Molecular Cancer

REVIEW

Open Access

Emerging artificial intelligence-driven precision therapies in tumor drug resistance: recent advances, opportunities, and challenges

Yuan Mao[1,5], Dangang Shangguan[2], Qi Huang[3], Ling Xiao[4], Dongsheng Cao[1], Hui Zhou[5,6*] and Yi-Kun Wang[1*]

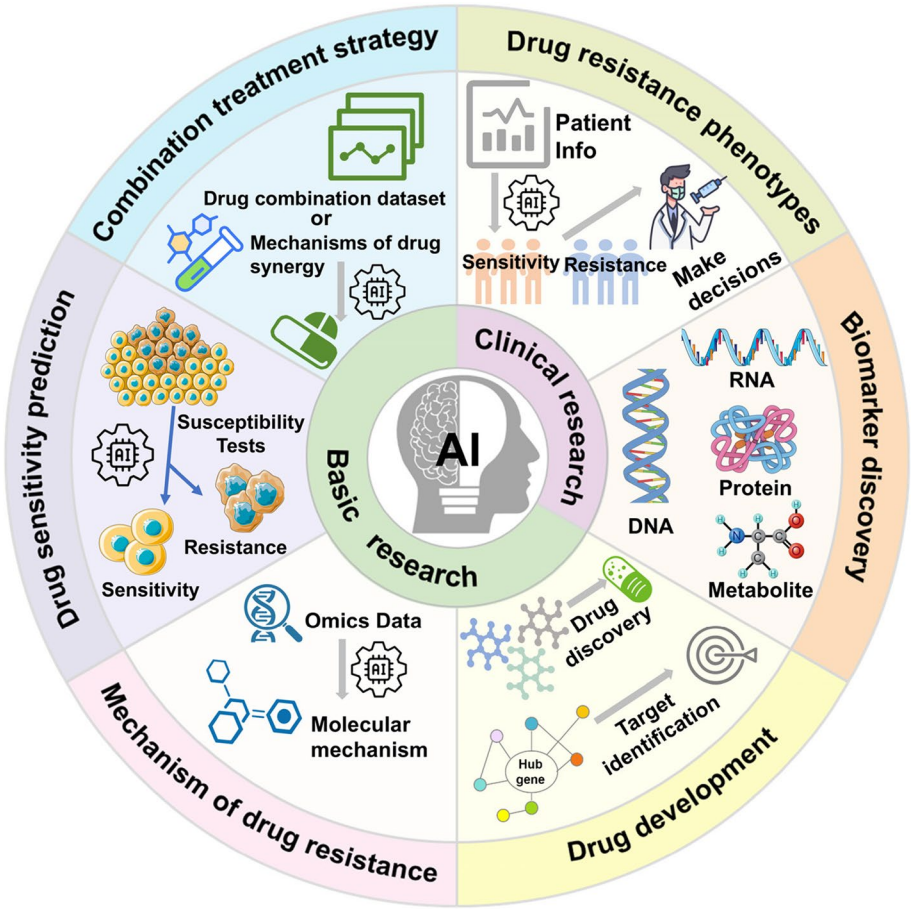
Abstract

Drug resistance is one of the main reasons for cancer treatment failure, leading to a rapid recurrence/disease progression of the cancer. Recently, artificial intelligence (AI) has empowered physicians to use its powerful data processing and pattern recognition capabilities to extract and mine valuable drug resistance information from large amounts of clinical or omics data, to study drug resistance mechanisms, to evaluate and predict drug resistance, and to develop innovative therapeutic strategies to reduce drug resistance. In this review, we proposed a feasible workflow for incorporating AI into tumor drug resistance research, highlighted current AI-driven tumor drug resistance applications, and discussed the opportunities and challenges encountered in the process. Based on a comprehensive literature analysis, we systematically summarized the role of AI in tumor drug resistance research, including drug development, resistance mechanism elucidation, drug sensitivity prediction, combination therapy optimization, resistance phenotype identification, and clinical biomarker discovery. With the continuous advancement of AI technology and rigorous validation of clinical data, AI models are expected to fuel the development of precision oncology by improving efficacy, guiding therapeutic decisions, and optimizing patient prognosis. In summary, by leveraging clinical and omics data, AI models are expected to pioneer new therapy strategies to mitigate tumor drug resistance, improve efficacy and patient survival, and provide novel perspectives and tools for oncology treatment.

Keywords Tumor drug resistance, Artificial intelligence-driven precision therapies, Machine learning, Deep learning

*Correspondence:
Hui Zhou
xiaolingcsu@csu.edu.cn
Yi-Kun Wang
wangyikun@csu.edu.cn
Full list of author information is available at the end of the article

Graphical Abstract



Introduction

Tumor drug resistance refers to the phenomenon of tumor cells evading the effects of anticancer drugs, leading to the failure of treatments such as chemotherapy, targeted therapy, or immunotherapy. Due to the influence of tumor burden, tumor heterogeneity, tumor microenvironment (TME), and other factors [1], the majority of traditional chemotherapy and radiotherapy fail to prevent the development of resistance during treatments. More seriously, current clinical methods for assessing tumor drug resistance have a significant lag effect, leading to poor therapeutic efficacy and serious toxic side effects for patients [2]. Notably, more than 90% of cancer-related deaths have been attributed to drug resistance [3]. Scientists and clinicians have long attempted to address this challenge from multiple dimensions and have developed a variety of methods to predict tumor drug resistance, including in vitro models [4], in vivo preclinical models [5], DNA sequencing technologies [6], immunohistochemistry [7], and liquid biopsies [8]. However, each of these methods has obvious limitations, including high workload, limited predictive accuracy, and difficulty in effectively utilizing the data. In particular, the massive amount of data generated by clinical omics, pathology, and imaging poses a great challenge for direct processing and analysis, thus hindering their effective application in tumor drug resistance practice.

For the large-scale and high-precision multimodal medical oncology data generated by the rapid development of high-throughput sequencing [9], mass spectrometry [10], radiology [11], and testing technologies [12], artificial intelligence (AI) technology has already shown great potential in integrating, analyzing and interpreting multisource tumor drug resistance data [13]. Indeed, by integrating multisource heterogeneous data, including omics data, medical images, and electronic medical records, AI can identify the key resistance features and construct more accurate and comprehensive diagnostic

and prognostic models of tumor resistance [14] to facilitate cross-modal information fusion, ultimately guiding clinical precision oncology and personalized therapy. Machine learning (ML), a prominent subset of AI, relies on algorithms that learn from available data to construct models to perform specific tasks [15]. Furthermore, deep learning is a particularly adept form of ML at handling and processing massive data from genomics, transcriptomics, metabolomics, proteomics, and radiomics [16]. For instance, Rathore et al. [17] applied transfer learning using a convolutional neural network pre-trained on 1.2 million ImageNet images to extract resistance features from brain scans of 270 glioblastoma patients. This approach effectively mined resistance-related information linked to O6-methylguanine-DNA methyltransferase promoter methylation status (MGMTpms), achieving robust MGMTpms prediction with cross-validated accuracies of 86.95%, 81.56%, and 82.43% across three independent cohorts.

Artificial intelligence has the potential to significantly advance tumor resistance practice, offering promising avenues for resistance prediction and the development of precision oncology. Given the significance of AI in tumor drug resistance, this review highlights the applications of AI in basic study and clinical practice, mainly including guiding the development of drugs against tumor drug resistance, advancing drug resistance mechanisms discovery, driving drug sensitivity prediction, optimizing combination therapy, facilitating tumor resistant phenotype prediction, and accelerating biomarker discovery. Additionally, we provided a practical workflow of AI-guided tumor resistance practice, applications and discussed the perspectives and challenges associated with its use in tumor drug resistance practice. This review provides novel insights into tumor resistance practice and precision therapy, presents a useful reference for the practice of combating drug resistance in clinical tumors.

Proposed feasible workflow for artificial intelligence-driven tumor resistance practice

A streamlined and practical workflow is crucial to enhance the efficiency and accuracy of tumor drug resistance evaluation and prediction. By summarizing the extensive literature, we propose a feasible and practical workflow (Fig. 1):

Tumor drug resistance data collection

Tumor drug resistance-related data collection represents the initial step in the AI-driven workflow, and the acquisition of high-quality data is essential for advancing drug resistance practice. Available clinical data include patient demographic and clinical information [18–20], genomic data [21, 22], transcriptomic data [23–25], metabolomic data [2], proteomic data [26, 27], imaging data [28–31], and physiological or biochemical pathology test results [26, 28, 29].

