Proceedings of ICBioMed 2026 Symposium: Biomedical Engineering and Public Health: Real-Time Wearables for Disease Detection
DOI: 10.54254/2753-8818/2026.ES38322

Advances in AI-Empowered Small-Molecule Drug Design for Vaccine Development and Immunotherapy

Zhiyi Zheng

School of Biotechnology, East China University of Science and Technology, Shanghai, China
MaxZheng21@outlook.com

Abstract. Artificial intelligence (AI) is reshaping drug discovery by shifting molecular design from empirical high-throughput screening toward data- and knowledge-guided workflows. This review summarizes the applications of graph neural networks, variational autoencoders, generative adversarial networks, biology-guided deep learning, and information-theoretic predictive models in vaccine development and cancer immunotherapy. Key tasks include molecular representation, de novo generation, virtual screening, ADMET prediction, antigen structure optimization, small-molecule adjuvant discovery, tumor microenvironment characterization, and precision dosing. Representative applications involving TLR/STING-related adjuvants, lipid nanoparticle delivery systems, and immunotherapy-response prediction are discussed to illustrate the complementarity of these computational approaches. Although AI can accelerate early-stage candidate identification and expand searchable chemical space, its clinical impact remains constrained by heterogeneous datasets, limited labeled samples, insufficient interpretability, weak cross-population generalization, and incomplete regulatory standards. Future research should emphasize standardized multimodal data, mechanism-aware model design, prospective validation, privacy-preserving collaboration, and closed-loop integration of computation, experimentation, and clinical feedback.

Keywords: artificial intelligence, small-molecule drug design, vaccine, graph neural network, biology-guided deep learning

1. Introduction

Public health emergencies keep emerging worldwide, and the disease burden from tumors keeps rising, which imposes stricter practical requirements on the molecular design work in vaccine development and tumor immunotherapy, requiring not only higher computational efficiency in design but also close attention to the actual translational value and computational accuracy of the results. The traditional model of new drug development as a whole shows real shortcomings such as long development cycles, large capital investment, and a persistently high probability of project failure. Existing statistical studies point out that a brand-new drug from early exploration to final launch costs on average about 2.6 billion US dollars, and the full development cycle is generally in the range of 10-15 years; even when a candidate molecule smoothly advances to the clinical trial stage, the proportion that finally gains approval for marketing remains low [1]. The industry calls

© 2026 The Authors. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

DOI: 10.54254/2753-8818/2026.ES38322

the objective phenomenon that R&D resources keep increasing but the efficiency of new drug output does not rise in step the 'reverse Moore's law' (Eroom's Law) [2]. Driven by the current reality of the industry, the R&D paradigm in pharmaceuticals is shifting; the past working mode that relied heavily on repeated trial and error based on experience is gradually being replaced, and the mode in which computational prediction and experimental validation cooperate with each other and jointly drive new drug development is being used more and more [3].

The rapid development of artificial intelligence (AI) technology provides solid technical support for the transformation of the R&D paradigm described above. Graph neural networks (GNNs) can directly read and learn the topological structure information of molecules themselves; variational autoencoders (VAEs) and generative adversarial networks (GANs), two types of models, are mostly used to explore the potential chemical space that has not been fully exploited; biology-guided deep learning (BgDL) introduces existing biological prior knowledge such as signaling pathways, cellular component features, and pathological states into the model construction process, so as to improve the degree of match between the model output and the real biological process, and at the same time enhance the interpretability of the model. The various technical means described above have penetrated multiple key steps of new drug development, covering research directions such as molecular feature representation, virtual screening of compounds, generation of brand-new candidate molecular structures, assessment of drug druggability, and improvement of dosing regimens [3-5].

In vaccine R&D, macromolecular antigens are the core substances that activate the body's specific immune response, while small molecules more often take on the function of adjuvants or immunomodulators, relying on their own actions to achieve the enhancement, remodeling, or prolongation of the immune response. At the present stage, the application boundary of AI technology keeps expanding and is no longer limited to the preliminary screening of candidate molecules; the application of AI technology can be seen in research steps such as antigen spatial conformation analysis and prediction, antigen epitope recognition and mining, adjuvant molecule screening, and delivery system scheme improvement [6-8]. Taking the structure prediction model built with deep learning as an example, this type of model can help researchers analyze the three-dimensional spatial conformation of viral proteins, providing a structural reference basis for the molecular design of prefusion-stable antigens. In the development of small-molecule adjuvants, AI algorithms can work together with actual compound libraries and high-throughput screening platforms of human peripheral blood mononuclear cells (PBMCs) to carry out research and screen out candidate compounds with specific immune-enhancing effects [7]. The relevant research case of the TLR7/8 agonist PVP-037 can prove that combining computational screening with experiments at the human-cell level can effectively improve the efficiency of discovering adjuvant molecules for specific populations [7].

