ISSN: 3060-4737 Volume 3, Issue 10 IF(Impact Factor) 10.98 / 2026

ARTIFICIAL INTELLIGENCE IN DRUG DISCOVERY AND DEVELOPMENT

Rishabh Jasoria , Medical Student , Samarkand state medical university, Uzbekistan

Adil Patel , Medical Student , Samarkand state medical university, Uzbekistan

Patel Mo.Uvesh , Medical Student , Samarkand state medical university, Uzbekistan

Aysha Sama , Medical Student , Samarkand state medical university, Uzbekistan

Bringing a new medicine to patients is slow, expensive and uncertain, and about nine in ten candidates that enter human trials never reach approval. This narrative review looks at how artificial intelligence (AI) is being used to attack that problem in three places: choosing drug targets, predicting toxicity before a compound is ever made, and shortening the time between an idea and a clinical trial. Using published studies and well-documented case examples, including AlphaFold, the antibiotic halicin and the AI-derived drug rentosertib, we describe what these tools do well and where they fall short. Early evidence is encouraging but uneven: molecules discovered with AI appear to pass Phase I safety trials far more often than the historical average, yet their Phase II success rate looks much like that of conventional drugs. AI is therefore speeding up and de-risking the early stages, while proving that a drug works in patients remains the hard part. Data quality, explainability and independent validation will decide how far the progress goes.

Keywords: *artificial intelligence; drug discovery; target identification; toxicity prediction; machine learning; generative chemistry; clinical trials*

1. Introduction

Ask anyone who has worked in pharmaceutical research how a medicine is born and you will hear a story about patience. A biologist spots a protein that seems to drive a disease. Chemists spend years shaping a molecule that can block it. Animal studies follow, then three phases of human trials. The whole journey typically takes 10 to 15 years and costs somewhere around a billion dollars or more, and roughly nine out of ten candidates that reach human testing are never approved (Paul et al., 2010; Sun et al., 2022).

Why do so many fail? Sun and colleagues (2022) analysed clinical trial data from 2010 to 2017 and attributed 40–50% of failures to a lack of clinical efficacy and about 30%

ISSN: 3060-4737 Volume 3, Issue 10 IF(Impact Factor) 10.98 / 2026

to unmanageable toxicity, with poor drug-like properties and weak commercial planning making up most of the rest (Figure 1). Those two large slices point to exactly the questions AI is now being asked to help with. Are we aiming at the right biological target? And will the molecule be safe enough to give to people?

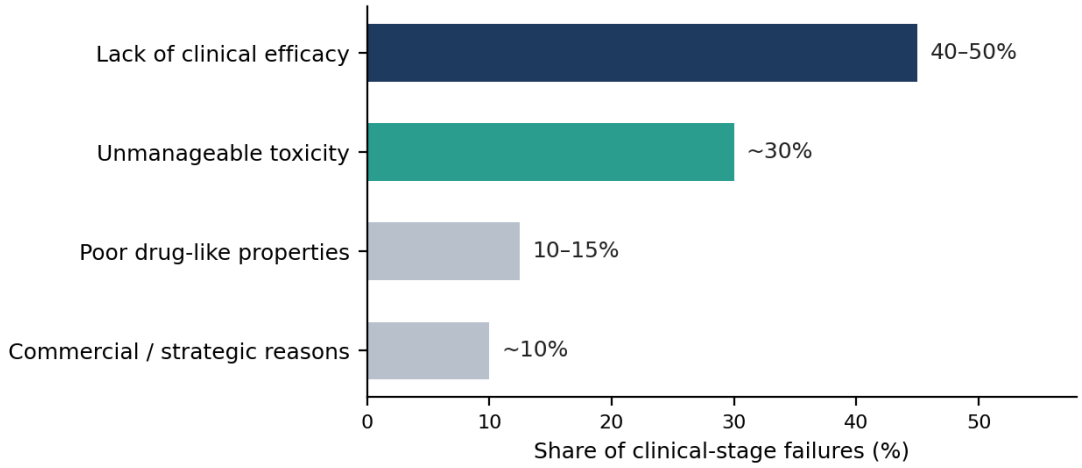


Figure 1. Main reasons for failure in clinical drug development. Bars show the midpoint of each reported range. Data from Sun et al. (2022).

This article follows those questions through the drug pipeline. It is a narrative review rather than a systematic one: sources were chosen for their influence and relevance, not through an exhaustive database search, and the aim is to give a clear, honest picture rather than a count of every publication.