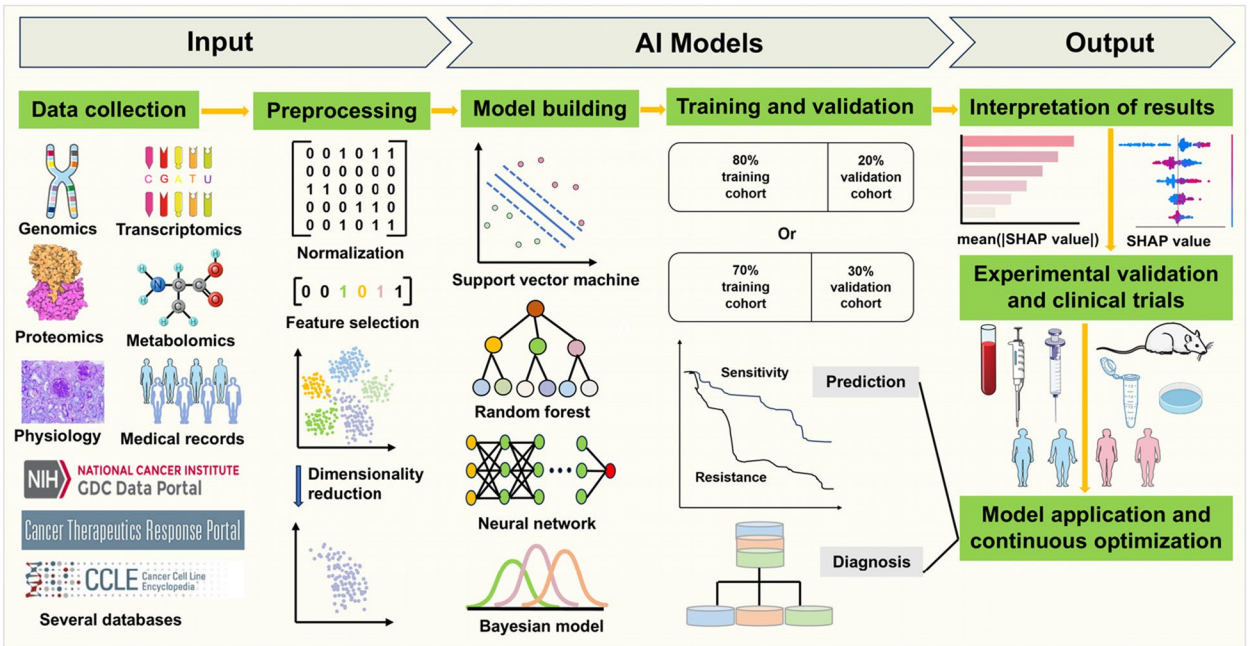


Fig. 1 A feasible workflow of artificial intelligence driven-tumor drug resistance model in basic studies and clinical practice

Preprocessing of tumor drug resistance data

Recently, massive tumor data in modern medicine have been rapidly increased and accumulated [32]. Multimodal information, including electronic health records, imaging reports, and genomic data, comprehensively covers the diagnosis and treatment process of cancer patients [33]. However, these data are scattered across different origins and systems, often containing missing values and outliers [34], and remain heterogeneous, posing significant challenges to data integration, analysis and utilize [35]. Therefore, tumor drug resistance-related data must undergo comprehensive preprocessing before being used to train AI models. This process includes several critical steps, such as coding of medical concepts, data cleaning, data standardization and normalization, and feature selection [36].

Tumor drug resistance modeling

Appropriate AI algorithms should be employed to develop diagnostic or predictive models tailored to the specific needs of tumor drug resistance trials. These models should be adept at discerning and interpreting the underlying correlations and patterns within the resistance data. Commonly used drug resistance data mining methods mainly include support vector machines (SVM), random forest (RF), logistic regression (LR), and deep learning [37]. deep learning model HECTOR was established for predicting distant recurrence risk in endometrial cancer, which extracted oncological pathology features from H&E-stained whole-slide images using a Vision Transformer, then integrated these features with image-based molecular classification and anatomical staging through a gating-based attention mechanism to generate prognostic predictions for tumors [29].

Tumor drug resistance model training and validation

Model training and validation are essential steps in applying AI to tumor drug resistance practice, ensuring that the models achieve optimal performance on the training datasets and exhibit robust generalization to unseen and unknown data. Typically, the tumor drug resistance datasets have been partitioned into a training set (often 80% or 70% of the total data) and a validation set (commonly 20% or 30%) [38]. The training set is utilized to train the model, while the validation set is employed to evaluate its performance. Commonly used validation methods include cross-validation, leave-one-out cross-validation, and k-fold cross-validation [39]. This approach can enhance the accuracy and generalization of a model, making it applicable to both basic and clinical studies on tumor drug resistance [40]. Ahn et al. [41] developed a pathology image-based deep learning classifier, PathoRiCH, to predict the response to platinum-based chemotherapy for high-grade serous ovarian cancer, employing pathology images from the SEV cohort for training and initial validation, and then utilizing images from the TCGA and SMC cohorts to further evaluate the generalization of the model.

Interpretation of tumor drug resistance results

The predictive output of AI models requires effective communication with healthcare professionals to ensure understanding and facilitate adoption [42]. To achieve this, it is essential to present model results in an interpretable manner that allows clinicians to understand the basis of the drug resistance prediction. Interpretable machine learning models have emerged as a key tool to address this challenge [43]. Specifically, the calculation of SHAP values can elucidate the biological characteristics or clinical factors with significant impact on tumor drug resistance. Guo et al. [44] constructed a more interpretable prediction model for distant metastasis in ovarian clear cell carcinoma using six different machine learning techniques, and the primary tumor stage (T) was identified as a critical clinical factor influencing metastasis risk through SHAP analysis, which also correlated with drug resistance development.

Validation of tumor drug resistance models in experimental and clinical studies

Once models screen for potential biomarkers or predict tumor resistance, these results must be further validated by molecular biology [45], cell biology [46], and cohort studies [23]. For instance, Cai et al. [47] utilized six machine learning algorithms and successfully identified the core gene RAC3, which is significantly associated with chemoresistance and immune infiltration characteristics of bladder cancer (BCa); Immunohistochemistry (IHC) staining, RT-qPCR, and Western blot were applied to validate the expression of RAC3 in BCa tumor tissues, which successfully provided potential markers for evaluating BCa resistance and addressed a critical gap in the risk assessment of BCa patients.

Application and continuous optimization of tumor drug resistance models

Initial resistance models are difficult to translate directly into clinical practice [48]. Continuous data collection and model optimization are critical to ensure the accuracy of drug resistance prediction.

Artificial intelligence facilitates discovery of tumor resistance mechanism

With the development and application of artificial intelligence, its enormous potential in biomedical and clinical fields is being continuously explored. Specifically, in basic

research on tumor drug resistance, AI can (1) identify new effective drugs against tumor resistance via facilitating the design and screening of novel drugs, predicting drug-target interactions [45], and identifying potential targets [49]; (2) elucidate the complex molecular mechanisms underlying tumor resistance in tumor cells through large-scale omics data analysis [50]; (3) construct drug sensitivity prediction models to assist clinicians in assessing the cytotoxicity of various drugs on tumor cells [51]; and (4) optimize drug combination strategies by analyzing the interactions between multiple antitumor agents to mitigate the resistance of monotherapy (Table 1) [52].

Artificial intelligence guides anticancer drug development to overcome tumor resistance

AI technologies hold tremendous promise for accelerating drug discovery and development. By leveraging machine learning and deep learning algorithms to analyze large biological and chemical datasets, AI can identify key biomarkers and molecular pathways associated with specific diseases or drug mechanisms of action [45], accelerating the drug discovery process [54], and facilitate rapid screening of potential drug candidates from large numbers of chemical compounds [53]. It is also capable of predicting the biological activity and safety of drug molecules [75], thus increasing the success rate of drug discovery (Fig. 2a).

(1) AI has facilitated antitumor drug discovery and design: AI can be applied to the design of potentially effective anticancer drugs by constructing chemoinformatics or pharmacoinformatics models to predict the properties (cytotoxicity, safety, metabolicity) of molecules. An advanced deep learning framework POLYGON was constructed via integration of variational autoencoder, reinforcement learning, and random forest regression models to embed chemical spaces and iteratively generate novel molecular structures. Among 32 generated lead compounds targeting MEK1 and mTOR, most significantly inhibited their activity at 1–10 μM and reduced tumor cell viability in vitro experimental validation [53]. Additionally, AI can facilitate the screening of potential antitumor drugs from large compound libraries. For instance, Wen et al. [54] developed an end-to-end deep learning framework combining a self-supervised graph neural network with a Transformer architecture. Fine-tuned on the BindingDB database, it screened 50 candidate clusters from 4,527,000 compounds, with further homogeneous time-resolved fluorescence assays identifying clusters exhibiting IC50 < 200 nM, accelerating cyclin-dependent kinase 12 inhibitor (CDK12i) discovery. Furthermore, AI models have been instrumental in predicting drug-target interactions and assessing drug cytotoxicity, selectivity, and risk profiles of drugs. The BipotentR model, a computational tool integrating linear mixed models and feed-forward neural networks, identified 38 immune-metabolic bifunctional regulators using single-cell data. Integration of experimental, bioinformatics and clinical validation demonstrated that regulator knockdown enhanced metabolic gene suppression and T-cell killing efficacy, with an Area Under the Curve (AUC) of 0.603 [45].