In the field of tumor immunotherapy, the research focus of related algorithms has shifted; it no longer simply focuses on the identification and extraction of single molecular features, but instead carries out systematic modeling analysis on the complex interaction relationships inside the tumor microenvironment (TME). After integrating metabolomics, transcriptomics, proteomics, medical imaging data, and clinical follow-up data, AI tools can help researchers lock onto the metabolic regulatory nodes, intercellular information transmission patterns, and PD-L1 protein expression features that are associated with immunosuppressive effects, supporting the efficacy prediction of immune checkpoint inhibitors and the division of patient subgroups, and can also assist in mining small-molecule drug targets with development potential [9, 10]. Emerging technical ideas such as

DOI: 10.54254/2753-8818/2026.ES38322

digital twins further provide a feasible implementation path for simulating the dynamic evolution process of diseases under different dosing strategies [11].

Although AI technology has already shown the application potential of shortening the R&D cycle in early R&D steps such as active compound discovery and lead compound optimization, existing studies still lack sufficient evidence to prove that this technology can truly improve the clinical translation success rate of candidate drugs. There is obvious heterogeneity among multi-center datasets, the number of fully labeled samples is insufficient, the cross-dataset migration effect of models is unsatisfactory, and in addition there is a lack of unified validation and evaluation standards at the clinical level; all the above factors reduce the reliability of algorithms in real research scenarios [3, 5]. Based on the industry status described above, this paper systematically reviews the core AI algorithms that are widely used in the field of small-molecule drug design, summarizes the actual application cases of these algorithms in vaccine R&D and tumor immunotherapy, and at the same time discusses the real bottlenecks existing in current technology, and expands the discussion on the subsequent research and development directions.

2. Core algorithms and their roles in molecular design

2.1. Molecular representation learning: from handcrafted descriptors to graph neural networks

In the early stage of small-molecule new drug development, a basic technical difficulty is encountered, namely how to convert the chemical structure information of compounds into a numerical expression form that computer models can recognize and process. In the past, scientific research mostly used molecular fingerprints, physicochemical property descriptors, or SMILES (simplified molecular input line entry system) sequences to complete the molecular information conversion work [4, 5]. This type of expression method has certain advantages in computational execution efficiency, but has obvious shortcomings in the characterization ability for molecular spatial conformation, long-range atomic interactions, and local chemical microenvironments.

GNNs convert a complete molecule into graph-structured data, where atoms correspond to nodes in the graph and chemical bonds serve as the edges connecting nodes. Relying on the message-passing mechanism, the model iteratively completes the information aggregation operation between adjacent atom nodes, so as to learn the various features of internal functional groups, chemical bond connection patterns, and the overall topological level of the molecule. Compared with molecular fingerprints that need to be manually preset, GNNs can autonomously learn useful molecular features in light of specific research tasks, and therefore can adapt to multiple research tasks such as binding affinity prediction, compound activity classification, drug toxicity assessment, and virtual compound screening. When carrying out the related prediction work of protein-ligand interactions, if three-dimensional spatial coordinates, atomic distance constraints, or equivariant network architectures are further introduced, the model can more fully exploit the three-dimensional spatial information of molecules [3-5].

However, the final output effect of GNN models is largely constrained by the quality of the training dataset itself and the data coverage. Once the training dataset has systematic deviations in molecular skeleton types, target types, and experimental measurement conditions, the model is likely to only memorize the local chemical patterns within the dataset, and it is difficult to give credible prediction extrapolation results for brand-new skeleton molecules and targets that have not been encountered. Based on this situation, evaluating the comprehensive performance of a model cannot only refer to the test results obtained by random division of the dataset; validation methods

DOI: 10.54254/2753-8818/2026.ES38322

such as skeleton splitting and time-series splitting, together with the evaluation work of independent external test sets, are also indispensable, so as to objectively judge the true generalization performance of the model.