2. Finding the Right Target

A drug target is usually a protein whose activity contributes to a disease. Picking the wrong one is the costliest mistake in the field, because it can take years and many millions before a clinical trial reveals it. Targets used to come from decades of bench biology, human genetics and a certain amount of luck. What has changed is the sheer volume of data: genome sequences, gene-expression profiles, proteomics, patient records and millions of research papers. No team can read all of it, but machine-learning systems can at least rank it.

Several approaches are common. Knowledge graphs link genes, proteins, diseases, drugs and publications into one network, so an algorithm can notice a connection nobody has stated outright. Models trained on omics data compare diseased and healthy tissue and flag genes whose changes look causal rather than incidental. Language models scan the literature at a scale no human reader could manage. Each method

ISSN: 3060-4737 Volume 3, Issue 10 IF(Impact Factor) 10.98 / 2026

produces a ranked list, and a scientist still has to decide which entries deserve a laboratory test (Vamathevan et al., 2019).

Structure matters just as much as selection. Once a target is chosen, chemists want to know its three-dimensional shape, because that shape decides which molecules can fit. AlphaFold 2 predicted protein structures with accuracy often close to that of laboratory experiments (Jumper et al., 2021), and its successor, AlphaFold 3, extends the idea to complexes of proteins with DNA, RNA and small molecules (Abramson et al., 2024). For many targets no experimental structure exists, so these predicted models give design teams somewhere to start.

The best-known example of an AI-nominated target is TNIK in idiopathic pulmonary fibrosis (IPF), a progressive scarring disease of the lungs. Insilico Medicine's platform ranked TNIK as a promising anti-fibrotic target and then generated a molecule, rentosertib, to inhibit it (Ren et al., 2025). In a randomised Phase IIa trial of 71 patients, those taking 60 mg daily gained an average of 98.4 mL in forced vital capacity over 12 weeks, while the placebo group lost 20.3 mL (Xu et al., 2025). That is encouraging, but the trial was small and short, more patients on the highest dose stopped treatment early than on placebo, and the authors themselves call for larger studies.

Table 1. Where AI is used along the drug-development pipeline

Stage	Typical AI approach	What it adds
Target identification	Knowledge graphs, omics mining, genetics-based ranking	Prioritises disease-linked proteins from very large datasets
Structure and binding	AlphaFold-type structure prediction; docking with learned scoring	Provides 3D models where no experimental structure exists
Hit finding and design	Virtual screening, graph neural networks, generative models	Explores far more chemical space than a physical library can
Safety and ADMET	QSAR and deep learning on assay data	Flags risky compounds before they are synthesised

ISSN: 3060-4737 Volume 3, Issue 10 IF(Impact Factor) 10.98 / 2026

Stage	Typical AI approach	What it adds
Clinical development	Patient stratification, site selection, health-record mining	Aims to reduce delays and mismatched patient groups

3. From Target to Molecule

With a target in hand, the job is to find or build a molecule that fits it. The space of possible drug-like molecules is astronomically large, far beyond anything that could be made and tested, so physical screening only samples a tiny corner of it. AI helps in two ways: it can score existing libraries on a computer, and it can invent new structures.

The halicin story shows the first route. In 2020 a team at MIT trained a graph neural network on a few thousand molecules labelled by whether they stopped E. coli from growing, then used it to screen a library of repurposed compounds. The model picked out halicin, a compound originally studied for diabetes, whose structure looks nothing like known antibiotics. It went on to show activity against a range of resistant bacteria in laboratory and mouse experiments (Stokes et al., 2020).

Generative models go a step further. In 2019, Zhavoronkov and colleagues used a generative approach to design inhibitors of DDR1, a kinase linked to fibrosis, in 21 days and then tested a lead compound in mice (Zhavoronkov et al., 2019). In the rentosertib programme, the same company reported getting from target discovery to a preclinical candidate in about 18 months (Ren et al., 2025), well under the multi-year timelines usually quoted for this stage. Table 2 collects a few of these milestones.

Table 2. Selected milestones in AI-supported drug discovery

Year	Milestone	Why it matters
2019	Generative design of DDR1 kinase inhibitors in 21 days (Zhavoronkov et al.)	Showed the design cycle could shrink from years to weeks
2020	Deep-learning discovery of the antibiotic halicin (Stokes et al.)	Found a structurally novel antibacterial among existing compounds

ISSN: 3060-4737 Volume 3, Issue 10 IF(Impact Factor) 10.98 / 2026

Year	Milestone	Why it matters
2021	AlphaFold 2 predicts protein structures at near-experimental accuracy (Jumper et al.)	Opened structure-based design to many previously unsolved targets
2024	Analysis of AI-discovered molecules in trials (Jayatunga et al.)	First systematic look at how such molecules perform in people
2025	Phase IIa results for rentosertib published (Xu et al.)	Early clinical signal for a drug whose target and structure were AI-derived

4. Predicting Toxicity Before It Happens

Toxicity is the second great killer of drug candidates, and it is also where computational methods have the longest track record, because the question can be framed as a plain prediction problem: given a structure, will it harm a liver, a heart or DNA? Early quantitative structure–activity relationship (QSAR) models relied on chemical descriptors chosen by hand. Modern deep-learning models learn their own features directly from the molecular graph. A landmark was DeepTox, which won the Tox21 Data Challenge by learning from roughly 12,000 compounds tested across 12 assays for nuclear-receptor and stress-response signalling (Mayr et al., 2016).