(2) AI has promoted the identification and screening of tumor drug resistance targets: AI can assist in transforming known or potential resistance-related genes or proteins into novel therapeutic targets to overcome resistance to existing therapies. For instance, Xiao et al. [75] trained ridge regression models using drug response data from GDSC and transcriptome data to predict drug sensitivity, validated across multiple colorectal cancer cohorts. Two BCL-XL inhibitors, navitoclax and WEHI-539, were identified and demonstrated the sensitivity towards high-chromosomal instability-colorectal cancer cells in vitro pharmacodynamic screening, thereby confirming CIN as a potential therapeutic target for colorectal cancer. Zhang et al. [59] employed a Bayesian model to integrate multi-omics data for predicting patient response to immune checkpoint inhibitors (ICI). They identified a stemness signature (Stem.Sig) negatively correlated with anti-tumor immunity and validated 20 stemness-related genes with immuno-resistance properties through CRISPR screening, suggesting their potential as immunotherapy.

Artificial intelligence advances molecular mechanisms underlying tumor drug resistance

Although the advent of targeted therapies and immunotherapy has significantly improved the survival rates of patients with advanced cancer, tumor resistance remains a major challenge in clinical cancer treatment. Emerging strategies such as genetic testing, liquid biopsy using circulating tumor DNA (ctDNA) technology [8], and single-cell sequencing [76] have elucidated the complex molecular mechanisms underlying tumor drug resistance. However, these approaches have generated massive amounts of data [56], which are challenging to accurately analyze and interpret using conventional statistical analysis methods.

AI can capture the complex nonlinear relationships inherent in tumor drug resistance data and extract characteristics of tumor resistance, including changes in the

Table 1 Application of artificial intelligence in basic research on tumor drug resistance

Sample or data sources	Algorithms	Validation	Key findings	Ref
List of P-gp modulator compounds from ChEMBL database	RF	Molecular docking	The model allows rapid screening of P-gp substrates and inhibitors from large chemical libraries	[49]
> 100,000 compounds binding data	Generative reinforcement learning + Variational autoencoder	Molecular docking + Cellular experiments	POLYGON can generate compounds that inhibit multiple targets with 81.9% accuracy	[53]
Drug-target interaction datasets from BindingDB	Self-supervised graphical neural network	Homogeneous time-resolved fluorescence CDK12 kinase assay + Cellular experiments	This study presents an efficient end-to-end deep learning approach to rapidly discover highly selective CDK12i	[54]
ctDNA	GBMs	Cross-validation	Drug development targeting ER, RTK, and cell cycle pathways may facilitate overcoming resistance to CDK4/6i	[55]
29 cancer types in TCGA database	Deep belief networks + Deep autoencoder	Internal validation + External validation	Deep learning-enabled genomic tumor stratification informs precision oncology decision-making	[50]
Differentiation markers of BRAFi resistant multiple myeloma cell line	RF	External validation on GDSC dataset and TCGA dataset	Targeting Rho regulated gene transcription pathways as therapeutic targets restored sensitivity to BRAFi-resistant tumors	[56]
GSE 91061 data set	XGBoost + CatBoost + RF + AdaBoost + LightGBM + Gradient Boosting	Internal validation + External validation + IHC staining + RT-qPCR + WB	The hub gene RAC3 was significantly associated with gemcitabine resistance	[47]
scRNA-seq profiles of neuroblastoma cell lines	Automatic consensus nonnegative matrix factorization	External validation on GOSH dataset + PDX + Genetically engineered mouse models	Chemotherapy rapidly induces neuroblastoma cells to adopt mesenchymal-like programs, potentially serving as a mechanism for tumor cells to evade chemotherapy-induced apoptosis	[57]
Bulk and single-cell datasets from various solid tumors	Elastic Net Regression + RF	EdU and phosphorylated Rb staining experiments + External data sets + siRNA knockdown assay	G0 stagnation underlay unfavorable responses to various therapies	[58]
Tumor RNA-seq data from anti-PD-1-treated mice	Linear mixed model + Feedforward neural network	CRISPR screening experiments	The key regulator ESRRA was activated in immunotherapy-resistant tumors	[45]
scRNA-Seq	Naive bayes	Independent test set of 149 patients	Genes implicated in cancer stemness, such as BECN1 and EMC3, could represent potential targets for immunotherapy	[59]
Abl kinase mutation dataset	Self-Paced Learning with Diversity and ExtraTrees (SPLDExtraTrees)	eightfold nested cross-validation on the TKI dataset	SPLDExtraTrees precisely forecasts the impact of protein mutations on the binding affinity of kinase inhibitors and discerns mutations conferring drug resistance	[60]
Replication stress induction drug response in GDSC and CTRP databases	VNN	CRISPR/Cas9 knockout screening + siRNA screening	The model discerns 41 protein assemblies that may serve as novel drug targets	[61]
692,859 cell line-drug pair data for GDSC and CTRP databases	VNN	fivefold cross-validation + PDX	The histone regulatory complex promoted S-phase entry through activation of KAT6A, TBL1XR1 and RUNX1	[62]

Table 1 (continued)

Sample or data sources	Algorithms	Validation	Key findings	Ref
CCLE database + TCGA database + Data from Zhongshan Hospital	LASSO regression + LR + Categorical Regression Tree + C4.5 decision trees	qRT-PCR + WB	TP53 mutations exhibited a robust association with cisplatin resistance, the expression levels of BATF3, IRF5, and ZBTB38 genes were correlated with cisplatin sensitivity	[46]
GDSC dataset + CellMiner dataset + GEO dataset	LASSO + Decision Trees + RF + SVM	External validation on GSE76092 dataset + RT-PCR + CCK8 experiment	This study identified several lncRNAs associated with oxaliplatin sensitivity	[63]
Mutation data from the GDSC and TCGA databases	Residual Genome Impact Transformer (ResGit) + Elastic Net	tenfold cross-validation	Compared to traditional genomic biomarker-based approaches, ResGitDR more effectively captures the influence of cellular state on drug response	[51]
RNA-seq data from TCGA and PCBC datasets	LASSO regression + RF + SVM-RFE + XGBoost	Internal validation on test set + External validation on ICGC dataset	Patients with high stem cell subtypes may respond poorly to immune checkpoint inhibitors	[64]
RNA-seq data from DepMap and TCGA datasets	TransCell (Self-Encoders + Migration Learning + Deep Feedforward Neural Networks)	External validation on proteomic data in CellMinerCDB and RNA-seq data in NCI60 cell line	TransCell improved drug susceptibility prediction performance by more than 50%	[65]
Gene expression data of ovarian cancer cell lines	TabNet	Cross-validation + ROC curve (mean AUC of 0.808)	BCL 2L1 was identified as an important gene contributing to cisplatin resistance	[66]
Pharmacogenomic data from 3D organoid models and web	Ridge Regression + LR + Support Vector Regression + Deep Neural Networks	Survival analyses + External validation	The model empowers physicians to predict a patient's response to specific drugs contingent upon their gene expression profile	[67]
Tumor transcriptomics from > 700 ICB-treated homozygous mice	Nonnegative matrix factorization	Internal validation on mouse tumor data + External validation on mouse tumor samples not involved in training	The model can be used to predict which cancer types or patients exhibit a higher propensity to respond to ICB therapy	[68]
Subcutaneous injection ovarian PDX tumor	Forward–backward feature selection classifier	fivefold cross-validation + Leave-one-out cross validation + Intraperitoneal injection PDX	The platform predicted intraperitoneal injection outcomes for three second-line cytotoxic therapies with a mean AUC of 0.91, applicable to new drug preclinical development	[69]
Breast cancer gene dataset from TCGA database	One-class logistic regression	Gene enrichment analysis + Breast cancer cell line + Mouse model	Combined treatment with paclitaxel and britannin fully eradicated breast cancer stem cells in mice	[70]
Generation of 1167 RNA-seq profiles from six DLBCL cell lines	Neural network + Bayesian network propagation analysis	6 different DLBCL cell lines + xenograft mouse models	Histone deacetylase inhibitors showed potent synergy when combined with JAK inhibitors	[71]
Temozolomide cellular pharmacology dataset + Dose resolution dataset	GBMs	LN229 glioblastoma cell line	The combination of quantitative systems pharmacology and ML enabled the design of multi-drug therapies to overcome initial tumor resistance	[52]
Integration of drug response data with genomic data	RF + Gradient Boosting + XGBoost	tenfold cross-validation + Single cell imaging cytometry	The model can select the optimal drug combination in accordance with the patient's molecular profile and drug response data	[72]

Table 1 (continued)

