Generative artificial intelligence in health sciences: A comprehensive review of emerging applications, recent advancements, challenges and future perspectives

Amit Chawla[*1], Nadeem[2], Jyoti Arora[3], Prabhjot Singh[4]
*[*1]Professor Cum Principal, Faculty of Pharmacy, Tantia University, Sri Ganganagar, Rajasthan, India*
[2]Research Scholar, Department of Public Health, St. Francis College, Brooklyn, New York, United States of America
[3]Vice Principal, Sri Ganganagar College of Nursing, Tantia University, Sri Ganganagar
[4]Deputy Director, Faculty of Nursing. Desh Bhagat University, Mandi Gobindgarh, Punjab, India

DOI: 10.5281/zenodo.23880499

Abstract:
Generative artificial intelligence (GenAI) is becoming an important computational approach in health sciences because it can create new molecular structures, protein sequences, scientific text, images and synthetic data from patterns learned in large datasets. Unlike conventional predictive artificial intelligence, which mainly estimates an outcome from existing inputs, generative systems can propose previously unobserved candidates and produce structured content for further testing. This review examines the principles, emerging applications, recent advances, limitations and future directions of GenAI across the pharmaceutical lifecycle. Major application areas include de novo small-molecule design, lead optimisation, protein and peptide engineering, target and mechanism research, drug repurposing, formulation development, preclinical assessment, clinical-trial operations, pharmaceutical manufacturing, quality systems and pharmacovigilance. Developments in variational autoencoders, generative adversarial networks, transformers, large language models and diffusion models have expanded the range of molecules and biological structures that can be explored computationally. Protein-structure tools such as AlphaFold 3 and diffusion-based design systems illustrate the growing integration of generative modelling with structural biology. Nevertheless, generated candidates are hypotheses rather than medicines: they require chemical synthesis, analytical characterisation, biological testing, toxicological assessment and clinical evaluation. Data quality, model hallucination, limited external validation, bias, interpretability, intellectual-property uncertainty, privacy and regulatory accountability remain major concerns. The review concludes that the most credible near-term model is human-supervised, experimentally grounded use of GenAI within a reproducible design–make–test–analyse workflow. GenAI is likely to strengthen pharmaceutical research productivity and broaden exploration of chemical and biological space, but it cannot replace experimental science, professional judgement or regulatory standards.

Keywords:
Generative artificial intelligence; pharmaceutical sciences; drug discovery; molecular design; large language models; protein design; pharmaceutical formulation; pharmacovigilance; drug development.

---

1. Introduction:
Health Science research is a long, multidisciplinary process involving the identification of a therapeutic need, selection of a biological target, discovery and optimisation of a candidate, formulation, preclinical testing, clinical development, manufacturing and post-marketing safety monitoring. Each stage produces complex data, including chemical structures, protein sequences, omics profiles, laboratory observations, imaging, clinical records and regulatory documents. Conventional computational methods have improved many of these activities, but the scale and complexity of modern biomedical data continue to challenge researchers.

Generative artificial intelligence (GenAI) offers a new way to explore these data by learning patterns and producing candidate outputs that can be assessed experimentally.

The term GenAI refers to computational models designed to generate new content that resembles, or is constrained by, patterns in training data. In pharmaceutical sciences, this content may be a molecular graph, a SMILES string, a three-dimensional ligand pose, a protein sequence, a protein backbone, a formulation proposal, a protocol draft or a structured summary of scientific evidence. The generated output is not automatically correct or useful. It is a proposal that must be assessed against scientific, clinical, manufacturing and regulatory requirements. Early reviews described the transition from conventional computer-aided molecular design towards deep generative approaches that can propose molecules with selected properties (Meyers et al., 2021; Tong et al., 2021). More recent reviews describe a broader role for GenAI across research and development, including scientific productivity and hypothesis generation (Gangwal et al., 2024; Doron et al., 2025).

The rapid growth of GenAI has also created a need to separate demonstrated capability from expectation. A model may generate a chemically valid structure without producing a potent drug; it may predict a protein fold without proving that the protein functions in a biological system; and it may draft a clinical document without ensuring that the content is accurate or compliant. Therefore, the value of GenAI should be judged by experimentally verified improvements in quality, speed, reproducibility or decision-making, rather than by the novelty of generated outputs alone.

2. Scope and review approach:

This article is a narrative review of published research and authoritative guidance relevant to GenAI in pharmaceutical sciences. The literature considered includes foundational studies of molecular generation, reviews of drug and protein design, studies of structure-prediction and diffusion-based methods, and publications addressing pharmaceutical research, medicine, formulation, pharmacovigilance and responsible AI. Regulatory and governance sources from the US Food and Drug Administration (FDA), European Medicines Agency (EMA), World Health Organization (WHO) and National Institute of Standards and Technology (NIST) were also considered. The discussion reflects literature and guidance available up to 31 October 2026. It is not a systematic review or meta-analysis; no pooled effect estimate or formal risk-of-bias assessment is claimed.