2.2. Generative models: VAE, GAN, and reinforcement learning

At present, the research thinking of molecular design is changing; it no longer simply screens molecules from within established physical compound libraries, but gradually transitions toward the direction of limiting multiple constraint conditions and autonomously generating brand-new molecular structures. VAEs rely on the encoder module to map discrete molecular samples to a continuous latent space, and then complete the molecular reconstruction or the generation of brand-new molecules through the decoder module. The continuous latent space can facilitate researchers to carry out molecular structure interpolation operations, targeted optimization of drug properties, and similar-molecule retrieval work, but the molecules output by this model also show various defects such as chemically invalid structures, highly repeated generated structures, or failure to reach the target activity properties [3-5].

GANs rely on the adversarial iterative training between the generator and the discriminator to continuously narrow the gap between the data distribution of generated molecules and the distribution of real compound samples. This training mode helps improve the structural diversity of the output candidate molecules, yet the unstable oscillation of the training process and mode collapse remain very prominent practical problems of this type of model. Reinforcement learning (RL) can integrate multiple indicators such as molecular binding activity, target selectivity, drug solubility, synthetic feasibility, and toxicity risk to construct a reward function, so as to carry out targeted optimization of the generation model for objective tasks. In actual research scenarios, it is a common phenomenon that the above optimization objectives check and balance each other and conflict with each other; the construction idea of the reward function, the setting scheme of various constraint conditions, and the search strategy for the Pareto optimal solution will all directly determine the quality of the molecular generation results.

The actual positioning of generative models is not to directly replace medicinal chemistry researchers; their core value lies in rapidly outputting a batch of structural hypotheses that can be tested by follow-up experiments. The candidate molecules output by the model still need to go through structural filtering and screening, synthetic route feasibility assessment, molecular docking operations, and molecular dynamics simulation, and then be combined with in vivo and in vitro multi-level experimental verification, before they can enter the next-stage structure optimization process.

2.3. Virtual screening, molecular generation, and ADMET prediction

When carrying out virtual screening research, the scoring function built by deep learning can simultaneously read molecular graph information, protein sequences, and part of the three-dimensional structure data, complete the ranking and screening of the possibility of binding between candidate compounds and targets, and thereby reduce the total amount of compounds that need to carry out physical experimental screening later (Table 1). Compared with traditional molecular docking scoring functions, models trained in a data-driven way can capture complex nonlinear association relationships, but the output results are very easily disturbed by the data distribution of the training set, the construction strategy of negative samples, and the complete quality of the protein input structure.

DOI: 10.54254/2753-8818/2026.ES38322

For the research direction of generating brand-new candidate molecules, combining generative models with RL algorithms can complete the optimization work at the level of target selectivity and drug druggability. The PKC-theta inhibitor EXS4318 developed by the company Exscientia is a typical example frequently cited in AI-assisted immunomodulator molecule design [3]. This type of R&D project can show that after target confirmation, AI tools can indeed accelerate the proposal and iteration rhythm of candidate molecules, but whether the subsequent R&D of candidate molecules can succeed is still jointly determined by multiple practical conditions such as pharmacological mechanism, in vivo drug exposure level, safety attributes, and clinical applicable diseases.

In the prediction work related to absorption, distribution, metabolism, excretion, and toxicity (ADMET), deep neural networks can rely on large-scale experimental measurement data to autonomously mine the internal connection between chemical structure and drug physicochemical properties, supporting the synchronous joint prediction of multiple evaluation endpoints [4, 5]. In recent years, some studies have introduced information-theoretic means and conditional probability models, striving to dynamically characterize the body's response effect produced by drugs under complex physiological environment conditions, rather than completely relying on artificially predefined explicit mathematical equations to carry out calculations [11]. This type of model expands the application boundary of traditional quantitative structure-activity relationship research, but the stability of the model itself must be confirmed by independent datasets and prospective physical experiments.