Sample or data sources	Algorithms	Validation	Key findings	Ref
Gene expression, mutation and copy number variation data	RF + XGBoost + CatBoost	2D cell lines + 3D tumor slice culture models + WB	The combination of lapatinib and pazopanib exhibited potent therapeutic effects on breast cancer by inhibiting the downstream PI3K/AKT/mTOR signaling pathway	[73]
Expression levels of 105 proteins from the TCGA database	RF	Proteomics of MMTV-R26 Met tumors	WEE1 and BCL-XL synergistically induced cell death in TNBC	[74]
mRNA expression data from colorectal cancer samples in the GDSC database	Ridge Regression	CIN-aneuploid cell model + Mouse xenograft model + Molecular biology experiments	BCL-XL inhibitors may yield superior therapeutic outcomes in patients with high chromosomal instability tumors	[75]

P-gp P-glycoprotein, *RF* Random Forest, *CDK12i* Cyclin-Dependent Kinase 12 Inhibitor, *ctDNA* Circulating tumor DNA, *GBMs* Gradient Boosting Machines, *IHC* Immunohistochemistry, *WB* Western Blot, *PDX* Patient-Derived Xenograft, *VNN* Visible Neural Network, *LASSO* Least Absolute Shrinkage and Selection Operator, *LR* Logistic Regression, *SVM* Support Vector Machine, *SVM-RFE* Support Vector Machine Recursive Feature Elimination, *ROC* Receiver Operating Characteristic, *AUC* Area Under the Curve, *TNBC* Triple-Negative Breast Cancer

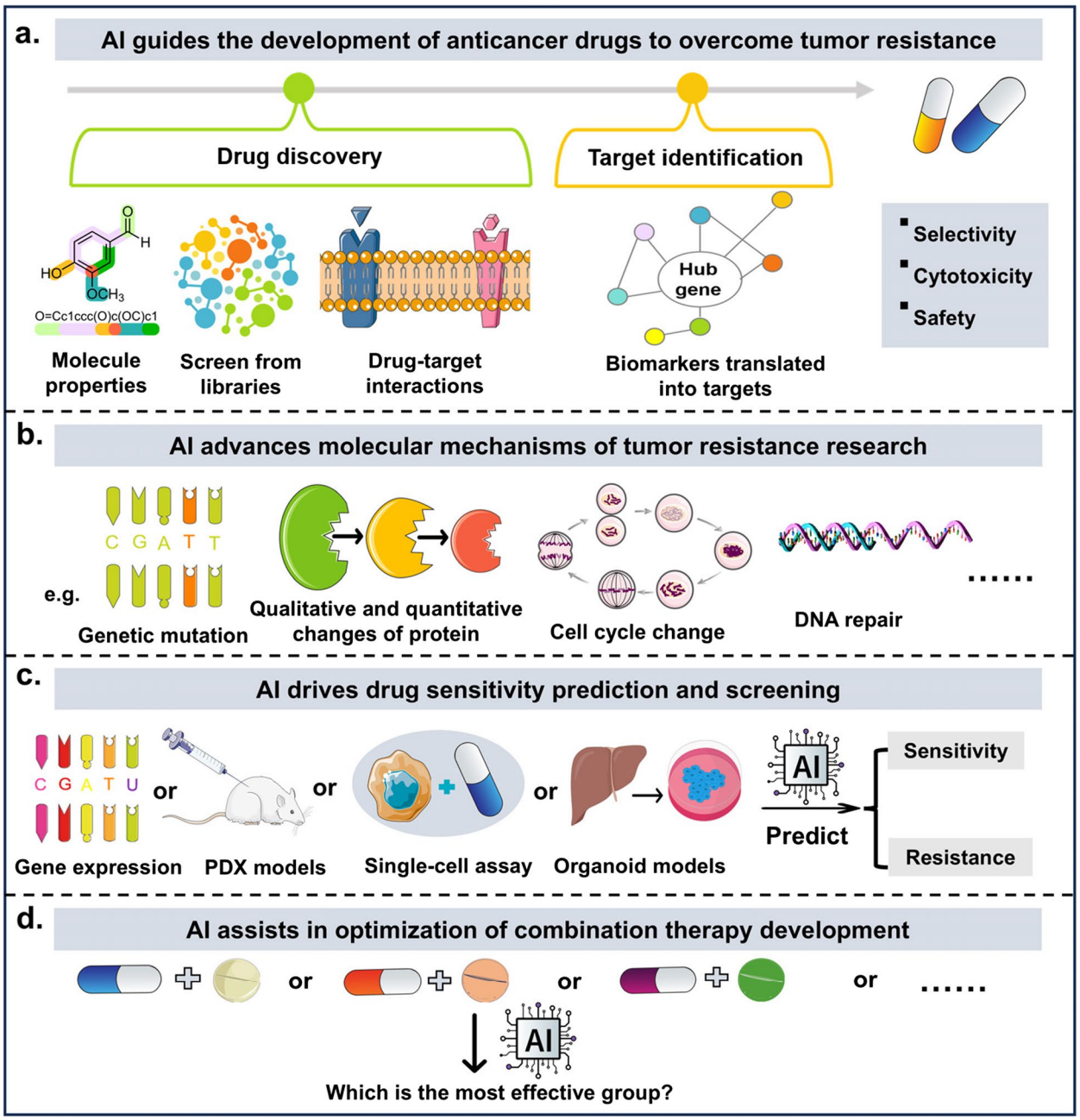


Fig. 2 Artificial intelligence in basic research on tumor drug resistance. **a** AI can facilitate the design and screening of drugs against tumor drug resistance by predicting molecular properties, screening effective lead compounds from libraries, predicting drug-target interactions, and identifying potential targets. **b** AI can help elucidate the complex molecular mechanisms underlying drug resistance in tumor cells. **c** AI can construct drug sensitivity prediction models to assess the inhibitory effects of various drugs on tumor cells. **d** AI can optimize drug combinations and explore combination strategies by analyzing the interactions between multiple antitumor agents

cell cycle, TME modes [64], modes of cell death [77], abnormal expression of tumor resistance-related proteins [60], gene regulation [66], and the mediation of various signaling pathways, providing valuable insights into the molecular mechanisms of tumor resistance from genomic and proteomic perspectives (Fig. 2b) [78]. Integrating a visible neural network (VNN) with hierarchical structures, backpropagation, AdamW optimizer, BatchNorm, and Dropout, an interpretable deep learning model NeST-VNN was established. Using the Cancer Multi-Protein Complexes Atlas (NeST) and data from GDSC and Cancer Therapeutics Response Portal

(CTRP), it analyzed 718 resistance relevant genes with mutations, copy number alterations, and deletions to uncover drug-resistance mechanisms. Based on in vitro cell line screening, patient-derived xenograft (PDX) modeling, clinical data, and CRISPR validation, it revealed that histone regulatory complexes, mediated by KAT6A, TBL1XR1, and RUNX1, promote S-phase entry, driving resistance of cyclin-dependent kinase 4/6 inhibitor (CDK4/6i) [62]. Gerratana et al. [55] employed gradient boosting machines to analyze baseline ctDNA data from 610 hormone receptor-positive/HER2-negative metastatic breast cancer patients, identifying resistance mechanisms of CDK4/6i, including alterations in ER, RTK, and cell cycle pathways.

Artificial intelligence drives drug sensitivity prediction and screening

AI can assist in predicting individual drug responses by rapidly identifying and obtaining specific sets of genes or genetic traits, protein profiles, and metabolic characteristics that correlate with treatment outcomes. Numerous studies have established accurate, efficient, and intuitive drug sensitivity prediction platforms by combining tumor drug sensitivity assessment models [65] with AI algorithms for evaluating the cytotoxicity of various drugs [67], predicting the individual treatment responses, and providing a scientific basis for clinicians to design personalized and precise treatment programs (Fig. 2c).