The review focuses on generative methods and on conventional AI methods only where they form part of a GenAI-supported workflow. For example, a predictive model that estimates solubility is not itself generative, but it can act as a scoring function to filter molecules proposed by a generative model. This distinction is important because the pharmaceutical literature often uses the wider term artificial intelligence to describe predictive analytics, automation and generative systems together (Huanbutta et al., 2024; Joshi and Sheth, 2025).

3. Foundations and major generative model architectures:

GenAI systems learn a representation of patterns in training data and sample from that learned representation to create new outputs. The data may be represented as molecular strings, molecular graphs, three-dimensional atomic coordinates, amino-acid sequences, text or linked multimodal records. Model performance depends strongly on the quality, coverage, consistency and relevance of these data. In pharmaceutical applications, a model may be pretrained on large general datasets and then adapted to a narrower task using domain-specific data, fine-tuning, retrieval-augmented generation or preference-based optimisation.

3.1. Variational autoencoders:

Variational autoencoders (VAEs) encode an input molecule into a compressed latent representation and decode a sampled point from that representation into a molecular output. By moving through latent space, researchers can explore structural variations and attempt to optimise desired properties. VAEs can support scaffold exploration and multi-parameter optimisation, but generated outputs may be chemically invalid, synthetically difficult or outside the model's learned chemical domain. Their practical value therefore depends on robust validity checks and experimental confirmation (Tong et al., 2021; Gangwal and Lavecchia, 2024).

3.2. Generative adversarial networks:

Generative adversarial networks (GANs) contain a generator that proposes samples and a discriminator that learns to distinguish generated samples from real training examples. During training, both components improve through competition. GAN-based approaches have been explored for generating molecular structures, peptides, proteins and biological data. They can model complex distributions, although training instability,

mode collapse and difficulty in enforcing precise chemical constraints can limit use. Reviews of GAN applications in drug discovery emphasise their potential but also the need for more realistic benchmarks and experimental validation (Tripathi et al., 2022).

3.3. Autoregressive transformers and large language models:

Transformers generate sequences by predicting the next token or sequence element from preceding context. In chemistry, tokens can represent atoms, bonds, ring closures or fragments in a molecular string; in protein modelling, they can represent amino acids; and in scientific assistants, they represent words or structured information. Large language models (LLMs) can summarise literature, draft technical documents, extract information and support conversational access to scientific databases. Their strength in language does not guarantee knowledge of molecular energetics, experimental conditions or clinical context. They can produce plausible but incorrect statements, making retrieval of source evidence and expert verification essential (Teo et al., 2025; Shanmugam et al., 2025).

3.4. Diffusion and score-based models:

Diffusion models learn to reverse a gradual process in which structured data are corrupted by noise. Starting from noise, the trained model iteratively generates a plausible sample. In molecular science, diffusion approaches can operate on molecular coordinates, conformations or protein backbones. They are attractive for three-dimensional design because they can represent spatial relationships and can be conditioned on a target pocket, motif or desired geometry. RFdiffusion demonstrated the ability to design protein structures and functional motifs, while AlphaFold 3 uses a diffusion-based architecture to predict complexes involving proteins, nucleic acids, small molecules, ions and modified residues (Watson et al., 2023; Abramson et al., 2024). Predicted structures still require experimental confirmation and should not be interpreted as proof of binding or function.

Table. 1: Comparison of common model families

Model family	Typical pharmaceutical use	Important limitation
VAE	Latent-space exploration and molecular variation	Validity and property control may be imperfect
GAN	Molecule, peptide or protein sample generation	Training instability and mode collapse
Transformer / LLM	Molecular sequence generation; literature and document support	Hallucinations, context limits and prompt sensitivity
Diffusion model	3D ligand conformations and protein-backbone design	Computational cost and need for structural/functional testing
Multimodal / agentic systems	Combining text, structures, assays and workflow tools	Complex error propagation and difficult oversight

4. Emerging applications across pharmaceutical sciences:

4.1. Target identification and biological hypothesis generation:

GenAI can assist researchers in connecting disease phenotypes, genes, proteins, pathways, chemical entities and published findings. LLMs can extract hypotheses from scientific articles, organise evidence by target and pathway, and generate structured summaries for expert review. Multimodal systems may combine sequence, structure, expression and assay information to suggest relationships that are difficult to identify through manual literature review. Retrieval-augmented generation (RAG) can improve traceability by requiring a model to draw on a defined collection of approved documents or databases. The generated hypothesis must still be checked for direction of effect, biological plausibility, tissue relevance and independent replication. A generated target association is a starting point for experiments, not proof of causality.