Table 1. Comparison of the functions of major AI methods in small-molecule design

Method	Core function	Main applications	Advantages	Main limitations
GNN	Learns molecular graphs and local topological features	Activity prediction, virtual screening, binding affinity prediction	Can learn structural features end-to-end	Depends on data distribution; underuses 3D information
VAE	Constructs a continuous latent space and generates molecules	Structure interpolation, similar-molecule search, multi-objective optimization	Facilitates continuous exploration of chemical space	May generate invalid or underperforming structures
GAN	Approaches the real molecular distribution via adversarial training	New-molecule generation, candidate library expansion	Improves the diversity of generated structures	Unstable training; risk of mode collapse
RL	Optimizes generation strategies based on a reward function	Optimization of selectivity, druggability, and synthetic feasibility	Suited to multi-objective targeted optimization	Results heavily depend on reward-function design
BgDL	Embeds biological pathways and pathological priors into the model	TME subtyping, prognosis, and immunotherapy-response prediction	Strong biological relevance and interpretability	High requirements for prior-knowledge quality and external validation
Information-theoretic models / digital twins	Simulates drug responses based on dynamic data	PK/PD prediction, individualized dose optimization	Supports continuous feedback and dynamic decision-making	Insufficient clinical data density and prospective evidence

2.4. Technical limitations and empirical reflection

At the present stage, the primary difficulty faced by AI-assisted drug design work is the insufficient supply of usable data that meets quality standards. Experimental datasets related to immune activity

DOI: 10.54254/2753-8818/2026.ES38322

generally have problems such as a small total number of samples, a low proportion of positive samples, large differences in measurement results across different experimental platforms, and non-unified criteria for defining research endpoints; the above problems easily cause the model to have overfitting and the output conclusions to fluctuate greatly. In addition, the internal decision logic of complex deep learning models is difficult to interpret intuitively, and the prediction scores given by the model often lack a clear corresponding relationship with the real biological action mechanism behind them, which also makes it difficult for clinical researchers and regulatory agencies to fully trust the model output results [3, 11].

It needs to be viewed objectively that the efficiency improvement in the early R&D stage cannot be equated with the rise in the success rate of the clinical stage of drugs. Some AI-driven R&D projects have compressed the cycle from hit compound screening to lead candidate determination to within 18 months, yet cases such as the candidate molecule DSP-1181, which announced the termination of development at the clinical trial stage, remind us that algorithm models cannot replace researchers in carrying out systematic research and understanding of the biological characteristics of targets, the heterogeneity of the disease itself, and the pharmacological laws of the human body [3]. It can be seen from this that AI drug design cannot simply rely on data-driven approaches, and needs to continue to develop toward a mode of collaborative driving by data, biological prior knowledge, and experimental result feedback.

3. Applications of AI in vaccine development and adjuvant design

3.1. Antigen structure prediction and vaccine molecule design

The core goal of vaccine development is to induce the body to produce a specific immune response with ideal effects and sufficient duration. To achieve this goal, researchers are required to fully grasp the three-dimensional spatial structure, conformational stability, and antigen epitope exposure of pathogen antigens. X-ray crystallography and cryo-electron microscopy can output protein structure information with very high accuracy, but carrying out experimental determination takes a long cycle. Structure prediction tools developed based on deep learning, such as AlphaFold and RoseTTAFold, can quickly output structural models of antigen proteins and receptor binding regions, providing a reference basis for the design modification of candidate antigens and the screening of mutation sites [12, 13].

In vaccine-related research work carried out for SARS-CoV-2, structure prediction means can judge the conformational stability of the spike protein, and at the same time analyze whether key antigen epitopes can be recognized and contacted by immune cells, so as to support the design work of prefusion-stable antigens [6]. Diffusion models such as TopoDiff can also generate protein backbone skeletons with specific topological attributes under global geometric constraints, opening up a brand-new computational idea for the structural optimization of antigen molecules [14]. However, the protein structures output by computational models still need to be verified by biochemical experiments, structural biology experiments, and immunogenicity-related experiments; a prediction structure with high confidence cannot be directly equated with the antigen having good immunogenicity.

3.2. Intelligent screening and targeted optimization of small-molecule adjuvants

Adjuvant substances can activate innate immune signaling pathways such as Toll-like receptors (TLRs) and stimulator of interferon genes (STING), thereby amplifying the immune response

DOI: 10.54254/2753-8818/2026.ES38322

induced by vaccines, improving the quality of the immune response, and prolonging the duration of immune protection [6, 8]. AI technology can be invested in multiple tasks such as adjuvant candidate molecule screening, targeted chemical structure modification, prediction of cellular-level immune response patterns, and potential adverse reaction risk assessment.