(1) Gene expression profiling: Gene expression profiling has become a widely used method for tumor drug sensitivity screening [79]. However, variations in sequencing depth across different technologies and laboratories, as well as batch effects and heterogeneity, have posed challenges for single-cell RNA sequencing (scRNA-seq) data analysis [80]. The integration of AI technology and gene expression profiling not only enhances data processing and model prediction accuracy but also facilitates a comprehensive understanding of the mechanisms underlying gene expression. TransCell, a two-step deep transfer learning framework integrating autoencoders, transfer learning, and deep feed-forward neural networks, was trained on pan-cancer tumor samples and validated using CellMinerCDB and DepMap 20Q1 data. It predicted drug sensitivities for 124 pediatric cell lines across 4686 drugs, identifying 29 broadly effective and cancer-specific agents [65].
(2) Patient-derived xenograft (PDX) model: The PDX model is widely considered to more accurately reflect tumor heterogeneity, with efficacy evaluation results closely resembling those observed in clinical patients [81]. However, the use of the PDX models has been constrained by several factors: tumor implantation and expansion, as well as extensive in vivo drug sensitivity testing, which typically requires 10–15 months [82], and has hindered the application of these models in tumor resistance studies. AI can be subjected to assist in identifying factors that influence the success rate of PDX model establishment, such as tumor tissue processing methods and transplantation site selection, to optimize the establishment process [83]. Furthermore, AI can analyze genomic and transcriptomic data from PDX models to reduce time consumption. For instance, Cotler et al. [69] developed a platform to accelerate ovarian cancer drug sensitivity testing, predicted outcomes of intraperitoneal injections for three second-line cytotoxic therapies, achieving an average AUC of 0.91, based on machine learning classifier with linear regression and forward–backward stepwise feature selection,
(3) Single-cell drug susceptibility testing: Single-cell drug susceptibility testing, a method developed in recent years, can evaluate the sensitivity of single cells to various antitumor drugs [84]. Cellular parameters obtained from single-cell cytotoxicity assays can be processed as input files and further analyzed by AI algorithms, providing deeper insights for clinical decision-making [85]. Additionally, integrating tumor sensitivity testing with AI to analyze various omics data can facilitate more personalized drug selection for cancer patients in the era of precision oncology. Based on RF, SVM, and k-nearest neighbors algorithms, the first single-cell transcriptome-based AI model, SCATTome, was developed predict individual cell responses to proteasome inhibitors, validated by in vitro pharmacodynamic screening, addressing the challenge of drug sensitivity heterogeneity in multiple myeloma [86].
(4) Organoid models: The use of organoids in tumor drug resistance studies has gained significant attention recently [87]. The integration of AI-driven data analysis can optimize quality control processes and culture conditions, alleviating the financial and technical hurdles inherently associated with organoid culture [88], while simultaneously enhancing the efficiency and accuracy of drug sensitivity predictions. Kong et al. [67] integrated ridge regression with linear regression, support vector regression, and deep neural networks to train a model using transcriptomic and pharmacovigilance data from colorectal and bladder cancer organoids. Based on TCGA patient transcriptomic data, the model pre-

dicted drug responses for 114 colorectal and 77 bladder cancer patients, with survival analysis significantly supporting the predictions ($P=0.014$ and $P=0.01$, respectively).

Artificial intelligence assists in optimizing combination therapy development

AI can integrate data from multiple biomedical sources [70], organize drug combination datasets [72], predict drug combination sensitivity, and save valuable time in drug combination screening [89], thereby overcoming tumor resistance, addressing combinatorial explosion challenges, and enhancing cost-effectiveness in oncology drug screening (Fig. 2d). scTherapy, a machine learning model based on LightGBM, integrated single-cell transcriptome and drug response data to predict combination therapies for metastatic/refractory tumors. Following validation of flow cytometry and bulk cell viability assays in acute myeloid leukemia patient samples, 96% of predicted combinations demonstrated selective efficacy or synergy [90]. ComboPred combining RF, gradient boosting, and XGBoost, identified synergistic drug combinations with high selectivity against ovarian cancer. It prioritized candidates like the mTOR inhibitor vistusertib and BCL2L1 inhibitor A1155463 (HSA score 9.7) for single-cell validation, guiding preclinical oncology drug testing [72].

Meanwhile, AI models can analyze the synergistic mechanism of antitumor agents, thereby optimizing drug combinations, facilitating the identification of the optimal therapeutic regimen, and maximizing the benefits for oncology patients. For instance, Zhou et al. [73] developed prediction models using RF, XGBoost, and CatBoost to analyze drug combination datasets from DrugComb, DrugCombDB, and SYNERGxDB, and identified the combination of lapatinib (RTK inhibitor) and pazopanib (multi-kinase inhibitor) as potent against breast cancer by blocking the PI3K/AKT/mTOR pathway, validated by in vitro cytotoxicity screening towards MDA-MB-231 and HCC1937 cell. Davis et al. [71] trained an AI model combining neural networks and Bayesian network propagation on IC50 values and RNA sequencing (RNA-seq) data from six DLBCL cell lines, predicting strong synergy between histone deacetylase inhibitors and JAK inhibitors. Experimental validation confirmed significant synergistic effects of these combination therapy ($p<0.01$), establishing a computational-experimental closed-loop framework for cancer combination therapy development.

Artificial intelligence assists reduce tumor drug resistance in clinical oncology

Tumor drug resistance poses a significant challenge in clinical oncology, making studies into this phenomenon essential for enhancing the effectiveness of cancer therapies [91], improving patient outcomes [92], reducing healthcare costs, and advancing medical science. Recently, AI has been increasingly applied to clinical trials of tumor drug resistance, offering promising solutions to address resistance issues. By integrating diverse sets, including genomic, transcriptomic, proteomic, imaging, and clinical data, machine learning models have been developed to predict patient responses to specific oncology drugs (Table 2) [93].

Artificial intelligence facilitates the prediction of tumor-resistant phenotypes

Undeniably, AI is a powerful tool for processing and analyzing patient genomic data [22] and clinical information [113]. Its introduction could assist physicians in identifying patients at higher risk of developing tumor resistance. Tumor drug resistance is influenced by multiple factors, such as gene mutations [101] and alterations in gene expression [75], which contribute to the tumor's increased tolerance to therapeutic agents. AI can predict drug resistance phenotypes by systematically analyzing large-scale clinical information and genomic data (Fig. 3a). Furthermore, factors such as tumor patients' genetic background and mutations, weight, gender, age, and lifestyle habits may influence treatment outcomes [114], and AI contributed to clinicians systematically evaluating patients' drug resistance and predicting resistance-prone populations or cohorts [115].

AI can fully mine tumor resistance-related data from vast amounts of data in single-cell omics to guide tumor drug resistance. A drug response prediction model PERCEPTION using transfer learning and elastic net regularization, predicted responses of transcriptional clones within tumors, with the most resistant clone indicating overall patient response. It stratified responders and non-responders in multiple myeloma ($AUC=0.83$) and breast cancer ($AUC=0.776$) experiments [116]. Liu et al. [2] mined and analyzed HCT-116 colon cancer single-cell mass spectrometry metabolomic data, then used RF, artificial neural networks, and penalized LR to predict drug resistance in individual cells, achieving 86.5% accuracy in validation, highlighting its clinical potential. Goldstein et al. [94] developed an AI model with SVM, RF, and XGBoost, classify drug resistance and metastatic potential with > 95% accuracy using lung cancer cell features (size, granularity, fluorescence intensity).

Table 2 Application of artificial intelligence in clinical research on tumor drug resistance

Sample or data sources	Algorithms	Validation	Clinical Insights	Ref
Cancer cell size, granularity, fluorescence intensity of five particles	XGBoost + RF + SVM	fivefold cross-validation + tenfold cross-validation + ROC curve	Cancer cell subtypes were accurately classified with over 95% accuracy	[94]
RNA expression of 1003 early gastric cancers in metastatic and non-metastatic lymph nodes	RF	Leave-one-out cross-validation + External validation + Mouse models	The model can identify early-stage gastric cancer patients at high metastatic risk, reducing recurrence and metastasis risks	[23]
Micro-organospheres from patient tumor tissue	Region-based convolutional neural network	Comparison with conventional cell viability assays	The model enables rapid evaluation of a patient's response to chemotherapy, targeted therapy, and radiotherapy	[93]
Preclinical cancer cell lines mRNA expression level	TINDL (Fully Connected Neural Network + CXPlain)	Mann–Whitney U + siRNA knockout assay	TINDL was a powerful deep learning framework for predicting drug response and identifying biomarkers of cancer drug response	[95]
Clinicopathologic data on patients in 3 tertiary care hospitals	LR + RF + SVM + Deep Neural Network	fivefold cross-validation + ROC curve	The LR-based model was the most effective in identifying platinum-resistant cases, with an AUC of 0.741	[28]
PIs-sensitized and PIs-treated multiple myeloma patient cells	SCATTome (RF + SVM + LASSO + kNN)	Correlation analysis with cell line drug sensitivity data + ROC curves + Patient survival analysis	SCATTome facilitates the identification of high-risk patients with poor response to specific treatments in clinical data analysis	[96]
Pathologic images of 814 high-grade serous ovarian cancer patients	PathoRiCH (MIL + Contrastive Self-Supervised Learning + DS-MIL)	fivefold cross-validation + External validation + Survival analysis + Cox regression	PathoRiCH analyzes histopathology images to facilitate early identification of patients with favorable or unfavorable response to platinum-based chemotherapy	[41]
scRNA-seq derived from acute myeloid leukemia patients	XGBoost	In vitro assay + Flow cytometry	Models can predict synergistic drug combinations and selectively inhibit leukemia cells in a given patient	[91]
Multimodal data from 147 breast cancer patients	MOMLIN (SMCCA + WMSCCA + LR)	fivefold cross-validation (100 repetitions) + AUC	MOMLIN revealed detailed molecular characterization of drug responses in breast cancer, with a mean AUC of 0.989	[97]
Preoperative biopsy specimens from 153 gastric cancer patients	SVM	Validation cohort (46 patients, AUC = 0.80)	This model successfully predicted treatment response to sintilimab combined with the SOX chemotherapy regimen in gastric cancer patients	[98]
Transcriptome data from 46 abiraterone-responsive metastatic castration-resistant prostate cancer patients	Sparse Bayesian model	Independent dataset validation + siRNA knockdown experiments	Disruption of ELK3, MXD1, and MYB signaling cascades affected abiraterone resistance	[99]
Multi-transcriptome sequencing of chronic myeloid leukemia patients	SVM-RFE + LASSO + RF	Multiple independent datasets (including GSE144119 and clinical samples) + ROC curve + In vitro assay	Targeting ferritin deficiency was a potential therapy to overcome resistance in patients with chronic myeloid leukemia	[100]
Clinicopathologic characteristics and drug response data in 1339 patients with EGFR mutations	D3EGFR model	Internal validation on D3EGFRdb database + External validation on data of 102 patients	D3EGFR accurately predicted drug response in patients with EGFR mutations, achieving validation accuracies of 0.81 internally and 0.85 externally	[101]