In practice, teams build panels of models, each aimed at one endpoint (Table 3). If a compound is predicted to block the hERG channel, which is a known cardiac risk, chemists can redesign it before anyone makes it. That is where the real saving lies. Failing a computer screen costs almost nothing, while failing in a Phase II trial costs years and tens of millions.

Table 3. Common toxicity endpoints modelled with machine learning

Endpoint	Why it matters	Typical data and models
Cardiotoxicity (hERG block)	Can disturb heart rhythm; a classic reason for late-stage withdrawal	Public hERG inhibition data; random forests, graph neural networks

ISSN: 3060-4737 Volume 3, Issue 10 IF(Impact Factor) 10.98 / 2026

Endpoint	Why it matters	Typical data and models
Liver injury (hepatotoxicity)	A leading safety reason for stopped trials and withdrawn drugs	Curated liver-injury labels; datasets are small and noisy
Mutagenicity	Flags compounds that may damage DNA	Large Ames-test datasets; structural alerts combined with ML
Nuclear-receptor and stress-response assays	Signals endocrine disruption and cellular stress	Tox21 library; multitask deep learning such as DeepTox

There are caveats. Toxicity data are patchy and biased towards compounds that were tested because they already looked promising. Many harms are idiosyncratic, striking a small number of people because of their genes or other medicines, and a model trained on cell assays will rarely capture that. Differences between animals and humans remain, too. Finally, deep models are often black boxes, yet regulators and chemists want to know which part of a molecule triggered an alert, so explainability tools matter nearly as much as accuracy.

5. Speeding Up Development

Speed is the headline promise. Faster target ranking, virtual screening and in-silico safety filters shorten the design–make–test cycle, and the DDR1 and TNIK examples suggest that early stages can take weeks to months rather than years. Beyond the laboratory, machine learning is being applied to clinical development itself: mining electronic health records to find eligible patients, choosing trial sites, predicting dropouts, and stratifying patients so a drug is tested in the people most likely to respond (Vamathevan et al., 2019). The evidence for real gains here is thinner than it is for early discovery.

The most informative numbers so far come from Jayatunga et al. (2024), who examined molecules from AI-native biotechnology companies. Of 24 molecules that had completed Phase I, 21 succeeded, a rate of about 80–90% against a historical range of roughly 40–65%. In Phase II, however, success was around 40%, based on only about ten completed trials, which is close to the historical figure of roughly 29% reported by

ISSN: 3060-4737 Volume 3, Issue 10 IF(Impact Factor) 10.98 / 2026

Wong et al. (2019) (Figure 2). One sensible reading is that AI is good at producing molecules that behave well in the body, meaning safe and drug-like, but has not yet shown it can choose biology that works in patients more reliably than humans do.

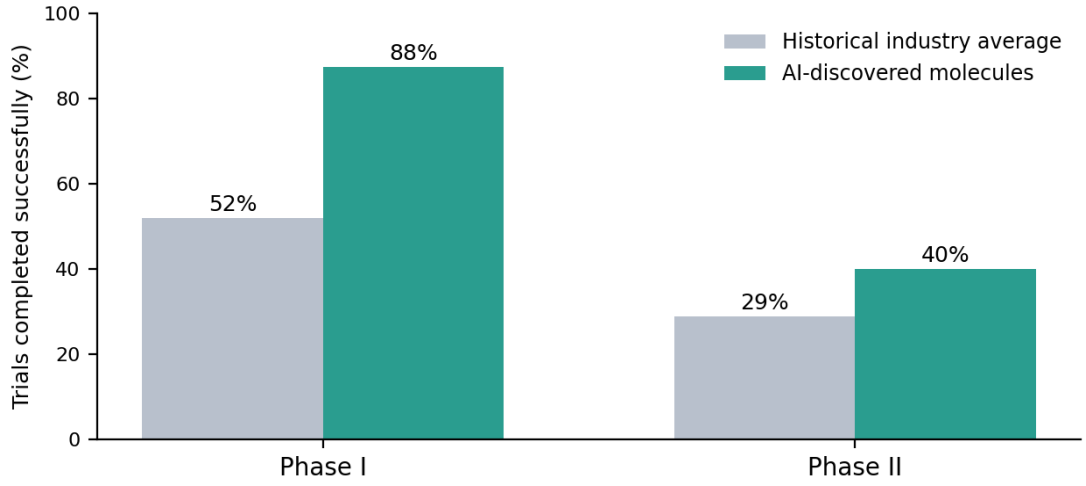


Figure 2. Clinical success rates: AI-discovered molecules versus historical averages. AI figures are 21 of 24 (Phase I) and about 4 of 10 (Phase II) completed trials (Jayatunga et al., 2024); historical figures from Wong et al. (2019).