In the past, researchers used the customized software 'Synthetic Chemist' together with machine learning algorithms to screen out the TLR9 ligand CPG55.2 with cross-species activity from more than 10^16 kinds of theoretically synthesizable oligonucleotide molecules. Although oligonucleotides do not belong to the narrow category of small molecules, this case can still prove that, in the face of an extremely large candidate molecular space, machine learning algorithms can efficiently complete the molecular priority ranking work [8]. For research on STING agonists, combining high-throughput virtual screening means with GNN molecular representation methods can jointly evaluate molecular structure characteristics, target binding potential, and drug druggability, and point out candidate modification directions for the optimization of molecular membrane permeability, intracellular stability, and target selectivity [6, 8].

Generative models can also produce brand-new molecular structures that meet the activity requirements of specific targets. The system GENTRL integrates generative models with reinforcement learning algorithms, and after given multiple objective constraints, it can quickly output a batch of candidate small-molecule structures [15]. From the methodological level, this system can shorten the cycle iteration time of 'structure generation - computational evaluation - molecular re-modification'. However, in actual implementation and use, conditions such as synthetic feasibility, potential toxicity risk, and formulation process adaptability still need to be included in the constraint scope, so as to avoid obtaining molecules with excellent computational scores but completely lacking subsequent development value.

3.3. Vaccine delivery systems and formulation optimization

After vaccination, the actual immune effect is influenced by the properties of the antigen and adjuvant themselves; the physical stability of the delivery system, the in vivo tissue distribution characteristics, and the cell uptake efficiency also have an important influence. In the field of nucleic acid vaccine research, the particle size, surface charge attributes, lipid composition ratio, nucleic acid encapsulation rate, and storage-condition stability of lipid nanoparticles (LNPs) all directly change the delivery effect and immune activation level.

Combining AI-assisted molecular dynamics simulation operations, Bayesian optimization algorithms, and formulation-effect correlation modeling can analyze the complex nonlinear associations between multiple physicochemical parameters and delivery efficiency, and can also complete the priority ranking of alternative formulations. Compared with the physical experimental screening mode that completely relies on exhaustive trial and error, computational models can screen out a large number of formulation combinations with no research value and improve the overall experimental design efficiency. However, LNP formulation systems are very easily disturbed by preparation batches and experimental operation conditions; therefore, such computational models must rely on standardized preparation processes and be combined with independent experimental datasets to complete parameter calibration work [5, 16].

3.4. Multi-algorithm collaboration and its application boundaries

By integrating the various research steps described above, it can be seen that AI technology has penetrated into work such as antigen structure analysis, adjuvant candidate molecule screening, and

DOI: 10.54254/2753-8818/2026.ES38322

delivery system formulation improvement. Structure prediction tools output antigen conformation information and mine potential antigen epitopes; generative models combined with virtual screening means expand the selectable chemical molecular space; modeling related to delivery systems can also feed back the problems encountered at the molecular modification and formulation levels to the front-end molecular design process in reverse. Integrating the various algorithms described above into a complete workflow can build a coherent research chain from antigen biological analysis and target discrimination to candidate molecule and formulation development.

However, the advantages of this intelligent R&D mode are more concentrated in the early molecular design and candidate priority ranking stages of a project, and can accelerate the early response speed to a certain extent when facing public health emergencies. As for whether this technology can truly lower the overall R&D investment, raise the clinical translation success rate, or improve the adaptation effect of vaccines among different populations, there is still a lack of sufficiently long-term project follow-up data as support [3, 5]. At the present stage, a more pragmatic research idea is to connect computational prediction results with human-cell experiments, organoid models, animal in vivo experiments, and clinical sample analysis in series, build a progressively advancing verification path, and rely on hierarchical verification to narrow the deviation between the model's computational output results and the body's real immune response.

4. Applications of BgDL in tumor microenvironment analysis and precision dosing

4.1. BgDL and TME metabolic profile characterization

High heterogeneity exists inside the TME, and this feature is regarded as one of the key inducements that interfere with the efficacy of tumor immunotherapy and promote the occurrence of drug resistance. Although traditional deep learning models can extract predictive features from high-dimensional biological data, they can hardly explain the real biological process corresponding to these features; there is a difficult-to-reconcile contradiction between the feature representation extracted inside the model and the interpretability required by clinical research [9]. The implementation idea of BgDL is to introduce existing biological priors in the network architecture design or loss function setting step, including immune cell infiltration, stromal components, metabolic signaling pathways, and gene regulatory relationships, so that the model output can maintain a higher consistency with the cognition at the level of biological mechanisms.