Table 2 (continued)

Sample or data sources	Algorithms	Validation	Clinical Insights	Ref
H&E-stained pathology slides	PhenoTIL (U-Net Convolutional Neural Network + SVM + Gaussian Mixture Model)	Kaplan–Meier + Cox regression + Quantitative immunofluorescence images + RNA-seq	For patients with low PD-L1 expression, PhenoTIL can identify those who may benefit from immune checkpoint inhibitor monotherapy	[31]
RNA-seq data and clinical data from liver hepatocellular carcinoma patients	Survival-SVM, GBM, RSF, SuperPC, PLS-Cox, CoxBoost, StepCox, Ridge, LASSO, Elastic Net	Internal validation on training cohorts + External validation on the ICGC-LIRI and GSE14520 cohorts + RT-qPCR + WB + IHC stain	PAK1IP1 was the most important gene for predicting prognosis in liver hepatocellular carcinoma	[26]
Transcriptome sequencing data from treatment-resistant TNBC	Glmnet + kNN + Neural network + RF + SVM + Kernel SVM	Cross-validated on BEAUTY and I-SPY1 datasets (AUC = 0.88) + Survival analysis on the independent TNBC dataset	The 17 identified genes, particularly those linked to inflammatory signaling pathways like NF-κB, may serve as potential therapeutic targets	[24]
Expression data of 46 RNA methylation regulators from the TCGA database	RF	tenfold cross-validation (AUC > 0.9) + scRNA-seq + in vivo + in vitro	Tumors with high expression of RNA methylation regulator showed increased genomic alterations and resistance to neoadjuvant chemotherapy	[102]
Gene expression data from 84 primary glioblastoma	Ridge Regression	Validated on an independent test set containing 23 primary glioblastoma samples (73. 9% accuracy)	The model aids in early identification of glioblastoma patients who may exhibit treatment resistance	[22]
Large tumor transcriptomes from over 700 patients	Network Propagation	Leave-one-out cross-validation (Median AUC = 0.79) + scRNA-seq + Survival analysis	Growth factor-mediated structural cell communication marked resistance to ICI	[25]
486 HNSCC transcriptomes and clinical data	SVM-RFE	Cellular assay + Xenograft mouse model + Immunohistochemistry + qRT-PCR	PYGL was a key metabolism-related biomarker	[103]
Metabolite profiles of pre-treatment serum in 80 female BC patients	RFE + LR	Leave-one-out cross-validation + ROC curve (AUC > 0.8)	Glyoxalate and dicarboxylic acid metabolism were associated with resistance to neoadjuvant chemotherapy	[104]
Pan-cancer scRNA-Seq dataset (65 datasets; 26 cancer types)	HistGradientBoost	Validation set (AUC = 0.74) + Confusion Matrix + Survival analysis	FOXO1 serves as a crucial regulator of immunotherapy resistance, and its inhibition enhances immunotherapy efficacy	[105]
Multi-omics data and scRNA-seq data on tumor samples	RF	Two independent validation cohorts (AUC = 0.711)	PDK1 was a hub gene associated with tumor hypoxia, glycolysis, and immunotherapy resistance	[106]
Ovarian tissue and plasma samples from 813 patients	XGBoost	External validation + Survival analysis + Log-rank test (p = 0.047)	Fusion of proteomics data and machine learning algorithms enables precise prediction of cancer recurrence risk	[107]
Liquid biopsy of CTCs	RF	3 patient samples (model recall 0.76, specificity 0.99, precision 0.17)	Combinatorial dual-color could identify CTCs, detect CTC clusters, and assess CTC heterogeneity	[108]
TCGA and CGGA databases + Clinical samples from Peking Union Medical College Hospital	LASSO + SVM + RF + XGBoost	ROC curve + External validation + Survival analysis	Immunotherapy achieved improved outcomes in patients with higher mRNAsi scores	[109]
Sequencing and clinical data from 544 adult patients with diffuse gliomas	CELLO2 model	Patient-derived glioma spheroids and established glioma cell lines	MYC was an early predictor of cancer evolution	[19]

Table 2 (continued)

Sample or data sources	Algorithms	Validation	Clinical Insights	Ref
Genomic information in exosomes from TNBC patients	CoxBoost	Survival analysis + ROC curve + External validation on GEO dataset	Exosome-related genes could aid risk stratification of TNBC and identify patients with poor prognosis	[92]
Plasma samples from 120 NSCLC patients treated with ICI	SVM + RF + LR + Neural Networks	Internal validation	By identifying predictive biomarkers, this model assists clinicians in personalizing ICI treatment for NSCLC patients	[110]
Clinical and RNA-seq data from patients ($n = 522$) with 4 cancers	AE-SDN (Auto-Encoder + Cox regression)	Survival analysis + Log-rank test	AE-SDN could identify prognostic biomarkers in ICI-treated cancer patients, outperforming CD3 + and CD8 + T cell immunity scores	[111]
Approximately 800 plasma proteins in 143 ICI-treated NSCLC patients + Different clinical parameters	XGBoost	ROC curve (AUC = 0.79) + Cross-validation	CXCL8, CXCL10, age, and sex could predict response to ICI therapy in NSCLC patients	[112]

RF Random Forest, *RSF* Random Survival Forest, *IHC* Immunohistochemistry, *WB* Western Blot, *LASSO* Least Absolute Shrinkage and Selection Operator, *LR* Logistic Regression, *SVM* Support Vector Machine, *SVM-RFE* Support Vector Machine Recursive Feature Elimination, *RFE* Recursive Feature Elimination, *kNN* k-Nearest Neighbor, *ROC* Receiver Operating Characteristic, *AUC* Area Under the Curve, *CTCs* circulating tumor cells, *PIs* Proteasome Inhibitors, *NSCLC* Non-Small Cell Lung Cancer, *ICI* Immune Checkpoint Inhibitors, *TNBC* Triple-Negative Breast Cancer, *EGFR* Epidermal Growth Factor Receptor, *MIL* Multiple Instance Learning, *DS-MIL* Deep Supervision-based Multiple Instance Learning, *SMCCA* Sparse Multiple Canonical Correlation Analysis, *WMSCCA* Weighted Multiblock Sparse Canonical Correlation Analysis

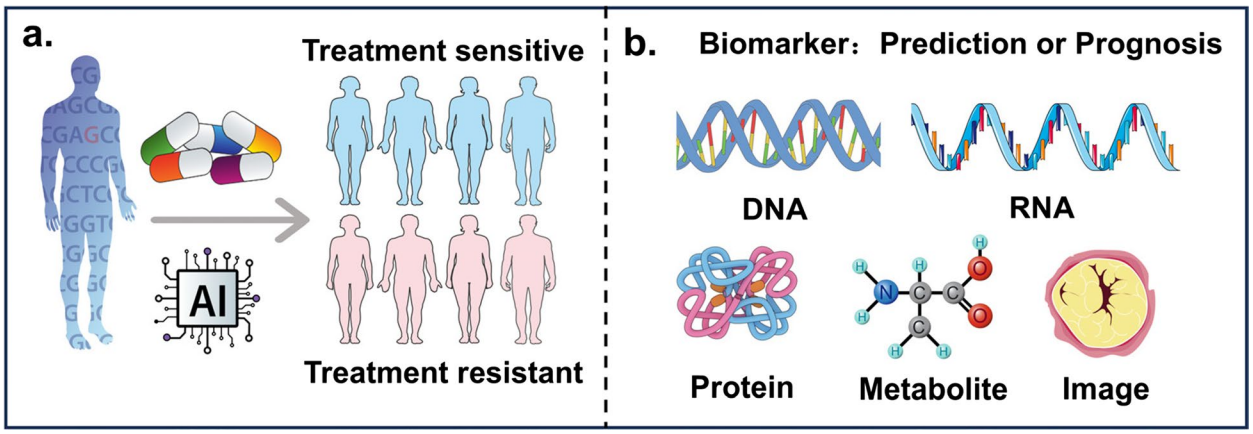


Fig. 3 Artificial intelligence-guided tumor drug clinical practice. **a** AI models predict patient responses to specific drugs and identify and differentiate individuals with distinct drug-resistant phenotypes. **b** AI models screen predictive and prognostic biomarkers associated with drug resistance

Additionally, liquid biopsy samples can provide genomic, transcriptomic, epigenomic, proteomic, and metabolomic information about tumor resistance through omics analysis [8]. AI can efficiently integrate and process these multidimensional datasets, enhancing assay precision and supporting more informed diagnostic and therapeutic decisions. Based on plasma proteomics data of 184 non-small cell lung cancer patients, Shaked et al. used supervised learning, and identified resistance-associated proteins and combined key clinical parameters to predict response of ICB therapy. Survival analysis revealed risk ratios of 4.5 (CI 2.07–9.77; $p < 0.0001$) for overall survival and 2.27 (CI 1.7–4.03; $p = 0.004$) for progression-free survival, enabling patient stratification and dynamic monitoring [117].

