American Journal of
Biomedical Science & Research

www.biomedgrid.com
ISSN: 2642-1747

Review Article

Risks of Using Language Models in Biology for the Creation of Biological and Pharmaceutical Products

Andrey Shcherbakov[1,2,3,4*] and Alina Ryazanova[5]

[1]*Doctor of Technical Sciences, Ph.D. Medicine, Russia*

[2]*World Academy of Artificial Consciousness (WAAC), correspondent academician, Paris, France*

[3]*Scientific Director of the Center for Advanced Fundamental Research, Israel*

[4]*Leading Researcher at the State University of Management, 109542 Moscow*

[5]*Postdoctoral researcher, Kaluga branch of the Federal State Budgetary Educational Institution of Higher Education "Russian State Agrarian University – Moscow Timiryazev Agricultural Academy", Russia*

***Corresponding author:** Andrey Shcherbakov, Postdoctoral researcher, Kaluga branch of the Federal State Budgetary Educational Institution of Higher Education "Russian State Agrarian University – Moscow Timiryazev Agricultural Academy", Russia.

To Cite This article: *Andrey Shcherbakov* and Alina Ryazanova, Risks of Using Language Models in Biology for the Creation of Biological and Pharmaceutical Products. Am J Biomed Sci & Res. 2026 32(2) AJBSR.MS.ID.004132,* **DOI:** 10.34297/AJBSR.2026.32.004132

Received: September 01, 2026 **Published:** September 08, 2026

Abstract

Language models trained on textual, molecular, genomic, and protein data are becoming important tools in bioinformatics and pharmaceutical research and development. They can accelerate the identification of biological targets, protein annotation, molecular and protein design, and the interpretation of experimental data. At the same time, their broader adoption creates a range of risks, from errors and biases in computational predictions to medical-data leakage, impaired reproducibility, and the potential for dual-use applications [1-4].

Keywords: Large language models, Biology, Bioinformatics, Protein design, Drug discovery, Biosecurity, Dual use, Artificial intelligence

Introduction

The integration of language models into biological research is transforming the conventional pipeline for creating biological and pharmaceutical products. Unlike systems that operate solely on natural language, specialized models can process amino acid sequences, protein structures, chemical molecular representations, genomic data, and extensive collections of scientific publications. This enables their use for pattern discovery, hypothesis generation, and the prioritization of promising candidates before expensive laboratory experiments are undertaken [1,2]. The practical value of these systems is particularly evident in identifying disease-target associations, virtual screening of compounds, prediction of drug interactions, optimization of pharmacological properties, and support for clinical research [1,3]. However, an output produced by a language model should be regarded as a probabilistic hypothesis rather than an independently established scientific fact; it requires independent computational, biochemical, toxicological, and clinical validation [1,5].

Areas of Application

In biology, language models are used to predict protein functions, analyze protein-protein interactions, identify functionally significant sequence regions, and design protein variants [2,4]. Protein language models are trained on large collections of natural sequences and can identify statistical relationships potentially associated with protein structure, stability, or function [2,4]. In pharmaceutical research and development, models are applied to generate and rank small molecules, assess their interactions with biological targets, predict ADMET parameters—absorption, distribution, metabolism, excretion, and toxicity—and identify new indications for existing drugs [1,3,6]. These approaches can reduce the number of candidates requiring synthesis and experimental testing, but they do not eliminate the need for a full preclinical and clinical evaluation of safety and efficacy [1,3] (Table 1).

Table 1

Application area	Expected outcome	Key risk
Protein annotation	A hypothesis concerning a protein's function, properties, or interactions	Misinterpretation due to incomplete, imbalanced, or conflicting data [2,4]
Protein design	Novel enzymes, binding proteins, or therapeutic candidates	Unpredictable activity, immunogenicity, toxicity, or inadequate stability of the construct [2,5]
Small-molecule generation	Candidates for subsequent drug development	Incorrect assessment of synthesizability, bioavailability, efficacy, or safety [1,3,6]
Analysis of biomedical literature	Target discovery, fact extraction, and formulation of research hypotheses	Hallucinations, fabricated references, and conflation of hypotheses with experimentally confirmed findings [1,5]
Automated "design-synthesis-testing" cycle	Faster development iterations	Reduced human oversight and increased dual-use risks [7,8]

Risks to Reliability

The first fundamental risk is that a model may generate persuasive but unreliable information, including descriptions of biological mechanisms, gene-disease associations, target characteristics, or rationales for the promise of a compound. In pharmaceutical research, this may result in incorrect prioritization, loss of time and resources, and the establishment of false-positive development pathways [1,5]. The second risk relates to the limited interpretability of generative systems. In complex models, it may be difficult to determine which elements of the training data and computational representations produced a particular prediction. This complicates expert review, comparison of model outputs with known biological mechanisms, and the timely detection of errors [3,5]. The third risk stems from the quality and structure of the underlying data. Biological databases are heterogeneous in completeness, provenance, and reliability; moreover, well-studied genes, diseases, populations, and protein families are generally represented much more comprehensively than rare diseases or poorly studied organisms [2,5]. Consequently, a model may reproduce or amplify systematic data biases [2, 5].

