

Advanced Artificial Intelligence and data science in bioinformatics-driven drug discovery for cancer: Pathways toward shorter and less toxic treatment

Yejiide Eniola Dabiri *

Department of Information Technology, University of Cumberlands, Kentucky, United States

Magna Scientia Advanced Research and Reviews, 2026, 17(02), 001–010

Publication history: Received on 25 May 2026; revised on 30 June 2026; accepted on 02 July 2026

Article DOI: <https://doi.org/10.30574/msarr.2026.17.2.0122>

Abstract

Cancer remains one of the leading causes of death worldwide, with the GLOBOCAN estimates placing the 2022 global burden at close to 20 million new cases and 9.7 million deaths (Bray et al., 2024), a burden projected by the American Cancer Society (2024) to rise to roughly 35 million annual cases by 2050. Conventional cytotoxic chemotherapy, though still central to treatment for many tumor types, is frequently associated with prolonged treatment courses, non-specific systemic toxicity, and reduced quality of life. This review synthesizes recent literature on the application of artificial intelligence (AI) and data science within bioinformatics-driven cancer drug discovery, examining how these tools are reshaping target identification, molecular design, biomarker discovery, and treatment personalization. The analysis shows that deep learning-based protein structure prediction (Jumper et al., 2021), generative molecular design (Gangwal & Lavecchia, 2024), multi-omics target identification (Bhinder et al., 2021; Wei et al., 2023), digital pathology and radiomics (Bera et al., 2022; Lu et al., 2024), and machine learning models for predicting chemotherapy toxicity (Huang et al., 2024; Moslemi et al., 2025) are collectively shortening discovery timelines, improving the precision of treatment selection, and reducing treatment-related adverse effects in reported studies. Case evidence is presented, including a generative-AI-designed molecule that reached Phase I clinical trials in under 30 months (Insilico Medicine, 2022) and the 2024 Nobel Prize in Chemistry awarded for the AlphaFold protein-structure-prediction system. While these advances present a credible pathway toward shorter, more targeted, and less toxic cancer treatment, and in specific molecular contexts may reduce reliance on conventional chemotherapy, the evidence does not yet support claims that AI will universally eliminate chemotherapy; rather, it points toward an increasingly personalized standard of oncologic care. The review concludes by discussing the ethical, regulatory, and data-governance barriers that must be addressed for these gains to be realized safely and equitably.

Keywords: Artificial Intelligence; Bioinformatics; Drug Discovery; Cancer Treatment; Precision Oncology; Chemotherapy Toxicity; Data Science

1. Introduction

Cancer is a heterogeneous group of diseases characterized by uncontrolled cellular proliferation, and it continues to impose an enormous and growing burden on global health systems. According to the most recent GLOBOCAN estimates, approximately 20 million new cancer cases and 9.7 million cancer deaths occurred worldwide in 2022, with roughly one in five people expected to develop cancer during their lifetime (Bray et al., 2024). The American Cancer Society (2024) further projects that annual incidence will climb to about 35 million cases by 2050, driven by population growth, ageing, and shifting risk-factor exposure. For decades, the therapeutic backbone of oncology has combined surgery, radiotherapy, and cytotoxic chemotherapy, the latter acting broadly on rapidly dividing cells and consequently producing well-documented adverse effects, including myelosuppression, neurotoxicity, cardiotoxicity, gastrointestinal injury, and mucositis (Agheli et al., 2024; Zhou et al., 2024).

* Corresponding author: Yejiide Dabiri

The convergence of artificial intelligence (AI), machine learning (ML), deep learning (DL), and bioinformatics has begun to reshape this paradigm. The explosive growth of genomic, transcriptomic, proteomic, and clinical data, coupled with advances in computational architecture such as transformers and graph neural networks, has created the conditions for AI to meaningfully accelerate drug discovery and enable more precise, less toxic cancer treatment (Askr et al., 2023; Hasselgren & Oprea, 2024). Data-driven approaches now support nearly every stage of the oncology drug-development pipeline, from target identification and protein structure prediction to de novo molecular design, toxicity prediction, and clinical trial optimization (Bhinder et al., 2021; Gangwal & Lavecchia, 2024; You et al., 2022).

This review examines how AI and data science, applied within a bioinformatics framework, are being leveraged to shorten cancer drug discovery timelines, personalize treatment selection, and reduce the side-effect burden traditionally associated with chemotherapy. It further considers, on the basis of the current evidence, the extent to which these technologies may plausibly reduce dependence on conventional chemotherapy for specific cancer subtypes, while cautioning against overstating what remains an evolving and unevenly distributed body of evidence.