4.2. De novo small-molecule design and lead optimization:

The most established pharmaceutical use of generative modelling is the proposal of novel small molecules. Models can generate structures that satisfy constraints such as molecular size, charge, scaffold, functional groups or predicted interaction with a target. A typical workflow begins with target and binding-site characterisation, followed by generation of candidate structures, computational filtering and selection for synthesis. Candidate ranking may combine predicted affinity, selectivity, solubility, permeability, metabolic stability, toxicity alerts, novelty and synthetic accessibility. Generative systems can also modify an existing

lead to explore analogues while preserving a pharmacophore or selected interactions. Reviews by (Tong et al., 2021), (Meyers et al., 2021), (Gangwal et al., 2024) and (Tang et al., 2024) describe the evolution of these methods from sequence-based generation towards graph-based, target-conditioned and three-dimensional approaches.

A key challenge is multi-objective optimisation. Improving predicted target affinity may worsen solubility, metabolic stability or off-target activity. Reward functions can encode these competing goals, but the results depend on the quality and calibration of the component predictors. A model may learn to exploit weaknesses in a scoring function and generate structures that score highly in silico but fail in the laboratory. Accordingly, medicinal chemists should review proposed structures for chemical reactivity, stereochemistry, intellectual-property constraints, route feasibility and fit with the intended mechanism. (Du et al., 2024) describe how generative molecular design can integrate generation and filtering, while newer target-specific systems seek to condition the output on the selected protein rather than generate compounds with generic drug-like features alone.

4.3. Protein, peptide and antibody engineering:

Protein therapeutics, enzymes, peptides and antibodies occupy a growing part of the pharmaceutical landscape. Generative protein-design systems can propose amino-acid sequences or three-dimensional backbones with specified structural motifs, binding surfaces or functional objectives. RFdiffusion demonstrated de novo protein backbone design and motif scaffolding, illustrating how diffusion models can generate structures that would be difficult to explore using conventional trial-and-error approaches (Watson et al., 2023). AlphaFold 3 expanded structure prediction to a range of biomolecular interactions, including complexes with small molecules and nucleic acids (Abramson et al., 2024). These tools are complementary: structure prediction estimates plausible configurations, whereas a generative design model proposes new sequences or structures.

Potential applications include designing enzyme replacements, stabilising therapeutic proteins, creating peptide binders, improving antibody binding interfaces and exploring protein scaffolds. Protein design remains a demanding experimental discipline. A sequence that appears structurally plausible may express poorly, aggregate, degrade, trigger an immune response or lack the required biological activity. Koh et al. (2025) describe an emerging workflow in which AI design tools are combined with synthesis and experimental validation. The practical lesson is that protein design should use iterative laboratory feedback rather than rely solely on computational confidence scores.