Artificial intelligence accelerates discovery of tumor resistance biomarkers

In addition to tumor drug resistance phenotypes, there is an urgent clinical need for specific and highly sensitive biomarkers to monitor tumor progression and resistance [118]. By leveraging AI-based bioinformatics tools and computational biology models, the gene expression levels of proto-oncogenes and oncogenes [99], enzyme activity, and metabolic reprogramming [103] in patients' tumors can be thoroughly analyzed to screen for molecular markers that are closely related to tumor drug resistance, providing reliable evidence for the progression, prediction [109], and prognosis [111], while also providing scientific rationale for the development of novel drug resistance detection technologies and therapeutic strategies (Fig. 3b).

The discovery of predictive biomarkers is critical for stratifying patients into distinct susceptibility subtypes and enabling personalized treatment [119]. AI can assist physicians in identifying resistance genes [106] or proteins [107] through minimally invasive manipulation, thereby enhancing patient risk stratification and clinical trial selection. Based on tissue information normalization and deep learning with a fully connected neural network, TINDL was established and trained on RMA-normalized data from 958 GDSC cancer cell lines, and tested on TCGA primary tumor RNA-seq data. It identified key resistant genes (such as SLFN11, RPS6, RPL13) and pathways, validated by siRNA knockdown experiments [95]. Lee et al. [25] used LR and network propagation to analyze TME interactions, predicting ICI responses in melanoma, lung, bladder, and gastric cancers, and achieved a median AUC of 0.79 across 11 ICI cohorts, with single-cell experiments and enrichment analyses identifying resistance-associated pathways as potential combination therapy targets.

Prognostic biomarkers are valuable tools for monitoring cancer progression and treatment efficacy, allowing physicians to make timely adjustments to treatment regimens [120]. By integrating single-cell and multi-patient sample sequencing [100] and leveraging AI to explore the relationship between specific tumor cell subpopulations and patient prognosis, disease progression and outcome can be predicted, providing novel insights for clinical diagnosis and treatment of cancer [105]. An AI model named AE-SDN, combining autoencoders and deep neural networks, extracted key features from tumor RNA-seq data into a Cox regression layer to output patient risk scores and identify immune-, oncogenic-, and tumor suppressor-related genes. Compared to $CD3^+/CD8^+$ T-cell-density-based immune scores, AE-SDN improved predictive power by > 20% [111]. Guan et al. [103] used

support vector machine-recursive feature elimination to screen PYGL, a prognostic metabolic gene, from 858 KEGG pathway genes and single-cell data. A xenograft model confirmed that PYGL knockdown inhibited tumor growth, supporting PYGL as a potential metabolic therapy target.

Available online tumor drug resistance databases or servers

Available, easy-to-use tumor resistance databases shorten the gap between basic research and clinical application, allowing physicians to quickly obtain therapeutic references and guidance based on laboratory data, contributing to precision oncology decision-making (Table 3) [121]. In addition to The Cancer Genome Atlas (TCGA) (https://www.cancer.gov/tcga) [122], based on the aggregation of drug sensitivity data from nearly 75,000 experiments, Yang et al. [123] developed the GDSC database (https://www.cancerrxgene.org/), which can identify molecular biomarkers of drug sensitivity by querying for specific anticancer drugs or cancer genes, facilitating the discovery of novel biomarkers for cancer therapy. The DRMref database (https://ccsm.uth.edu/DRMref/) [124], which analyzed tumor cell composition, intra-tumor heterogeneity, and epithelial-mesenchymal transition scores, provides a comprehensive characterization of drug resistance mechanisms and supports the development of drug combinations and innovative therapeutic targets. The Cancer Therapeutics Response Portal (CTRP) (http://portals.broadinstitute.org/ctrp/) [125], links genetic and cellular characteristics of 860 cancer cell lines to their sensitivity to 481 small molecule probes and drugs, accelerating the discovery of patient-matched therapies. ncRNADrug (http://www.jianglab.cn/ncRNADrug) [126] catalogs non-coding RNAs (ncRNAs) associated with drug resistance and targets, and predicts drug-ncRNA

Table 3 Available databases on tumor drug resistance

Databases	Websites	Functions	Ref
DRMref	https://ccsm.uth.edu/DRMref/	DRMref provides comprehensive characterization of drug resistance mechanisms using single-cell data obtained from drug treatment settings	[124]
DRESIS	https://idrblab.org/dresis/	DRESIS is a comprehensive list that characterizes drug-resistant diseases and all types of resistance mechanisms	[128]
GDSC	https://www.cancerrxgene.org/	GDSC integrates large-scale drug susceptibility and genomic datasets to aid in the identification of novel biomarkers for cancer therapy	[123]
CTRP	http://portals.broadinstitute.org/ctrp/	CTRP links genetic, genealogical and cellular characteristics of cancer cell lines to drug sensitivity to accelerate the discovery of patient-matched cancer therapies	[125]
CTR-DB 2.0	http://ctrdb.ncpsb.org.cn	With gene set enrichment and tumor microenvironment analysis, CTR-DB 2.0 assists in elucidating tumor resistance mechanisms and identifying potential combination therapies and their predictive biomarkers	[129]
CancerDR	http://crdd.osdd.net/raghava/cancerdr	CancerDR offers pharmacological data for 148 anticancer drugs across 952 cancer cell lines, aiding in the identification of genetic alterations in drug target-encoding genes	[130]
ncRNADrug	http://www.jianglab.cn/ncRNADrug	ncRNADrug enables the prediction of drug-ncRNA associations based on ncRNA expression profiles, aiding drug development	[126]
MdrDB	https://quantum.tencent.com/mdrdb/	MdrDB captures the biochemical impact of mutations on protein–ligand affinity, providing insights into mutation-driven drug resistance, combination therapy development, and novel drug discovery	[131]
CCLE	https://portals.broadinstitute.org/ccle	CCLE integrates pharmacological data for 24 anticancer drugs across 479 cell lines, enabling the identification of predictors for drug sensitivity	[132]
TCGA	https://www.cancer.gov/tcga	TCGA provides comprehensive genomic, epigenomic, transcriptomic, proteomic, and clinical data to support cancer diagnosis, treatment, and prevention	[122]
PharmGKB	https://www.pharmgkb.org	PharmGKB is a pharmacogenomics database containing genetic information on tumor drug resistance, aiding the identification of factors influencing individualized therapy	[133]
COSMIC	https://cancer.sanger.ac.uk/cosmic	COSMIC provides data on somatic mutations in cancer, enabling the assessment of their impact on disease progression	[134]
OncoKB	https://www.oncokb.org	OncoKB is a cancer genome database that offers information on cancer-related mutations, drug resistance, and treatment response	[135]
ScDrugAct	http://bio-bigdata.hrbmu.edu.cn/scDrugAct	ScDrugAct dissects cellular heterogeneity and the tumor microenvironment, shedding light on mechanisms of drug action and resistance	[127]
CancerTracer	http://cailab.labshare.cn/cancertracer	CancerTracer facilitates tracking and characterization of tumor evolution in individual patients, aiding the identification of predictive biomarkers for personalized cancer therapy	[136]

interactions to support drug development and cancer treatment. ScDrugAct (http://bio-bigdata.hrbmu.edu.cn/scDrugAct) [127] compiles 17,274 drug-related genes and 276,559 associations between over 10,000 drugs and 53 cell types, linking drugs, genes, and cells to support cell type-specific therapies and the identification of therapeutic biomarkers.

Current challenges and future perspectives

AI models outperform traditional methods in data integration, handling complex data, and adaptability, offering deeper insights into biomedical data for clinical decision-making and drug development [137, 138]. However, their effectiveness depends on high-quality data [139], and their "black-box" nature poses interpretability challenges, particularly with complex omics data and resistance mechanisms [61]. Therefore, standardized data management and acquirement is essential to ensure data quality and consistency [140], while efforts should prioritize enhancing model interpretability and visualization [141]. Moving forward, multimodal AI models should be utilized to integrate diverse data sources, emphasizing key interactions between oncological data modalities to boost predictive accuracy for resistance. Moreover, strengthening collaboration among computer scientists [142], biologists [143], clinicians, and pharmacologists [144] will be vital for translating research into practical applications and advancing precision oncology.