1.1. Problem Statement and Rationale

Traditional small-molecule drug discovery is notoriously slow, expensive, and failure-prone. Widely cited industry estimates place the average cost of bringing a single new drug to market at well over two billion dollars and the average development timeline at ten to fifteen years, with most candidate compounds failing during preclinical or clinical development (DiMasi et al., 2016). In oncology specifically, this inefficiency is compounded by tumor heterogeneity, evolving drug resistance, and the narrow therapeutic index of many cytotoxic agents, which limits dose intensity and drives the adverse-effect profile that patients must tolerate (Adamopoulos et al., 2024; Pandiyan & Wang, 2022).

Chemotherapy, despite decades of refinement, remains a relatively blunt instrument: because most cytotoxic agents act on rapidly dividing cells generally rather than cancer cells specifically, healthy tissues such as bone marrow, gastrointestinal epithelium, and hair follicles are also affected, producing the fatigue, nausea, infection risk, and organ toxicity that define the patient experience of treatment (Agheli et al., 2024). A substantial research effort has therefore focused on using AI and bioinformatics to (a) identify more selective molecular targets, (b) design molecules and biologics with greater specificity, (c) predict which patients are likely to experience severe toxicity so that dosing can be personalized, and (d) identify existing or repurposable drugs and combination regimens that could reduce or replace generalized cytotoxic chemotherapy for defined patient subgroups (Rehman et al., 2024; Zhou et al., 2024).

Aim and Objectives

This review aims to critically examine how AI and data science, applied through bioinformatics, are transforming cancer drug discovery and treatment. The specific objectives are to:

- Identify the principal AI and data-science methods currently applied across the oncology drug-discovery pipeline;
- Present documented empirical and case evidence of efficiency gains, timeline compression, and toxicity reduction;
- Assess, on the basis of the evidence, the realistic prospects for shorter-duration, chemotherapy-sparing cancer treatment; and
- Identify the ethical, regulatory, and data-governance challenges that must be addressed for these gains to be realized responsibly.

1.2. Approach

This review adopts a narrative, integrative approach consistent with prior AI-in-oncology reviews (Wei et al., 2023). Peer-reviewed literature published mainly between 2020 and 2025 was drawn from sources including PubMed/PMC, Nature-family journals, ScienceDirect, and MDPI journals, supplemented by verifiable industry and press documentation describing AI-designed drug candidates that have entered clinical trials. Sources were synthesized thematically around the stages of the drug-discovery-to-treatment continuum shown in Figure 1.

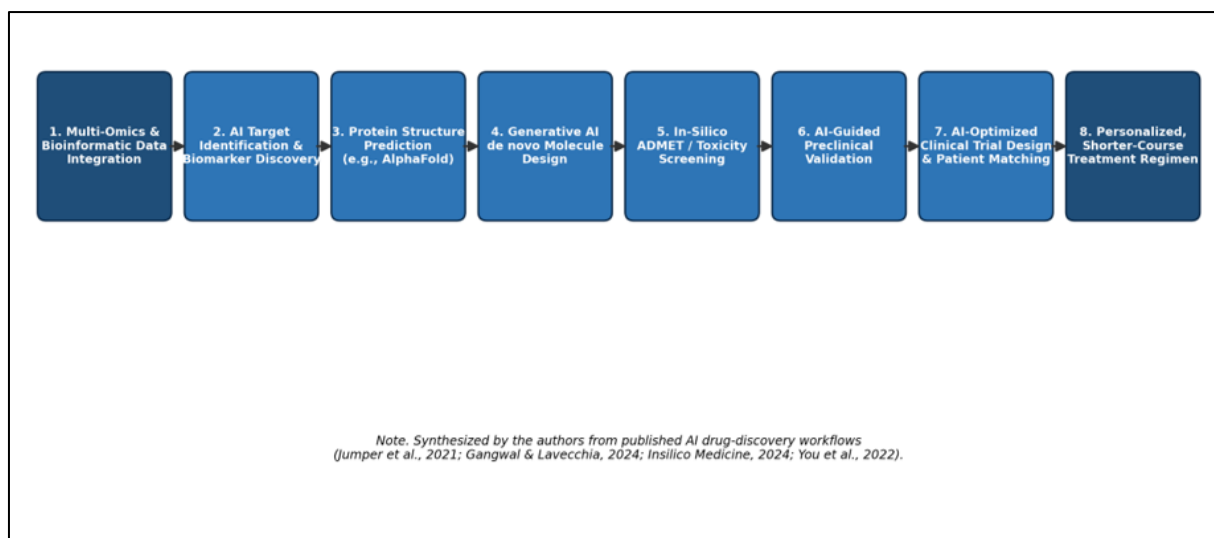


Figure 1 Conceptual pipeline for AI- and bioinformatics-driven cancer drug discovery, from multi-omics data integration through personalized treatment delivery.