When carrying out TME metabolic profile analysis, AI can integrate metabolomics, transcriptomics, and proteomics datasets to identify various tumor microenvironment states such as massive lactate accumulation, glutamine consumption, lipid metabolism reprogramming, and excessive adenosine signal activation. The multi-omics network model built relying on the graph attention mechanism can characterize the interrelationships between different cell subpopulations and metabolic pathways, helping researchers lock onto patient subpopulations that may benefit from IDO1 inhibitors or other metabolism-regulating drugs [10]. However, the results output by the model can mostly only show that there is a correlation between variables; if one wants to further confirm the causal-level action logic, it is still necessary to combine spatial omics detection, functional verification experiments, and treatment response follow-up data for comprehensive analysis.

DOI: 10.54254/2753-8818/2026.ES38322

4.2. From static dose prediction to dynamic response modeling

The core demand of individualized precision dosing is to find the balance point between treatment benefit and drug toxicity for each patient at different treatment stages. The application of AI in this direction is no longer limited to ADMET property prediction at a single time node; dynamic modeling of pharmacokinetics/pharmacodynamics (PK/PD) relying on longitudinal follow-up data has gradually become a research hotspot [11]. Information-theoretic prediction models use conditional probability relationships to characterize the drug-body response under different physiological conditions, and can serve as a type of supplementary modeling means for complex biological systems that lack mature mechanistic equations.

Taking the platform curate.ai as an example, this platform can read the series of detection data of patients that change over time, continuously update the corresponding relationship between dose and effect, and thereby provide a reference basis for the dose adjustment work of single-drug or combination-drug regimens [17]. Digital twin technology attempts to build a dynamic computational model of an individual patient, simulating the evolution process of the disease and the risk of adverse reactions under different treatment plans. This type of technology has development potential at the clinical application level, but simulation cannot achieve zero error and cannot be directly equated with the precise prediction of the real situation. When put into actual clinical scenarios, this type of tool can only be embedded in the whole process of doctors' clinical decision-making, real-time condition monitoring, and result traceable verification, and absolutely cannot directly replace clinicians in making diagnosis and treatment judgments.

4.3. Case study: gastric cancer immunotherapy response prediction

In research related to gastric cancer, the BgDL model built by Jiang et al. is used to predict the prognostic outcome of patients and the treatment effect of anti-PD-1 drugs, and at the same time attempts to mine the patient subpopulation within deficient mismatch repair (dMMR) tumors that still responds poorly to immunotherapy. Published research results have confirmed that, when the TME subtype results divided by the BgDL model are used together with clinical biomarkers such as the combined positive score (CPS), the area under the curve (AUC) of treatment response prediction can reach above 0.80 [9].

This research case fully reflects the application value of the mutual fusion of biological prior knowledge and clinical detection indicators. In addition to outputting risk scores, the model can also correspond the prediction results to an interpretable TME state classification. However, the model performance obtained by relying on a single-center retrospective dataset cannot be directly equated with the actual effect in a real clinical environment; external cohort validation and model fairness analysis of different subgroups still need to be completed, and prospective clinical trials need to be advanced for corroboration.

4.4. Implications for clinical practice

BgDL has a very practical application value: it tries to open up the pathway between underlying biological mechanism analysis and clinical outcome prediction, and to a certain extent alleviates the 'interpretability difficulty' commonly encountered when AI models move toward clinical translation. The metabolic characteristics and cellular component information inside the TME can both provide a reference for patient stratification and point out a direction for the screening of small-molecule immunomodulatory drug targets; at the same time, continuously collected clinical treatment

DOI: 10.54254/2753-8818/2026.ES38322

feedback can also iteratively correct the model in reverse, so that it is continuously optimized in the individualized diagnosis and treatment process. In addition, models built by combining radiomics with multi-omics have the opportunity to realize minimally invasive or even non-invasive monitoring of TME states, reducing the detection deviation brought by tumor spatial heterogeneity in a single tissue biopsy [9, 10]. As privacy-preserving technologies such as federated learning continue to iterate, multiple medical institutions can collaboratively complete model training work without directly sharing raw data, but the model iteration update rules, data governance norms, and rights-and-responsibilities division issues still need clear and explicit institutional constraints [18].