Pharmaceutical Risks

A molecule generated by a model may appear promising in a computational prediction but prove unstable, difficult to synthesize, toxic, or pharmacologically inactive in a real biological system. In silico predictions of binding and toxicity do not replace experiments in cell cultures, model organisms, or clinical trials, because many effects depend on dose, metabolism, tissue context, and interactions with other substances [1,3,6]. A substantial risk arises when decision-making functions are delegated to a model. If a model is used to rank candidates, the model version, data sources, query parameters, filters, and subsequent validation results should be documented. Such records create an audit trail necessary for reproducibility, error analysis, and accountability [1,5].

Privacy and Rights

The training or operation of biomedical models may involve highly sensitive information, including genomic sequences, clinical records, data on rare diseases, and multi-omics patient profiles [2,5]. Principal threats include unauthorized access to datasets, re-identification of individuals after insufficient de-identification, and the possible reproduction of training-data fragments in model outputs [5]. To mitigate these risks, organizations should minimize the scope of processed data, apply de-identification, implement role-based access controls, use secure computing environments, and obtain legally valid informed consent. Identifiable clinical data should be transferred to external cloud services only after assessing the legal regime governing their storage, processing, and cross-border transfer [5].

Dual Use

The most serious societal risk derives from the dual-use nature of biotechnology. Methods intended to design therapeutic proteins, enzymes, vaccine candidates, or drug molecules may, in principle, be repurposed to develop toxic compounds and other hazardous biological agents [7,8]. Particular concern arises from integrating language models with biological-design tools, scientific databases, and automated laboratory platforms. In a closed "model-synthesis-experiment-feedback" loop, both iteration speed and the scale of

optimization increase, while researcher oversight may become insufficient [7,8]. Risk assessment should therefore account not only for a model's capabilities but also for its access to tools, data, laboratory infrastructure, and means of practical implementation [7].

Organizational Measures

The safe use of language models in biology and pharmaceutical development should follow the principle that models generate and rank testable hypotheses, whereas final decisions remain with qualified professionals. The following measures are recommended:

a. Independently validate every biologically significant prediction experimentally.

b. Clearly distinguish model-generated hypotheses from experimentally confirmed scientific findings in documentation.

c. Involve ethics and biosecurity review in projects with potential dual-use characteristics.

d. Restrict access to high-risk models, biological data, specialized tools, and laboratory platforms.

e. Maintain logs of model versions, prompts, data sources, selection criteria, and validation results.

f. Evaluate models through specialized tests addressing biosecurity risks and potentially harmful capabilities.

g. Establish incident-response procedures for data leaks, suspicious requests, and violations of applicable-use policies [7,8].

Conclusion

Language models can substantially accelerate the development of biological and pharmaceutical products, but their practical value depends not only on computational capability, but also on the quality of scientific, ethical, and organizational oversight [1,7]. A robust implementation model should combine algorithmic analysis, experimental validation, data protection, decision documentation, and biosecurity review [1,5,7].

The most justified approach is to use language models as tools supporting research rather than as autonomous developers of biological products. This approach makes it possible to realize their potential for science and pharmaceutical development while reducing the likelihood of errors, sensitive-data leakage, and harmful dual-use applications [2,7,8].

Acknowledgments

None.

Conflict of Interest

None.

References

1. Zheng Y, Koh HY, Ju J, Madeleine Yang, Lauren T May, et al. (2025) Large language models for drug discovery and development. Patterns 6(10): 101346.
2. Liu XH, Lu ZH, Wang T, Liu F (2024) Large language models facilitating modern molecular biology and novel drug development. Frontiers in Pharmacology 15: 1458739.
3. Othman ZK, Ahmed MM, Okesanya OJ (2025) AI-enabled language models to large language models in drug discovery and development: A comprehensive review. Journal of Advanced Research.
4. Liu Y, Ding S, Zhou S, Fan W, Tan Q (2024) MolecularGPT: Open large language model for few-shot molecular property prediction. arXiv.
5. Hu Z, Liu M, Liu Y (2024) A survey for large language models in biomedicine. arXiv.
6. Liu H, Li J, Zhou Y (2025) Large language models and their applications in drug discovery and development: A primer. Clinical and Translational Science. 18(4): e70205.
7. Esvelt K, Gopal A, Jeyapragasan T (2024) Recommendations to NIST on biosecurity-specific risk assessments and benchmarks for AI models. NIST Request for Information response.
8. (2024) Johns Hopkins Center for Health Security. Response to AISI Chem-Bio AI Model Request for Information: risk evaluation and mitigation for biological AI models.