2. AI and Data Science Applications Across the Bioinformatics Drug-Discovery Continuum

2.1. Multi-Omics Integration and AI-Driven Target Identification

The starting point of modern oncology drug discovery is the identification of druggable molecular targets from the vast volume of genomic, transcriptomic, epigenomic, and proteomic data generated by next-generation sequencing and mass spectrometry (Wei et al., 2023). Machine learning methods, including random forests, support vector machines, and increasingly deep multimodal neural networks, are used to integrate these heterogeneous "omics" layers to identify mutations, gene-expression signatures, and pathway alterations that drive tumor growth (Bhinder et al., 2021; He et al., 2023). Capobianco (2022) notes that the high dimensionality of cancer genomic data, often far exceeding sample size makes AI-based dimensionality reduction and feature-selection techniques indispensable for robust biomarker discovery. Protein-language models such as ESM, trained on hundreds of millions of sequences, extend this capability by learning evolutionary and structural information directly from amino-acid sequences, aiding the identification of functionally important mutations (Rives et al., 2021).

At the pan-cancer level, deep learning models such as multimodal neural networks and transformers have been used to identify shared and cancer-type-specific molecular features, enabling cross-cancer subtype clustering and drug-response prediction, although such models require large, well-curated datasets to avoid bias from imbalanced representation across cancer types (Wei et al., 2023). You et al. (2022) similarly describe how AI-enabled target identification has moved beyond single-gene analysis toward network- and pathway-level reasoning that better reflects the systems biology of tumorigenesis.

2.2. Protein Structure Prediction and Structure-Based Drug Design

A foundational bottleneck in structure-based drug design has historically been the difficulty of determining three-dimensional protein structures experimentally. This bottleneck was substantially reduced by AlphaFold, a deep learning system that predicts protein structure from amino-acid sequence with an accuracy approaching that of experimental methods such as X-ray crystallography (Jumper et al., 2021). The achievement was recognized with the 2024 Nobel Prize in Chemistry, awarded jointly to the AlphaFold developers and to David Baker for computational protein design (Chemistry Nobel goes to developers of AlphaFold AI, 2024). The successor model, AlphaFold3, extends prediction to complexes involving proteins, nucleic acids, and small-molecule ligands, directly supporting drug-target interaction modeling (Abramson et al., 2024).

In oncology specifically, AlphaFold-predicted structures have already been combined with generative chemistry engines to identify novel inhibitors. For example, Insilico Medicine (2024) reported the use of AlphaFold-predicted structures of salt-inducible kinase 2 (SIK2), a target overexpressed in roughly 30% of human ovarian cancers, to design and prioritize a small library of candidate inhibitors, from which compounds with favorable potency and selectivity profiles were identified. McGibbon et al. (2024) review the parallel evolution of small-molecule representation methods,

from simple molecular fingerprints to learned graph and geometric representations that make such structure-based generative design increasingly accurate.

2.3. Generative AI and De Novo Molecular Design

Generative AI models, including generative adversarial networks, variational autoencoders, transformer-based architectures, and reinforcement learning agents, can propose entirely novel molecular structures optimized against multiple objectives (binding affinity, synthesizability, metabolic stability, and toxicity) rather than merely modifying known compounds (Gangwal & Lavecchia, 2024; Blanco-González et al., 2023). Comprehensive reviews of computer-aided and AI-assisted drug design describe how these generative engines are now integrated into end-to-end discovery platforms spanning target discovery, molecule generation, and outcome prediction (Vemula et al., 2023; Askr et al., 2023).

The most extensively documented case example is Insilico Medicine's Pharma.AI platform, which links a target-discovery engine (PandaOmics) to a generative chemistry engine (Chemistry42) and a clinical-outcome prediction engine (InClinico). Using this pipeline, the company took a novel anti-fibrotic small molecule from target identification to a Phase I clinical trial in under 30 months, reportedly at a fraction of conventional cost (Insilico Medicine, 2022, 2023). The same underlying platform has since been applied to oncology, including a clinical-stage USP1 inhibitor (ISM3091) targeting BRCA1/2-mutant solid tumors and additional preclinical programs for MTAP-deleted cancers and hormone-receptor-positive breast cancer (Insilico Medicine, 2024). Independent reporting corroborates that the approach compresses design-build-test-learn cycles that traditionally require years into months (NVIDIA, 2024).

2.4. AI-Driven Drug Repurposing for Oncology

Because developing an entirely new oncology drug is costly and slow, AI-based drug repurposing, identifying new anticancer applications for drugs already approved for other indications, or for natural products has emerged as a complementary, faster-to-clinic strategy (Pandiyan & Wang, 2022; Rehman et al., 2024). Liang et al. (2020) describe how AI models trained on large-scale pharmacogenomic, and cell-line datasets can screen thousands of existing compounds against novel cancer targets far more rapidly than traditional high-throughput screening alone. Natural-product-derived compounds, historically difficult to characterize structurally, are similarly benefiting from AI-assisted dereplication and structure elucidation pipelines (Mullowney et al., 2023). Chen et al. (2023) note that repurposing candidates carry an inherent safety advantage, since their pharmacokinetic and toxicity profiles are already partly characterized in humans, potentially shortening the path to clinical use relative to entirely novel chemical entities.