5. Key challenges and future development trends

5.1. Data quality and standardization

The quality ceiling of the dataset itself basically determines the performance ceiling that AI models can reach. Research related to vaccines and tumor immunotherapy involves chemical structure information, in vitro activity measurement results, multi-omics high-throughput data, medical imaging, and various clinical outcome indicators, and belongs to a typical multimodal research scenario; different detection platforms, sample pre-processing methods, and efficacy evaluation endpoints all cause huge differences between data [3, 5]. In subsequent research work, a data dictionary with complete traceability should be gradually built, the judgment criteria for compound activity should be unified, the processing strategy for missing data should be clarified, and a cross-research-center quality control process should be established. At the same time, attention should be paid to the data archiving records of negative experiments and failed projects, so as to weaken the interference of publication bias on model training.

5.2. Interpretability and mechanistic consistency

If research work one-sidedly pursues the numerical level of prediction indicators, the finally obtained models mostly can only reflect the correlation between variables, and it is difficult to carry out experimental confirmation. Subsequent research should fully use multiple means such as pathway constraint construction, causal graph analysis, attention mechanism analysis, counterfactual reasoning, and experimental perturbation verification, to check whether the features learned by the model conform to the reported biological action mechanisms [9, 11]. Models developed for clinical decision-making also need to clearly explain the uncertainty of prediction results, the boundary of the applicable population of the model, and the failure scenarios of the model; one must avoid directly interpreting the probability-form model output as a definite clinical conclusion.

5.3. Generalization validation, privacy protection, and regulation

The generalization ability of models across different populations, disease subtypes, and experimental platforms is the core evaluation indicator of their clinical translation. In addition to internal cross-validation, external time validation, external geographic validation, and prospective validation should be adopted to report the performance differences of different subgroups. At the same time, technologies such as federated learning, trusted execution environments, and differential privacy can reduce the risk of data leakage in cross-institution collaboration, but may also introduce performance loss and system complexity [18]. At the regulatory level, the training data source,

DOI: 10.54254/2753-8818/2026.ES38322

model version, update rules, audit records, and human supervision responsibility need to be clarified.

5.4. Building a human-machine collaborative closed-loop R&D system

The future development direction of AI-empowered drug R&D is not to build a fully automatic R&D system completely free from manual intervention, but to build a closed-loop working mode of 'computational generation of scientific hypotheses - physical experiments complete verification - experimental results are fed back - model iterates and updates' [3, 5]. Medicinal chemistry researchers, immunology researchers, clinicians, and algorithm developers need to participate together and clarify the objective function and verification judgment criteria that the model needs to optimize. Active learning combined with Bayesian optimization algorithms can, based on already obtained experimental results, preferentially select the next batch of candidate molecules with the highest information gain to carry out experiments, and improve the iterative R&D efficiency under the realistic conditions of limited experimental funds and sample resources. Only when the model can continuously receive real experimental data and clinical sample data to complete correction can the prediction advantage brought by the algorithm be transformed into stable and implementable R&D benefits.

6. Conclusions

AI technology has already shown application potential in many research steps of small-molecule drug design, and is gradually producing deep intersections with vaccine development and tumor immunotherapy research. GNNs, VAEs, GANs, RL, BgDL, and information-theoretic models respectively provide researchers with brand-new research tools from multiple angles such as molecular structure representation, brand-new molecule generation, druggability optimization, biological mechanism analysis, and individualized dynamic dosing. The application value of this type of technology is more reflected in improving the screening efficiency of early candidate compounds, expanding the explorable chemical molecular space, integrating multimodal biological evidence, and assisting the completion of patient subgroup division.

However, from the perspective of the actual development level at the present stage, AI can only serve as an auxiliary decision-making tool, and absolutely cannot replace the verification work at the physical experiment and clinical levels. A series of practical problems such as data heterogeneity, insufficient model interpretability, limited cross-scenario generalization ability, patient data privacy protection, and industry regulatory rights-and-responsibilities division directly determine whether the algorithm model can truly integrate into a reproducible, auditable, and translatable new drug R&D process. If sustained breakthroughs can be made in directions such as biological mechanism knowledge embedding, standardized data system construction, prospective verification system construction, and the implementation of human-machine collaborative closed-loop R&D mode, the practical application value of AI technology in vaccine and immunomodulatory small-molecule drug development will be further released.

References

[1] DiMasi, J. A., Grabowski, H. G., & Hansen, R. W. (2016). Innovation in the pharmaceutical industry: New estimates of R&D costs. *Journal of Health Economics, 47*, 20-33. DOI:10.1016/j.jhealeco.2016.01.012.