6. Limits and Cautions

It would be easy to read the success stories as proof that the problem is solved. It is not. Models are only as good as the data behind them, and public datasets lean heavily towards well-studied targets and chemical families, so performance tends to drop on genuinely new biology. Many published results come from the companies that built the tools, and independent replication is still scarce. The Phase II samples are small enough that the numbers could shift as more trials report. Questions about patient-data privacy, regulatory acceptance of AI-derived evidence and the ownership of AI-designed molecules are also unsettled. At the time of writing, industry trackers report that no AI-discovered drug has yet completed the full path to regulatory approval, so the strongest claim anyone can make is that the early stages look faster, not that the whole process is solved.

7. Conclusion

AI is changing drug discovery in ways that are real but specific. It helps researchers sift enormous biological datasets for plausible targets, predict protein structures, design and screen molecules in silico, and screen out likely toxic compounds before the expensive work begins. The best current evidence, high Phase I success but only average Phase II success, suggests that these tools make candidate molecules safer and

ISSN: 3060-4737 Volume 3, Issue 10 IF(Impact Factor) 10.98 / 2026

better behaved without yet guaranteeing that they work in patients. The next stage of progress will depend less on cleverer algorithms and more on better data, transparent models and careful clinical validation. If those come together, AI will not replace the scientists, the patience or the clinical trials, but it may make the road to a new medicine noticeably shorter.

References

Abramson, J., Adler, J., Dunger, J., et al. (2024). Accurate structure prediction of biomolecular interactions with AlphaFold 3. Nature, 630, 493–500.

Jayatunga, M. K. P., Ayers, M., Bruens, L., Jayanth, D., & Meier, C. (2024). How successful are AI-discovered drugs in clinical trials? A first analysis and emerging lessons. Drug Discovery Today, 29(6), 104009. https://doi.org/10.1016/j.drudis.2024.104009

Jumper, J., Evans, R., Pritzel, A., et al. (2021). Highly accurate protein structure prediction with AlphaFold. Nature, 596, 583–589.

Mayr, A., Klambauer, G., Unterthiner, T., & Hochreiter, S. (2016). DeepTox: Toxicity prediction using deep learning. Frontiers in Environmental Science, 3, 80.

Paul, S. M., Mytelka, D. S., Dunwiddie, C. T., et al. (2010). How to improve R&D productivity: The pharmaceutical industry's grand challenge. Nature Reviews Drug Discovery, 9, 203–214.

Ren, F., Aliper, A., Chen, J., et al. (2025). A small-molecule TNIK inhibitor targets fibrosis in preclinical and clinical models. Nature Biotechnology, 43(1), 63–75. https://doi.org/10.1038/s41587-024-02143-0

Stokes, J. M., Yang, K., Swanson, K., et al. (2020). A deep learning approach to antibiotic discovery. Cell, 180(4), 688–702.

Sun, D., Gao, W., Hu, H., & Zhou, S. (2022). Why 90% of clinical drug development fails and how to improve it? Acta Pharmaceutica Sinica B, 12(7), 3049–3062. https://doi.org/10.1016/j.apsb.2022.02.002

Vamathevan, J., Clark, D., Czodrowski, P., et al. (2019). Applications of machine learning in drug discovery and development. Nature Reviews Drug Discovery, 18, 463–477.

Wong, C. H., Siah, K. W., & Lo, A. W. (2019). Estimation of clinical trial success rates and related parameters. Biostatistics, 20(2), 273–286.

ISSN: 3060-4737 Volume 3, Issue 10 IF(Impact Factor) 10.98 / 2026
Xu, Z., Ren, F., Wang, P., et al. (2025). A generative AI-discovered TNIK inhibitor for idiopathic pulmonary fibrosis: A randomized phase 2a trial. Nature Medicine. https://doi.org/10.1038/s41591-025-03743-2

Zhavoronkov, A., Ivanenkov, Y. A., Aliper, A., et al. (2019). Deep learning enables rapid identification of potent DDR1 kinase inhibitors. Nature Biotechnology, 37, 1038–1040.