2.5. Digital Pathology, Radiomics, and Immunotherapy-Response Prediction

Beyond molecule discovery, AI is transforming how treatment decisions are made once a diagnosis is established. Deep learning applied to radiomic features extracted from CT, MRI, and PET imaging can predict tumor behavior, treatment response, and outcome with a precision that complements or exceeds conventional radiological assessment (Bera et al., 2022). In digital pathology, foundation models trained on large collections of whole-slide histology images can infer tumor-microenvironment composition, immune-cell infiltration, and molecular subtype directly from routinely collected hematoxylin-and-eosin slides, without requiring costly additional sequencing (Lu et al., 2024; Ma et al., 2024). Liang et al. (2023) demonstrated that deep learning applied to histopathological images can discover prognostic biomarkers for liver cancer that were not previously identified through conventional analysis.

These capabilities are increasingly used to predict which patients will respond to immune checkpoint inhibitors and other immunotherapies, allowing clinicians to reserve broad cytotoxic chemotherapy for patients unlikely to benefit from more targeted options (Khalighi et al., 2024). Because immunotherapies and targeted agents generally carry a different, and often more favorable, toxicity profile than cytotoxic chemotherapy, better patient selection through AI-based imaging and pathology analysis represents one of the more concrete near-term mechanisms by which AI can reduce a population's overall exposure to chemotherapy-related harm.



Figure 3 Core application domains of AI and data science across the oncology drug-discovery and treatment continuum.

2.6. Predictive Toxicology and Personalized Dosing: Toward Reduced Chemotherapy Side Effects

A rapidly growing body of work applies machine learning directly to the problem of chemotherapy-related toxicity. Huang et al. (2024) developed an ensemble machine learning framework, combining balanced bagging classifiers with gradient boosting to predict drug-induced agranulocytosis, achieving an external validation area-under-the-curve (AUC) of 0.916 and identifying structural and molecular features associated with toxicity risk before treatment begins. In breast cancer, Moslemi et al. (2025) used CT textural features combined with machine learning to predict chemotherapy response prior to treatment initiation in locally advanced disease, comparing multiple feature-selection strategies to optimize predictive accuracy. Agheli et al. (2024) integrated radiomic, dosimetric, and clinical features to predict severe radiation-induced oral mucositis in head and neck cancer patients receiving combined chemoradiotherapy, with the explicit aim of enabling personalized treatment planning that minimizes toxicity. Zhou et al. (2024) built a chemotherapy-response prediction model based on hypoxia-related gene expression in oral cancer, while Xiang et al. (2025) combined machine learning with bulk and single-cell RNA sequencing to characterize cell-death patterns predictive of neoadjuvant chemotherapy outcomes in breast cancer.

Collectively, these studies illustrate a consistent pattern: rather than eliminating chemotherapy outright, AI models are enabling clinicians to (a) identify, before treatment, which patients are likely to experience severe toxicity; (b) personalize dose intensity and regimen selection to the individual's genomic and clinical profile; and (c) in some cases substitute cytotoxic agents with better-tolerated targeted therapies when biomarkers indicate a favorable predicted response. Figure 2 summarizes illustrative efficiency and performance figures drawn from this literature.