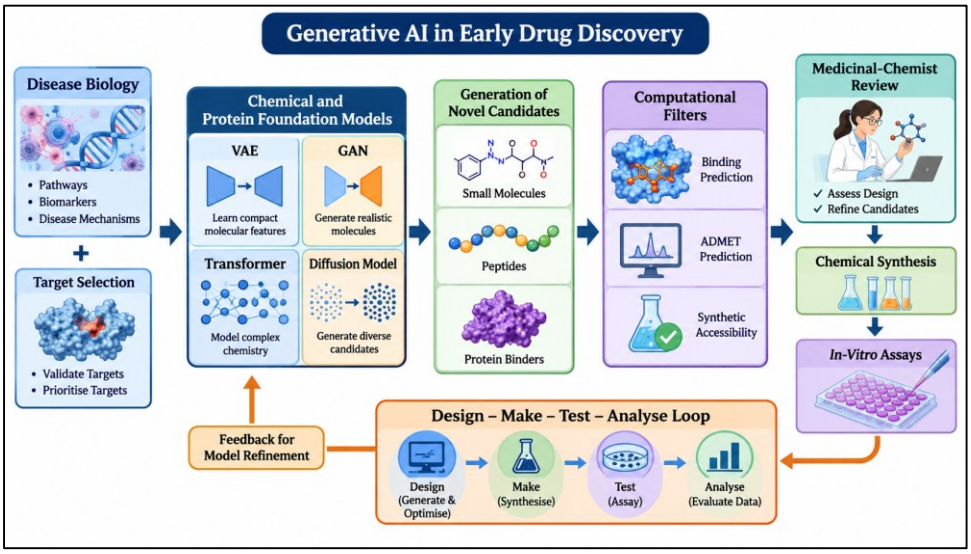


Figure. 1: Generative AI in Early Drug Discovery

4.4 Drug repurposing and natural-product research:

GenAI may support drug repurposing by organising evidence about drug targets, disease pathways, adverse effects and molecular similarity. An LLM can help summarise prior reports or formulate testable hypotheses, while generative models may propose analogues of known active compounds or explore derivatives of natural-product scaffolds. In natural-product research, models can assist with structure representation, analogue generation and prioritisation for isolation or synthesis. However, generated similarity does not establish a shared mechanism, and text-based associations may reflect publication bias. Repurposing hypotheses need

confirmation through target-engagement studies, pharmacological assays, exposure analysis and appropriately designed clinical studies.

4.5. Pharmaceutical formulation and drug delivery:

Formulation development involves selecting active ingredients, excipients, proportions, process conditions and packaging to obtain a stable, manufacturable and effective dosage form. GenAI could propose candidate compositions, generate experimental plans and help explore combinations that are not obvious from conventional design spaces. A language model can structure formulation knowledge from laboratory notebooks and published studies, while a generative model may suggest compositions or process recipes. Predictive models for dissolution, stability, viscosity or release kinetics can then score proposed candidates. It is important to recognise that much of the current formulation literature concerns predictive AI rather than generative AI specifically. Huanbutta et al. (2024) and Joshi and Sheth (2025) describe broader AI applications in formulation and dosage calculations; the direct evidence for fully generative, experimentally validated formulation design remains more limited.

Possible future applications include personalised dosage forms, small-batch 3D-printed medicines, taste-masked formulations, controlled-release systems and excipient selection for patients with special requirements. GenAI could also help generate design-of-experiment plans that prioritise informative experiments and reduce unnecessary combinations. Nevertheless, each proposed formulation requires compatibility testing, stability studies, dissolution testing, content uniformity assessment, and microbial evaluation where relevant and process validation. A model-generated recipe cannot substitute for pharmacopeial testing or good manufacturing practice.

4.6. Preclinical pharmacology, toxicology and ADMET assessment:

Absorption, distribution, metabolism, excretion and toxicity (ADMET) properties are essential to candidate selection. Generative models can be conditioned on desired property ranges, while predictive models estimate specific endpoints such as permeability, metabolic stability, clearance or toxicity. GenAI can also help organise toxicology reports and identify missing evidence. A particularly important risk is that a generated compound may appear favourable because the training data under-represent certain chemical classes, assay conditions or toxicological mechanisms. Computational alerts should be treated as prioritisation tools, not definitive safety conclusions. Candidates need appropriate in-vitro, in-vivo and mechanistic assessment, with attention to species differences and the relevance of the test system to human exposure.

4.7. Clinical research and trial operations:

Clinical development generates large volumes of protocols, investigator brochures, consent materials, case-report forms, statistical analysis plans, monitoring reports and safety narratives. LLMs can support drafting, summarisation, consistency checks, and extraction of eligibility criteria and translation of technical information into patient-friendly language. GenAI may also help researchers generate synthetic datasets for software testing or explore recruitment strategies. Teo et al. (2025) and Shanmugam et al. (2025) describe wider biomedical applications of GenAI, including support for research design and complex multistage tasks.

The use of GenAI in clinical trials requires strict controls. Synthetic data must not be misrepresented as observations from real participants, and generated eligibility criteria must be checked against the protocol and clinical rationale. Patient-facing content requires review for readability, cultural suitability and accuracy. Models should not independently make eligibility, treatment or safety decisions. Any use that affects trial conduct, endpoint determination or regulatory evidence should be validated for the specific context, documented and governed through version control and change management.

4.8. Manufacturing, quality assurance and process development:

In pharmaceutical manufacturing, GenAI may help interpret deviation reports, prepare draft investigation summaries, retrieve standard operating procedures, generate training scenarios and support process-development documentation. It may also be used with digital twins and predictive process models to explore operating conditions or create synthetic data for system testing. These capabilities should be distinguished from established process analytical technology and conventional machine-learning models that monitor process parameters. The latter may detect unusual patterns or predict quality attributes, while a generative system produces candidate text, scenarios or process hypotheses. Huanbutta et al. (2024) review AI across manufacturing and quality control, but direct evidence that GenAI can autonomously manage a validated pharmaceutical process is not established.

Quality-critical activities require attributable records, audit trails, data integrity, validated software and human approval. A generated deviation report could omit a critical observation or invent a plausible explanation. Consequently, outputs should be linked to source records, checked by qualified personnel and retained in accordance with applicable quality-system requirements. GenAI can reduce administrative burden, but it should not approve batches, release products or replace investigation responsibility unless a formally validated and legally acceptable system has been established.

