hassanien AE. Deep learning in drug discovery: an integrative review and future challenges. Artif Intell Rev [Internet]. 2023 Jul 1 [cited 2025 Oct 20];56(7):5975–6037. Available from: https://pubmed.ncbi.nlm.nih.gov/36415536/
88. Qi X, Zhao Y, Qi Z, Hou S, Chen J. Machine Learning Empowering Drug Discovery: Applications, Opportunities and Challenges. Molecules [Internet]. 2024 Feb 1 [cited 2025 Oct 20];29(4). Available from: https://pubmed.ncbi.nlm.nih.gov/38398653/
89. Alam F, Mohammed Alnazzawi TS, Mehmood R, Al-maghthawi A. A Review of the Applications, Benefits, and Challenges of Generative AI for Sustainable Toxicology. Curr Res Toxicol [Internet]. 2025 Jan 1 [cited 2025 Oct 20];8. Available from: https://pubmed.ncbi.nlm.nih.gov/40331045/
90. 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 [Internet]. 2019 Jun 1 [cited 2025 Oct 20];18(6):463–77. Available from: https://pubmed.ncbi.nlm.nih.gov/30976107/
91. Hasselgren C, Oprea TI. Artificial Intelligence for Drug Discovery: Are We There Yet? Annu Rev Pharmacol Toxicol [Internet]. 2024 Jan 23 [cited 2025 Oct 20];64:527–50. Available from: https://pubmed.ncbi.nlm.nih.gov/37738505/
92. Kumar M, Nguyen TPN, Kaur J, Singh TG, Soni D, Singh R, et al. Opportunities and challenges in application of artificial intelligence in pharmacology. Pharmacol Rep [Internet]. 2023 Feb 1 [cited 2025 Oct 20];75(1):3–18. Available from: https://pubmed.ncbi.nlm.nih.gov/36624355/
93. Bi X, Wang Y, Wang J, Liu C. Machine Learning for Multi-Target Drug Discovery: Challenges and Opportunities in Systems Pharmacology. Pharmaceutics [Internet]. 2025 Sep 1 [cited 2025 Oct 20];17(9). Available from: https://pubmed.ncbi.nlm.nih.gov/41012523/
94. Gangwal A, Lavecchia A. Unleashing the power of generative AI in drug discovery. Drug Discov Today [Internet]. 2024 Jun 1 [cited 2025 Oct 20];29(6). Available from: https://pubmed.ncbi.nlm.nih.gov/38663579/
95. Jusoh AS, Remli MA, Mohamad MS, Cazenave T, Fong CS. How generative Artificial Intelligence can transform drug discovery? Eur J Med Chem [Internet]. 2025 Oct 5 [cited 2025 Oct 20];295. Available from: https://pubmed.ncbi.nlm.nih.gov/40456205/
96. Fu C, Chen Q. The future of pharmaceuticals: Artificial intelligence in drug discovery and development. J Pharm Anal [Internet]. 2025 Aug 1 [cited 2025 Oct 20];15(8). Available from: https://pubmed.ncbi.nlm.nih.gov/40893437/
97. Holzinger A, Langs G, Denk H, Zatloukal K, Müller H. Causability and explainability of artificial intelligence in medicine. Wiley Interdiscip Rev Data Min Knowl Discov [Internet]. 2019 Jul 1 [cited 2025 Oct 20];9(4). Available from: https://pubmed.ncbi.nlm.nih.gov/32089788/
98. Zack M, Stupichev DN, Moore AJ, Slobodchikov ID, Sokolov DG, Trifonov IF, et al. Artificial Intelligence and Multi-Omics in Pharmacogenomics: A New Era of Precision Medicine. Mayo Clinic Proceedings: Digital Health [Internet]. 2025 Sep 1 [cited 2025 Oct 20];3(3). Available from: https://pubmed.ncbi.nlm.nih.gov/40881104/

99. Lang O, Yaya-Stupp D, Traynis I, Cole-Lewis H, Bennett CR, Lyles CR, et al. Using generative AI to investigate medical imagery models and datasets. EBioMedicine [Internet]. 2024 Apr 1 [cited 2025 Oct 20];102. Available from: https://pubmed.ncbi.nlm.nih.gov/38565004/
100. Zhang K, Yang X, Wang Y, Yu Y, Huang N, Li G, et al. Artificial intelligence in drug development. Nat Med [Internet]. 2025 Jan 1 [cited 2025 Oct 20];31(1):45–59. Available from: https://pubmed.ncbi.nlm.nih.gov/39833407/
101. Albani FG, Alghamdi SS, Almutairi MM, Alqahtani T. Artificial Intelligence-Driven Innovations in Oncology Drug Discovery: Transforming Traditional Pipelines and Enhancing Drug Design. Drug Des Devel Ther [Internet]. 2025 [cited 2025 Oct 20];19:5685–707. Available from: https://pubmed.ncbi.nlm.nih.gov/40626099/
102. El-Deiry WS, Bresson C, Wunder F, Carneiro BA, Dizon DS, Warner JL, et al. Worldwide Innovative Network (WIN) Consortium in Personalized Cancer Medicine: Bringing next-generation precision oncology to patients. Oncotarget [Internet]. 2025 [cited 2025 Oct 20];16(1):140–62. Available from: https://pubmed.ncbi.nlm.nih.gov/40073368/.

HOW TO CITE: Gaurav Palsamkar*, Aditya Gurav, Generative AI and Machine Learning in Drug Discovery, Int. J. of Pharm. Sci., 2025, Vol 3, Issue 11, 3061-3082 https://doi.org/10.5281/zenodo.17658072