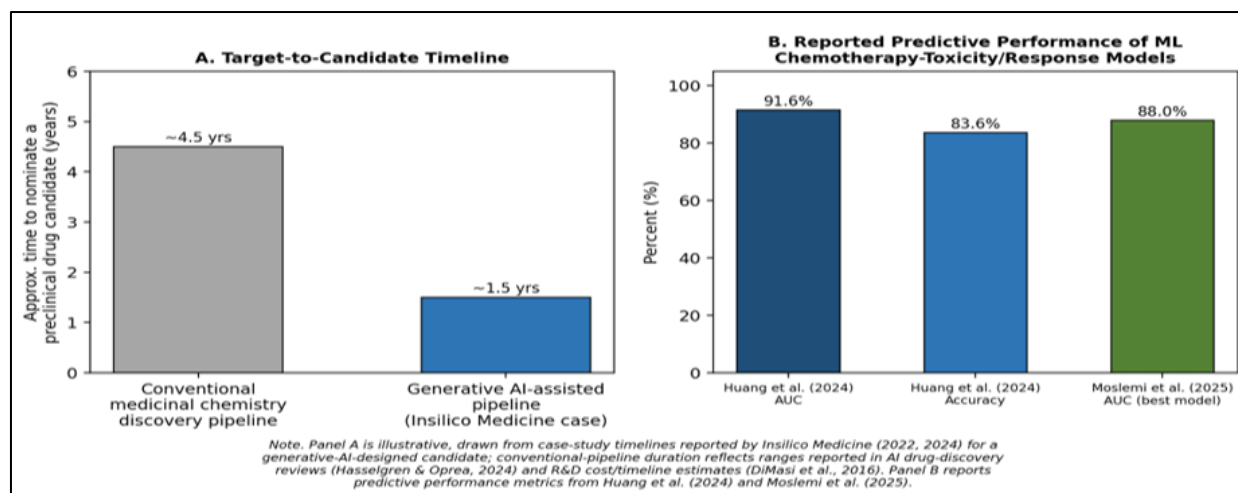


Figure 2 Illustrative timeline compression reported for an AI-assisted drug-discovery case study (Panel A) and predictive performance reported for machine learning chemotherapy-toxicity/response models (Panel B).

2.7. AI-Optimized Clinical Trial Design and Patient Matching

Clinical trials represent one of the most time- and cost-intensive stages of oncology drug development, frequently constrained by slow patient recruitment and high rates of protocol amendment. AI-based tools are increasingly used to optimize trial design, stratify patients by molecular and clinical features, and predict the probability that a given trial will meet its endpoints. Insilico Medicine's InClinico platform, for instance, uses transcriptomic, chemical, and clinical-trial-protocol data to forecast Phase II outcomes, and in a 2022 validation exercise correctly anticipated the outcome of a majority of a sample of historical oncology and non-oncology trials (Insilico Medicine, 2023). More broadly, Corti et al. (2023) argue that AI-enabled precision medicine tools must be paired with attention to equitable trial recruitment, since biased training data can otherwise translate into biased trial populations and, ultimately, into therapies that are less effective or less safe for underrepresented groups.

3. Empirical Grounding

Table 1 Selected empirical and case-study evidence discussed above, illustrating the range of AI applications and their reported outcomes across the cancer drug-discovery and treatment continuum.

Source	Focus / AI Method	Key Reported Finding
<i>Jumper et al. (2021)</i>	Deep learning protein-structure prediction (AlphaFold)	Predicted protein structures approaching experimental accuracy, since applied to more than 200 million proteins.
<i>Insilico Medicine (2022, 2023)</i>	End-to-end generative AI pipeline (PandaOmics, Chemistry42, InClinico)	Novel small molecule taken from target discovery to Phase I trial in under 30 months.
<i>Insilico Medicine (2024)</i>	AlphaFold + generative chemistry for SIK2 (ovarian cancer target)	8,918 candidate molecules generated; selective inhibitors identified with favorable potency.
<i>Huang et al. (2024)</i>	Ensemble ML toxicity prediction	External validation AUC of 0.916 for predicting drug-induced agranulocytosis toxicity.
<i>Moslemi et al. (2025)</i>	CT-radiomics + ML chemotherapy-response prediction (breast cancer)	Pre-treatment imaging features predicted eventual chemotherapy response.
<i>Agheli et al. (2024)</i>	Radiomic + clinical ML toxicity model	Model predicted severe radiation-induced oral mucositis to enable personalized planning.
<i>Bera et al. (2022)</i>	Radiomics and AI in radiology	AI-derived imaging features improve prediction of cancer outcomes and treatment response.

<i>Wei et al. (2023)</i>	Multiomics integration review	DL/ML methods enable cross-cancer subtyping and drug-response prediction from multiomics data.
<i>Adamopoulos et al. (2024)</i>	AI/ML in anti-cancer drug discovery	DL models elucidated resistance mechanisms to CDK4/6 inhibitors in breast cancer.
<i>Faheem & Dutta (2023)</i>	Case analysis of IBM Watson for Oncology	Documented failure attributed to biased/limited training data, underscoring implementation risk.

4. Toward Shorter, Less Toxic, and Chemotherapy-Sparing Cancer Care: Prospects and Realistic Limits

The evidence reviewed above supports a cautiously optimistic, but bounded, conclusion. AI and data science are demonstrably compressing discovery timelines (Insilico Medicine, 2022; NVIDIA, 2024), improving the precision with which molecular targets and patient subgroups are identified (Bhinder et al., 2021; You et al., 2022), and enabling toxicity to be anticipated and mitigated before it occurs (Huang et al., 2024; Agheli et al., 2024). In defined molecular contexts, for example, BRCA1/2-mutant tumors targeted by USP1 inhibition, or tumors with biomarkers predictive of strong immunotherapy response, these advances plausibly allow clinicians to reduce or avoid broad cytotoxic chemotherapy in favor of more targeted, better-tolerated regimens (Insilico Medicine, 2024; Khalighi et al., 2024).