4.9. Pharmacovigilance and medication safety:

Pharmacovigilance depends on the extraction and interpretation of safety information from spontaneous reports, clinical records, medical literature, social media and regulatory documents. LLMs can assist with adverse drug reaction (ADR) entity extraction, case summarisation, coding support, literature screening and preparation of safety narratives. A 2026 systematic review by Kim et al. reported that the evaluated LLM studies were concentrated in constrained tasks such as information extraction, classification, signal evaluation and documentation. The authors also identified hallucination, low specificity, prompt sensitivity and weak external validation as recurring limitations. This supports the use of LLMs as supervised workflow aids rather than autonomous safety decision-makers.

In practical settings, a model may help identify a possible drug-event relationship or organise a case, but it cannot independently establish causality. Safety professionals must assess temporal relationships, alternative causes, dechallenge and rechallenge information, seriousness, expectedness and completeness. Any automated triage must be monitored for missed cases and disproportionate performance across populations. GenAI can improve the speed of handling unstructured information only when it preserves source traceability and allows expert review.

5. Recent advancements and emerging technical directions:

5.1. From single-task models to foundation models:

A major development is the movement from models trained for one narrow task towards foundation models pretrained on large collections of chemical, biological, clinical or scientific information. These models can be adapted for multiple downstream tasks, potentially reducing the need to train a new system for every application. Chemical language models can learn common substructures and sequence patterns; protein models can capture relationships between amino-acid sequences and structure; and multimodal systems can connect text, images, molecular graphs and experimental measurements. Foundation models may improve transfer to new targets or assay settings, although performance still depends on whether the training data represent the intended application (Teo et al., 2025; Koh et al., 2025).

5.2. Target-conditioned and three-dimensional generation:

Newer systems increasingly condition molecular generation on target information, binding-pocket geometry or desired interactions. This is more relevant to medicinal chemistry than generating molecules solely from a general distribution of drug-like structures. Three-dimensional approaches can represent conformational flexibility and spatial fit, although docking scores and predicted poses remain imperfect approximations of binding energetics. The combination of structure prediction, diffusion-based generation and medicinal-chemistry filters is a promising direction for target-specific design (Abramson et al., 2024; Tang et al., 2024; Doron et al., 2025).

5.3. Retrieval-augmented generation and scientific agents:

RAG connects an LLM to a defined evidence base, such as approved product information, standard operating procedures, scientific publications or institutional databases. The system retrieves relevant material and uses it to formulate an answer, making it easier to inspect the supporting sources. This approach can reduce unsupported responses but does not eliminate errors: retrieval may miss relevant evidence, documents may be outdated, and the model may misinterpret retrieved passages. Agentic systems extend this concept by planning and executing a sequence of tasks, such as searching a database, extracting candidate targets, generating structures and preparing a report. In pharmaceutical research, every tool action should be permissioned, logged and reviewed, especially when it affects regulated records or experimental decisions.

5.4. Closed-loop design–make–test–analyse systems:

The most consequential future workflow may be the integration of generative models with robotic synthesis, high-throughput screening, analytical instruments and automated data capture. In a closed loop, a model

proposes candidates, a laboratory system prepares or tests them, experimental results are returned to the model, and the next design cycle is updated. This approach could make optimisation more efficient by selecting experiments that are expected to provide useful information. It also creates new demands for equipment interoperability, assay quality, material traceability and reproducibility. A closed-loop system is only as reliable as the laboratory measurements and decision rules that guide it.

5.5. Synthetic data and privacy-preserving research:

GenAI can generate synthetic molecular, image or tabular data for method development, software testing and education. Synthetic data may reduce direct exposure of identifiable records, but it should not automatically be considered anonymous. A generative model can memorise training examples or reproduce rare records, and synthetic datasets can preserve biases present in the original data. Before synthetic data are used to support regulatory or clinical conclusions, their statistical properties, privacy risks and suitability for the intended purpose must be evaluated. Privacy-preserving methods such as federated learning may support collaboration without centralising all source data, but they do not remove the need for governance and security safeguards.

6. Benefits and potential impact:

The main potential benefit of GenAI is not simply faster text production or a larger number of candidate molecules. It is the possibility of exploring a wider design space, integrating fragmented knowledge and reducing repetitive tasks while allowing specialists to focus on decisions that require scientific judgement. In drug discovery, models can propose diverse chemical structures and protein designs that may not be obvious through manual design. In research operations, language models can help summarise literature, draft routine documents and structure complex information. In formulation and manufacturing, generative tools may help formulate hypotheses, retrieve process knowledge and plan experiments. Doron et al. (2025) highlight the potential for increased productivity and scientific discovery across pharmaceutical R&D.