[2] Scannell, J. W., Blanckley, A., Boldon, H., et al. (2012). Diagnosing the decline in pharmaceutical R&D efficiency. *Nature Reviews Drug Discovery, 11*(3), 191-200. DOI:10.1038/nrd3681.

DOI: 10.54254/2753-8818/2026.ES38322

[3] Zhang, K., Yang, X., Wang, Y., et al. (2025). Artificial intelligence in drug development. *Nature Medicine, 31*(1), 45-59. DOI:10.1038/s41591-024-03434-4.
[4] Martinelli, C., Repetto, M., & Curigliano, G. (2024). Artificial intelligence in small-molecule drug discovery. In *Artificial intelligence for medicine* (pp. 37-45). Elsevier. DOI:10.1016/B978-0-443-13671-9.00012-0.
[5] Serrano, D. R., Luciano, F. C., Anaya, B. J., et al. (2024). Artificial intelligence (AI) applications in drug discovery and drug delivery: Revolutionizing personalized medicine. *Pharmaceutics, 16*(10), 1328. DOI:10.3390/pharmaceutics16101328.
[6] Gupta, A., Rudra, A., Reed, K., et al. (2024). Advanced technologies for the development of infectious disease vaccines. *Nature Reviews Drug Discovery, 23*, 914-938. DOI:10.1038/s41573-024-01041-z.
[7] Soni, D., Borriello, F., Scott, D. A., et al. (2024). From hit to vial: Precision discovery and development of an imidazopyrimidine TLR7/8 agonist adjuvant formulation. *Science Advances, 10*(27), adg3747. DOI:10.1126/sciadv.adg3747.
[8] Zhang, W. Y., Zheng, X. L., Coghi, P. S., et al. (2024). Revolutionizing adjuvant development: Harnessing AI for next-generation cancer vaccines. *Frontiers in Immunology, 15*, 1438030. DOI:10.3389/fimmu.2024.1438030.
[9] Jiang, Y., Zhang, Z., Wang, W., et al. (2023). Biology-guided deep learning predicts prognosis and cancer immunotherapy response. *Nature Communications, 14*, 5135. DOI:10.1038/s41467-023-40890-x.
[10] Hai, L., Jiang, Z., Zhang, H., et al. (2024). From multi-omics to predictive biomarker: AI in tumor microenvironment. *Frontiers in Immunology, 15*, 1514977. DOI:10.3389/fimmu.2024.1514977.
[11] Saied, H., Alfahad, O., Aljaffer, A. A., et al. (2026). From data to discovery: The rise of information-theoretic predictive models in drug development. *Scientific Reports, 16*, 12857. DOI:10.1038/s41598-026-45644-5.
[12] Jumper, J., Evans, R., Pritzel, A., et al. (2021). Highly accurate protein structure prediction with AlphaFold. *Nature, 596*(7873), 583–589. DOI:10.1038/s41586-021-03819-2.
[13] Baek, M., DiMaio, F., Anishchenko, I., et al. (2021). Accurate prediction of protein structures and interactions using a three-track neural network. *Science, 373*(6557), 871–876. DOI:10.1126/science.abj8754.
[14] Zhang, Y., Liu, Y., Ma, Z., et al. (2025). Improving diffusion-based protein backbone generation with global-geometry-aware latent encoding. *Nature Machine Intelligence, 7*, 1104–1118. DOI:10.1038/s42256-025-01059-x.
[15] Zhavoronkov, A., Ivanenkov, Y. A., Aliper, A., et al. (2019). Deep learning enables rapid identification of potent DDR1 kinase inhibitors. *Nature Biotechnology, 37*(9), 1038–1040. DOI:10.1038/s41587-019-0224-x.
[16] Hou, X., Zaks, T., Langer, R., et al. (2021). Lipid nanoparticles for mRNA delivery. *Nature Reviews Materials, 6*, 1078–1094. DOI:10.1038/s41578-021-00358-0.
[17] Blasiak, A., Khong, J., Kee, T. (2020). CURATE.AI: Optimizing personalized medicine with artificial intelligence. *SLAS Technology, 25*(2), 95–105. DOI:10.1177/2472630319890316.
[18] Rieke, N., Hancox, J., Li, W., et al. (2020). The future of digital health with federated learning. *npj Digital Medicine, 3*, 119. DOI:10.1038/s41746-020-00323-1.