At the same time, the claim that AI will eliminate chemotherapy altogether or reduce all cancer treatment to a uniformly short course, is not supported by the current evidence base. Adamopoulos et al. (2024) note that even AI-informed targeted therapies encounter acquired resistance, as illustrated by deep-learning analyses of resistance mechanisms to CDK4/6 inhibitors in breast cancer, meaning that combination strategies, sometimes still including cytotoxic agents will likely remain necessary for many patients. Cancer's biological heterogeneity, both between patients and within a single tumor over time, means that no single technological advance is likely to make chemotherapy universally obsolete in the near term. A more defensible characterization of the trajectory is that AI and bioinformatics are enabling a shift from uniform, broadly toxic regimens toward increasingly personalized treatment sequences in which chemotherapy, when still required, is better targeted, better dosed, and better timed and, for a growing subset of molecularly defined cancers, can be substantially reduced or replaced by targeted or immune-based therapies identified and optimized with AI assistance.

5. Ethical, Regulatory, and Data-Governance Challenges

Realizing these benefits responsibly requires confronting significant ethical and regulatory challenges. Algorithmic bias arising from non-representative training data remains a central concern; the well-documented failure of IBM Watson for Oncology, whose recommendations were found in some cases to reflect the biases and limitations of its training data rather than robust clinical evidence, is frequently cited as a cautionary case for the field (Faheem & Dutta, 2023). Data privacy is similarly prominent, since AI models in oncology are typically trained on large volumes of sensitive genomic and clinical data; safeguarding this information while still enabling the data sharing necessary for model development requires careful governance frameworks and, in many jurisdictions, compliance with regulations such as the General Data Protection Regulation and oversight by bodies such as the U.S. Food and Drug Administration for AI-enabled medical devices.

Corti et al. (2023) further argue that without deliberate attention to equity, AI-driven precision oncology risks widening existing disparities in cancer care, since the datasets underlying many models are drawn disproportionately from well-resourced health systems and specific population groups. Model interpretability, or the so-called "black box" problem, also remains a barrier to clinical trust and regulatory approval, particularly for deep learning models whose internal reasoning is not easily explained to clinicians or patients. Addressing these challenges will require continued investment in explainable AI methods, bias auditing, diverse and representative data collection, and regulatory frameworks that keep pace with the technology's rapid evolution.

6. Conclusion

This review has examined how artificial intelligence and data science, applied within a bioinformatics framework, are reshaping cancer drug discovery and treatment. Across target identification, protein structure prediction, generative molecular design, drug repurposing, digital pathology, toxicity prediction, and clinical trial optimization, the literature converges on a consistent finding: AI is measurably accelerating discovery timelines, improving the precision of patient

and target selection, and, in numerous documented studies, reducing the incidence and severity of treatment-related toxicity. Case evidence from AlphaFold's Nobel-recognized breakthrough in protein structure prediction to generative-AI-designed molecules reaching human trials in under three years demonstrates that these are not merely theoretical possibilities but active, ongoing developments in the field.

At the same time, this review cautions against the expectation that these advances will imminently or universally eliminate chemotherapy. The more accurate and evidence-grounded conclusion is that AI and bioinformatics are driving a transition toward increasingly personalized cancer care, in which treatment, including chemotherapy where it remains necessary is more precisely targeted, better dosed, and, for a growing number of molecularly defined cancers, substantially shortened or substituted with better-tolerated targeted and immune-based therapies. Realizing this potential equitably and safely will require sustained attention to data quality, algorithmic bias, patient privacy, and regulatory oversight alongside continued technical innovation.