These benefits should be measured against relevant baselines. Useful metrics include the proportion of generated structures that are chemically valid, synthesised and experimentally active; improvement in hit rate; time and cost per validated candidate; reproducibility across laboratories; reduction in document-processing time without loss of accuracy; and the number of safety events or quality deviations missed by an automated workflow. Benchmark scores alone are insufficient because they may reward model performance on artificial datasets without demonstrating practical value. Du et al. (2024) and Tang et al. (2024) emphasise the importance of evaluation design and task-relevant benchmarks in generative molecular research.

7. Challenges, risks and limitations:

7.1. Data quality, coverage and bias:

Generative models learn from available data, which may be incomplete, inconsistent, duplicated or biased towards particular diseases, chemical scaffolds, populations, laboratories and publication practices. Negative results and failed experiments are often under-reported, creating an incomplete picture of the chemical and biological search space. Assay values may also differ because of protocol, instrument, cell line or laboratory conditions. A model trained on narrow datasets may perform poorly for new targets, rare diseases, under-represented populations or unfamiliar molecular classes. Data provenance, curation, harmonisation and explicit description of dataset limitations are therefore essential.

7.2. Hallucination and scientific reliability:

LLMs can produce statements that are fluent and convincing but factually incorrect. They may invent references, misstate a result, confuse a drug name or omit an important safety qualification. Molecular generation has a related but distinct problem: an output can look plausible while containing invalid valence, unstable substructures, unrealistic stereochemistry or an impractical synthetic route. In both settings, plausibility is not evidence. RAG, constrained decoding, curated knowledge bases, chemical validators and expert review can reduce risk, but all high-impact outputs need source checking or experimental confirmation.

7.3. Interpretability, reproducibility and benchmark limitations:

Many deep generative models have complex internal representations that are difficult to interpret. This makes it challenging to identify why a molecule was proposed, which training example influenced a result or how sensitive an output is to small changes in prompt or input data. Benchmark leakage, repeated compounds and data overlap can produce overly optimistic estimates. Reproducibility also requires recording model versions, training data, prompts, software dependencies, random seeds, scoring functions and experimental protocols.

Without these records, another laboratory may be unable to reproduce either the computational output or the experimental outcome.

7.4. Translation from computational novelty to a medicine:

A generated molecule must satisfy a demanding sequence of conditions before it can become a therapeutic product. It must be synthesised reliably, possess suitable physicochemical properties, reach the intended site, engage the target, produce a meaningful biological effect and demonstrate acceptable safety. It must also be formulated, manufactured consistently and supported by adequate clinical evidence. Computational optimisation often focuses on a subset of these requirements. A structure with high predicted affinity may have poor solubility or unacceptable toxicity, while a protein with a desirable predicted fold may fail during expression or purification. This gap between in-silico performance and biological activity remains a central limitation of generative pharmaceutical research.

7.5. Regulatory accountability and validation:

Regulatory expectations depend on the intended use and the potential impact of an AI system on product quality, safety, efficacy or regulatory decisions. The EMA reflection paper published in 2024 discusses principles for AI and machine learning across the medicinal-product lifecycle, including development, manufacturing and post-authorisation activities (EMA, 2024). The FDA issued draft guidance in January 2025 on considerations for AI used to support regulatory decision-making for drug and biological products (FDA, 2025). The draft guidance is not a substitute for applicable law or binding regulations; it signals the need for a risk-based, context-specific approach. Sponsors should define the model's role, assess credibility for the intended context of use, document data and model development, and maintain change control and human oversight.

7.6. Privacy, intellectual property and cybersecurity:

Pharmaceutical datasets may contain confidential compound structures, unpublished research, clinical participant information, manufacturing parameters and commercially sensitive documents. Uploading these data to an external model without appropriate contractual, technical and legal safeguards may expose intellectual property or personal information. Copyright and inventorship questions may also arise when generated content is based on protected data or when AI contributes to an invention. Organisations should establish clear policies for approved platforms, access controls, data retention, model training, output ownership and disclosure. WHO and NIST guidance highlight the importance of accountability, transparency, privacy, security and risk management in AI systems (WHO, 2021; Autio et al., 2024).

7.7. Biosafety and dual-use concerns:

The capacity to design biological sequences or optimise molecular properties has potential benefits for therapeutics and biotechnology, but it may also create dual-use risks. Systems that generate proteins, peptides or biological functions should include appropriate screening, access restrictions and oversight proportionate to the potential harm. Safety controls should be integrated at the data, model, user and laboratory levels. Responsible research requires evaluating not only whether a model can generate a desired output but also whether the output could enable harmful use.