References

- [1] Abramson, J., Adler, J., Dunger, J., Evans, R., Green, T., Pritzel, A., Ronneberger, O., ... Jumper, J. M. (2024). Accurate structure prediction of biomolecular interactions with AlphaFold3. *Nature*, 630, 493–500. <https://doi.org/10.1038/s41586-024-07487-w>
- [2] Adamopoulos, C., Papavassiliou, K. A., & Papavassiliou, A. G. (2024). Fusing artificial intelligence and machine learning for anti-cancer drug discovery. *Cancers*, 16(20), 3522. <https://doi.org/10.3390/cancers16203522>
- [3] Agheli, R., Siavashpour, Z., Reiazi, R., Azghandi, S., Cheraghi, S., & Paydar, R. (2024). Predicting severe radiation-induced oral mucositis in head and neck cancer patients using integrated baseline CT radiomic, dosimetry, and clinical features: A machine learning approach. *Heliyon*, 10(2), e24866. <https://doi.org/10.1016/j.heliyon.2024.e24866>
- [4] American Cancer Society. (2024). Global cancer statistics, 2024 [Press release]. <https://pressroom.cancer.org/GlobalCancerStatistics2024>
- [5] Askr, H., Elgeldawi, E., Aboul Ella, H., Elshaier, Y. A. M. M., Gomaa, M. M., & Hassanien, A. E. (2023). Deep learning in drug discovery: An integrative review and future challenges. *Artificial Intelligence Review*, 56(7), 5975–6037. <https://doi.org/10.1007/s10462-022-10306-1>
- [6] Bera, K., Braman, N., Gupta, A., Velcheti, V., & Madabhushi, A. (2022). Predicting cancer outcomes with radiomics and artificial intelligence in radiology. *Nature Reviews Clinical Oncology*, 19, 132–146.
- [7] Bhinder, B., Gilvary, C., Madhukar, N. S., & Elemento, O. (2021). Artificial intelligence in cancer research and precision medicine. *Cancer Discovery*, 11(4), 900–915. <https://doi.org/10.1158/2159-8290.CD-21-0090>
- [8] Blanco-González, A., Cabezón, A., Seco-González, A., Conde-Torres, D., Antelo-Riveiro, P., Piñeiro, Á., & Garcia-Fandino, R. (2023). The role of AI in drug discovery: Challenges, opportunities, and strategies. *Pharmaceuticals*, 16(6), 891. <https://doi.org/10.3390/ph16060891>
- [9] Bray, F., Laversanne, M., Sung, H., Ferlay, J., Siegel, R. L., Soerjomataram, I., & Jemal, A. (2024). Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: A Cancer Journal for Clinicians*, 74(3), 229–263. <https://doi.org/10.3322/caac.21834>
- [10] Capobianco, E. (2022). High-dimensional role of AI and machine learning in cancer research. *British Journal of Cancer*, 126(4), 523–532.
- [11] Carracedo-Reboredo, P., Liñares-Blanco, J., Rodríguez-Fernández, N., Cedrón, F., Novoa, F. J., Carballal, A., Maojo, V., Pazos, A., & Fernandez-Lozano, C. (2021). A review on machine learning approaches and trends in drug discovery. *Computational and Structural Biotechnology Journal*, 19, 4538–4558.
- [12] Chemistry Nobel goes to developers of AlphaFold AI that predicts protein structures. (2024). *Nature*. <https://doi.org/10.1038/d41586-024-03214-7>
- [13] Chen, W., Liu, X., Zhang, S., & Chen, S. (2023). Artificial intelligence for drug discovery: Resources, methods, and applications. *Molecular Therapy — Nucleic Acids*, 31, 691–702.
- [14] Corti, C., Cobanaj, M., Marian, F., Dee, E. C., Lloyd, M. R., Marcu, S., Curigliano, G., et al. (2023). Artificial intelligence in cancer research and precision medicine: Applications, limitations and priorities to drive transformation in the delivery of equitable and unbiased care. *Cancer Treatment Reviews*, 112, 102498.

- [15] 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. <https://doi.org/10.1016/j.jhealeco.2016.01.012>
- [16] Faheem, H., & Dutta, S. (2023). Artificial intelligence failure at IBM Watson for Oncology. *IUP Journal of Knowledge Management*, 21(3), 47–75.
- [17] Gangwal, A., & Lavecchia, A. (2024). Unleashing the power of generative AI in drug discovery. *Drug Discovery Today*, 29(6), 103992. <https://doi.org/10.1016/j.drudis.2024.103992>
- [18] Hasselgren, C., & Oprea, T. I. (2024). Artificial intelligence for drug discovery: Are we there yet? *Annual Review of Pharmacology and Toxicology*, 64, 527–550. <https://doi.org/10.1146/annurev-pharmtox-040323-040828>
- [19] He, X., Liu, X., Zuo, F., Shi, H., & Jing, J. (2023). Artificial intelligence-based multi-omics analysis fuels cancer precision medicine. *Seminars in Cancer Biology*, 88, 187–200.
- [20] Huang, X., Xie, X., Huang, S., Wu, S., & Huang, L. (2024). Predicting non-chemotherapy drug-induced agranulocytosis toxicity through ensemble machine learning approaches. *Frontiers in Pharmacology*, 15, 1431941. <https://doi.org/10.3389/fphar.2024.1431941>
- [21] Insilico Medicine. (2022). From start to Phase 1 in 30 months: AI-discovered and AI-designed anti-fibrotic drug enters Phase I clinical trial. <https://insilico.com/phase1>
- [22] Insilico Medicine. (2023). First generative AI drug begins Phase II trials with patients. https://insilico.com/blog/first_phase2
- [23] Insilico Medicine. (2024). Insilico Medicine leverages generative AI and AlphaFold structures for novel cancer drug [Press release]. *EurekAlert!* <https://www.eurekalert.org/news-releases/997469>
- [24] Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., ... Hassabis, D. (2021). Highly accurate protein structure prediction with AlphaFold. *Nature*, 596, 583–589. <https://doi.org/10.1038/s41586-021-03819-2>
- [25] Khalighi, S., Reddy, K., Midya, A., Pandav, K. B., Madabhushi, A., & Abedalthagafi, M. (2024). Artificial intelligence in neuro-oncology: Advances and challenges in brain tumor diagnosis, prognosis, and precision treatment. *NPJ Precision Oncology*, 8(1), 80.
- [26] Liang, G., Fan, W., Luo, H., & Zhu, X. (2020). The emerging roles of artificial intelligence in cancer drug development and precision therapy. *Biomedicine & Pharmacotherapy*, 128, 110255.
- [27] Liang, J., Zhang, W., Yang, J., Wu, M., Dai, Q., Yin, H., Xiao, Y., & Kong, L. (2023). Deep learning supported discovery of biomarkers for clinical prognosis of liver cancer. *Nature Machine Intelligence*, 5(4), 408–420.
- [28] Lu, M. Y., Chen, B., Williamson, D. F. K., Chen, R. J., Liang, I., Ding, T., Jaume, G., Odintsov, I., Le, L. P., Gerber, G., Parwani, A. V., Zhang, A., & Mahmood, F. (2024). A visual-language foundation model for computational pathology. *Nature Medicine*, 30(3), 863–874.
- [29] Ma, J., He, Y., Li, F., Han, L., You, C., & Wang, B. (2024). Segment anything in medical images. *Nature Communications*, 15(1), 654.
- [30] McGibbon, M., Shave, S., Dong, J., Gao, Y., Houston, D. R., Xie, J., Yan, Y., Ren, J., & Feng, Y. (2024). From intuition to AI: Evolution of small molecule representations in drug discovery. *Briefings in Bioinformatics*, 25(1), bbad422.
- [31] Moslemi, A., Osapoetra, L. O., Dasgupta, A., Halstead, S., Alberico, D., Trudeau, M., Gandhi, S., Eisen, A., Wright, F., Look-Hong, N., Curpen, B., Kolios, M., & Czarnota, G. J. (2025). Prediction of chemotherapy response in locally advanced breast cancer patients at pre-treatment using CT textural features and machine learning: Comparison of feature selection methods. *Tomography*, 11(3), 33. <https://doi.org/10.3390/tomography11030033>
- [32] Mulleney, M. W., Duncan, K. R., Elsayed, S. S., Garg, N., van der Hooft, J. J. J., Martin, N. I., ... Linington, R. G. (2023). Artificial intelligence for natural product drug discovery. *Nature Reviews Drug Discovery*, 22(11), 895–916.
- [33] Pandiyan, S., & Wang, L. (2022). A comprehensive review on recent approaches for cancer drug discovery associated with artificial intelligence. *Computers in Biology and Medicine*, 150, 106140.
- [34] Rehman, A. U., Li, M., Wu, B., Ali, Y., Rasheed, S., Shaheen, S., Liu, X., Luo, R., & Zhang, J. (2024). Role of artificial intelligence in revolutionizing drug discovery. *Fundamental Research*, 5(3), 1273–1287.
- [35] Rives, A., Meier, J., Sercu, T., Goyal, S., Lin, Z., Liu, J., Guo, D., Ott, M., Zitnick, C. L., Ma, J., & Fergus, R. (2021). Biological structure and function emerge from scaling unsupervised learning to 250 million protein sequences. *Proceedings of the National Academy of Sciences*, 118(15), e2016239118.

- [36] Vemula, D., Jayasurya, P., Sushmitha, V., Kumar, Y. N., & Bhandari, V. (2023). CADD, AI and ML in drug discovery: A comprehensive review. *European Journal of Pharmaceutical Sciences*, 181, 106324.
- [37] Wei, L., Niraula, D., Gates, E. D. H., Fu, J., Luo, Y., Nyflot, M. J., Bowen, S. R., El Naqa, I. M., & Cui, S. (2023). Artificial intelligence (AI) and machine learning (ML) in precision oncology: A review on enhancing discoverability through multiomics integration. *British Journal of Radiology*, 96(1150), 20230211. <https://doi.org/10.1259/bjr.20230211>
- [38] Xiang, L., Yang, J., Rao, J., Ma, A., Liu, C., Zhang, Y., Huang, A., Xie, T., Xue, H., Chen, Z., Yuan, J., & Yan, H. (2025). Integrating machine learning and bulk and single-cell RNA sequencing to decipher diverse cell death patterns for predicting the prognosis of neoadjuvant chemotherapy in breast cancer. *International Journal of Molecular Sciences*, 26(8), 3682. <https://doi.org/10.3390/ijms26083682>
- [39] You, Y., Lai, X., Pan, Y., Zheng, H., Vera, J., Liu, S., Deng, S., & Zhang, L. (2022). Artificial intelligence in cancer target identification and drug discovery. *Signal Transduction and Targeted Therapy*, 7, 156.
- [40] Zhou, C., Jia, H., Jiang, N., Zhao, J., & Nan, X. (2024). Establishment of chemotherapy prediction model based on hypoxia-related genes for oral cancer. *Journal of Cancer*, 15(15), 4890–4901. <https://doi.org/10.7150/jca.96654>