8. Implementation framework for pharmaceutical organizations:

A practical implementation programme should begin with a clearly defined problem rather than with the purchase of a general-purpose model. The organisation should specify the intended users, input data, generated output, decision affected and consequences of error. A low-risk literature-search assistant has different validation needs from a model that prioritises compounds for synthesis or supports a regulatory submission. The following principles provide a practical framework:

First, select a bounded use case with measurable value, such as literature triage, structured extraction from approved documents or generation of candidate molecular analogues for expert review. Second, establish data governance, including provenance, consent or other lawful basis where applicable, quality checks, access controls and retention rules. Third, choose a model whose architecture and training are appropriate for the task, and compare it with a non-generative baseline. Fourth, validate the system on representative external data and define acceptance criteria before deployment. Fifth, retain human review for decisions affecting product quality, patient safety, clinical evidence or regulatory commitments. Sixth, monitor output quality, drift, bias, user overrides and adverse events after deployment. Finally, maintain records of prompts, model versions, datasets, validation results and changes.

A cross-functional governance group should include pharmaceutical scientists, medicinal chemists, formulation experts, clinicians, biostatisticians, quality-assurance professionals, regulatory specialists, data scientists, information-security staff and legal advisers. Training should explain both the capabilities and limitations of GenAI. Staff should know how to verify sources, protect confidential information, recognise hallucinations and escalate unexpected outputs. The WHO governance principles and the NIST GenAI risk-management profile provide useful foundations for institutional policies (WHO, 2021; Autio et al., 2024).

9. Future perspectives:

Future pharmaceutical systems are likely to combine chemical, biological, clinical and manufacturing data within multimodal foundation models. Such systems may reason across sequences, molecular structures, assay results and literature, allowing researchers to formulate and compare hypotheses more efficiently. Agentic tools could coordinate database searches, computational design, simulation and laboratory scheduling. Their value will depend on reliable tool use, transparent intermediate steps and explicit permission boundaries. In the longer term, models may become integrated with automated laboratories to support iterative design–make–test–analyse cycles.

Personalised pharmaceutical development may also benefit from GenAI. Models could help propose dosage forms or delivery strategies that consider patient characteristics, drug properties and treatment requirements. However, individualised formulations require clinically justified inputs, robust quality controls and manufacturing processes that can deliver consistent products. In parallel, advances in protein generation may expand the range of engineered enzymes, binding proteins and therapeutic biologics. The roadmap described by Koh et al. (2025) suggests that generative design will increasingly be coupled with laboratory synthesis and testing rather than used as a standalone computational exercise.

Several research priorities remain important. Shared benchmarks should evaluate not only novelty and predicted drug-likeness but also synthesizability, target-specific activity, selectivity, ADMET performance and experimental reproducibility. Dataset documentation should capture assay conditions and negative findings. Independent laboratories should test leading claims, and prospective studies should compare GenAI-assisted workflows with conventional research. Regulatory agencies and standards organisations will need to clarify expectations for model credibility, data provenance, change control and post-deployment monitoring. In pharmacovigilance, the emerging evidence supports constrained extraction and triage applications with human review, while autonomous safety assessment remains premature (Kim et al., 2026).

The central future opportunity is a shift from isolated prediction towards evidence-linked scientific collaboration between people, models and laboratories. Pharmaceutical scientists will remain responsible for selecting meaningful questions, interpreting results, designing experiments and protecting patients. GenAI can expand the range of hypotheses considered and help manage information, but its success will be judged by the medicines and reliable knowledge produced through rigorous science.

10. Conclusion:

Generative artificial intelligence is extending the computational toolkit of pharmaceutical sciences. Its strongest emerging applications include de novo molecular design, lead optimisation, protein and peptide engineering, scientific literature support and structured information extraction. Potential downstream roles in formulation, clinical research, manufacturing and pharmacovigilance are developing, but the evidence is uneven and often less mature than in molecular design. Recent transformer, foundation-model and diffusion-based advances have increased the range of outputs that can be generated, while RAG and agentic workflows offer new ways to connect models with scientific tools and institutional knowledge. Significant limitations remain, including biased or incomplete data, hallucination, weak interpretability, uncertain synthetic feasibility, limited prospective validation, privacy and intellectual-property risks, and evolving regulatory expectations. GenAI should therefore be implemented as a governed decision-support and discovery technology, not as an autonomous replacement for pharmaceutical expertise. Its long-term contribution will depend on transparent evaluation, strong data stewardship, human accountability and repeated experimental verification.

11. References:

(1) Tong X, Liu X, Tan X, Li X, Jiang J, Xiong Z, et al. Generative models for de novo drug design. J Med Chem. 2021;64(19):14011-27. doi:10.1021/acs.jmedchem.1c00927.
(2) Meyers J, Fabian B, Brown N. De novo molecular design and generative models. Drug Discov Today. 2021;26(11):2707-15. doi:10.1016/j.drudis.2021.05.019.

(3) Tripathi S, Augustin AI, Dunlop A, Sukumaran R, Dheer S, Zavalny A, et al. Recent advances and application of generative adversarial networks in drug discovery, development, and targeting. Artif Intell Life Sci. 2022;2:100045. doi:10.1016/j.ailsci.2022.100045.
(4) Watson JL, Juergens D, Bennett NR, Trippe BL, Yim J, Eisenach HE, et al. De novo design of protein structure and function with RFdiffusion. Nature. 2023;620:1089-1100. doi:10.1038/s41586-023-06415-8.
(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;15:1331062. doi:10.3389/fphar.2024.1331062.
(6) Du Y, Jamasb AR, Guo J, Fu T, Harris C, Wang Y, et al. Machine learning-aided generative molecular design. Nat Mach Intell. 2024;6:589-604. doi:10.1038/s42256-024-00843-5.
(7) Abramson J, Adler J, Dunger J, Evans R, Green T, Pritzel A, et al. Accurate structure prediction of biomolecular interactions with AlphaFold 3. Nature. 2024;630(8016):493-500. doi:10.1038/s41586-024-07487-w.
(8) 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. 2024;25(4):bbae338. doi:10.1093/bib/bbae338.
(9) Gangwal A, Lavecchia A. Unleashing the power of generative AI in drug discovery. Drug Discov Today. 2024;29(6):103992. doi:10.1016/j.drudis.2024.103992.
(10) Huanbutta K, Burapapadh K, Kraisit P, Sriamornsak P, Ganokratanaa T, Suwanpitak K, et al. Artificial intelligence-driven pharmaceutical industry: a paradigm shift in drug discovery, formulation development, manufacturing, quality control, and post-market surveillance. Eur J Pharm Sci. 2024;203:106938. doi:10.1016/j.ejps.2024.106938.
(11) European Medicines Agency. Reflection paper on the use of artificial intelligence (AI) in the medicinal product lifecycle. EMA/CHMP/CVMP/83833/2023. Amsterdam: European Medicines Agency; 2024.
(12) US Food and Drug Administration. Considerations for the use of artificial intelligence to support regulatory decision-making for drug and biological products: draft guidance for industry and other interested parties. Silver Spring (MD): US Food and Drug Administration; 2025 Jan 7.
(13) World Health Organization. Ethics and governance of artificial intelligence for health: WHO guidance. Geneva: World Health Organization; 2021.
(14) Autio C, Schwartz R, Dunietz J, Jain S, Stanley M, Tabassi E, et al. Artificial intelligence risk management framework: generative artificial intelligence profile. Gaithersburg (MD): National Institute of Standards and Technology; 2024. (NIST AI 600-1). doi:10.6028/NIST.AI.600-1.
(15) Doron G, Genway S, Roberts M, Jasti S. Generative AI: driving productivity and scientific breakthroughs in pharmaceutical R&D. Drug Discov Today. 2025;30(1):104272. doi:10.1016/j.drudis.2024.104272.
(16) Shanmugam D, Agrawal M, Movva R, Chen IY, Ghassemi M, et al. Generative artificial intelligence in medicine. Annu Rev Biomed Data Sci. 2025;8:199-226. doi:10.1146/annurev-biodatasci-103123-095332.
(17) Teo ZL, Thirunavukarasu AJ, Elangovan K, Cheng H, Moova P, Soetikno B, et al. Generative artificial intelligence in medicine. Nat Med. 2025;31(10):3270-82. doi:10.1038/s41591-025-03983-2.
(18) Koh HY, Zheng Y, Yang M, Arora R, Webb GI, Pan S, et al. AI-driven protein design. Nat Rev Bioeng. 2025;3:1034-56. doi:10.1038/s44222-025-00349-8.
(19) Joshi S, Sheth S. Artificial intelligence (AI) in pharmaceutical formulation and dosage calculations. Pharmaceutics. 2025;17(11):1440. doi:10.3390/pharmaceutics17111440.
(20) Kim TY, Oh WS, Jeong DH. Large language models in adverse drug reaction detection and pharmacovigilance: a systematic review of current applications, challenges, and future directions. Diagnostics (Basel). 2026;16(15):2435. doi:10.3390/diagnostics16152435.