Quality and standardization of tumor drug resistance data

The scarcity of high-quality and suitable datasets is a major challenge for the application of AI algorithms to tumor drug resistance [145]. Currently, AI models based on in vitro cancer cell lines show limited translational potential in forecasting clinical drug responses in real-world scenarios [146]. Although many studies have moved towards AI models using clinical data, appropriate resistance datasets remain limited [139]. Even in large databases such as TCGA, clinical drug response data are typically sparse [95].

The integrity and comprehensiveness of clinical data are fundamental to clinical research and evidence-based decision-making, necessitating rigorous quality control and validation protocols [147]. Effective integration and processing of heterogeneous clinical data are critical for ensuring the reliability of AI models [148]. In particular, the National Cancer Institute Genomic Data Commons (GDC) dataset, which integrated data from multiple cancer genome programs, provided comprehensive clinical drug response data alongside multi-omics profiles. By processing patient-derived molecular data through standardized GDC workflows, researchers can easily achieve data normalization, thereby enhancing the quality of tumor resistance modeling [149].

Interpretable and transparent AI models urgently needed in clinical oncology

Tumor drug resistance prediction models based on AI algorithms are often considered "black box" models due to the difficulty of explaining how these models actually arrive at their decisions [43]. When faced with the challenge of balancing performance and interpretability, scientists often prioritize performance metrics such as accuracy, precision, and recall. However, healthcare decisions require weighing complex, sometimes conflicting data, and clinicians value the interpretability and practical applicability of tumor resistance models [150].

Enhancing the interpretability of models can assist physicians in gaining a deeper understanding of the molecular basis of drug resistance toward tumor and developing more effective oncotherapy accordingly [43], potentially improving the feasibility and practicability of drug resistance models in clinical [62]. Zhao et al. [61] developed a series of "visible" neural network (VNN) models that linked genetic alterations to drug responses, utilizing knowledge maps of biological components and functions to guide the internal architecture of the model. Unlike traditional "black box" neural networks, VNN predictions of biomedical outcomes could be mapped to changes in molecular mechanisms and pathways, thereby enhancing the interpretability of clinical decisions. Ogunleye et al. [139] constructed a patient-interpretable machine learning model, where expression levels of selected miRNAs were nonlinearly combined by the CART algorithm in a correlated manner, supporting the model's predictive outcomes.

Emerging multimodal artificial intelligence models for enhanced robustness and accuracy in tumor drug resistance

Clinical data sources are massive and diverse, encompassing a wide range of data types and variables, such as patient charts, hospital records, laboratory test results, radiological imaging, histologic and histopathologic analyses, genomic profiling, and electronic health records [110]. These sources contain structured data, including clinical tests [151], semi-structured data like patient questionnaires [152], and unstructured data, such as physician's medical records [153].

The integration of such multimodal data has significantly enhanced the robustness and accuracy of diagnostic or prognostic models, driving advancements in AI applications within clinical settings [154]. MOMLN, an advanced multimodal and multi-omics machine learning integration framework, has demonstrated exceptional

predictive performance by incorporating comprehensive input data, including clinical characteristics, DNA mutation profiles, gene expression signatures, TME features, and molecular pathway information. This framework achieved a remarkable mean AUC of 0.989 in classifying drug response types among 147 breast cancer patients [97].

Concluding remarks

Artificial intelligence (AI), with its powerful data processing and analysis capabilities, has shown significant potential in both basic and clinical studies on tumor resistance. By analyzing clinical data and omics data, AI provides innovative perspectives and tools to understand the onset and progression of tumor drug resistance, driving advances in cancer prediction, treatment, and prognosis. Its successful application not only underscores the potential of AI in the medical field, but also points to new development directions for precision oncology.

However, the application of AI technology in tumor resistance practice still faces several major challenges, particularly the incompleteness and bias of medical data, model interpretability, and robustness. To tackle these recent challenges, it is essential to implement standardized protocols for data collection, integration, processing, analyzing, modelling and validation, and to focus on developing robust, interpretable AI systems. The adoption of emerging technologies, particularly multimodal AI models, can greatly advance tumor drug resistance research by enabling more effective synthesis and analysis of diverse data types. Clinical validation of AI models is also crucial to ensure their reliable application in real-world studies.

Despite these above challenges, AI is poised to play an increasingly pivotal role in mitigating tumor drug resistance in clinical practice as technology continues to advance and more comprehensive clinical data becomes available. In the near future, AI is expected to predict and combat tumor drug resistance with higher efficiency and precision and become an integral part of every stage of tumor screening strategy, patient management, and prognosis, thus realizing personalized treatment and precision oncology.

Abbreviations

AI	Artificial Intelligence
AUC	Area Under the Curve
BCa	Bladder Cancer
CDK4/6i	Cyclin-Dependent Kinase 4/6 Inhibitor
CDK12i	Cyclin-Dependent Kinase 12 Inhibitor
CIN	Chromosomal Instability
CRC	Colorectal Cancer
ctDNA	Circulating tumor DNA
CTRP	Cancer Therapeutics Response Portal
GDSC	Genomics of Drug Sensitivity in Cancer
ICB	Immune Checkpoint Blockade
ICI	Immune Checkpoint Inhibitors
PDX	Patient-Derived Xenograft
RF	Random Forest
RNA-seq	RNA sequence
scRNA-seq	Single-cell RNA sequencing
SVM	Support Vector Machine
TCGA	The Cancer Genome Atlas
TME	Tumor Microenvironment
VNN	Visible neural network

Acknowledgements

Yes.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Authors' contributions

Y.M. and DG.SG.: Literature search,wrote the main manuscript text. Q.H. L.X. and DS.C: Literature search. H.Z and YK.W: Conceptualization, Funding and revision of manuscripts.

Funding

This work was supported by the National Natural Science Foundation of China [No. 82003929], Provincial Natural Science Foundation of Hunan [No. 2023JJ40857] and Scientific Research Program of Hunan Provincial Department of education [No. 23A0665].

Data availability

No datasets were generated or analysed during the current study.

Declarations

Ethics approval and consent to participate

Not applicable.

Competing interests

The authors declare no competing interests.

Author details

[1]Present Address: Hunan Cancer Hospital/The Affiliated Cancer Hospital of Xiangya School of Medicine, Central South University, Changsha, Hunan, China. [2]Xiangya School of Pharmaceutical Sciences, Central South University, Changsha, China. [3]Department of Pharmacy, Xiangya Hospital, Central South University, Changsha, Hunan, China. [4]Department of Histology and Embryology of Xiangya School of Medicine, Central South University, Changsha, Hunan, People's Republic of China. [5]Department of Lymphoma and Hematology, The Affiliated Cancer Hospital of Xiangya School of Medicine, Central South University, Changsha, Hunan, People's Republic of China. [6]Department of Lymphoma and Hematology, Hunan Cancer Hospital, Changsha, Hunan, People's Republic of China.

Received: 7 February 2025 Accepted: 2 April 2025
Published online: 23 April 2025

References

1. Vasan N, Baselga J, Hyman DM. A view on drug resistance in cancer. Nature. 2019;575:299–309.
2. Liu R, Sun M, Zhang G, Lan Y, Yang Z. Towards early monitoring of chemotherapy-induced drug resistance based on single cell metabolomics: combining single-probe mass spectrometry with machine learning. Anal Chim Acta. 2019;1092:42–8.
3. Bukowski K, Kciuk M, Kontek R. Mechanisms of multidrug resistance in cancer chemotherapy. Int J Mol Sci. 2020;21:3233.
4. Rahmanian M, Seyfoori A, Ghasemi M, Shamsi M, Kolahchi AR, Modarres HP, Sanati-Nezhad A, Majidzadeh-A K. In-vitro tumor microenvironment models containing physical and biological barriers for modelling multidrug resistance mechanisms and multidrug delivery strategies. J Control Release. 2021;334:164–77.

5. Kang JH, Heo MH, Kim HK, Cho J, Kim Y, Lee H, Lee S-C, Park SH. Development of a novel patient-derived preclinical model from malignant effusions in patients with tyrosine kinase inhibitor resistant clear cell renal cell carcinoma. J Clin Oncol. 2017;35:445–445.
6. Spalato Ceruso M, Toulmonde M, Blouin L, Italiano A. Circulating tumour DNA sequencing reveal mechanisms of resistance to BRAF-targeted therapies in BRAF-mutated gastrointestinal stromal tumour. Eur J Cancer. 2023;179:25–7.
7. Lin Q, Wu Z, Yue X, Yu X, Wang Z, Song X, Xu L, He Y, Ge Y, Tan S, et al. ZHX2 restricts hepatocellular carcinoma by suppressing stem cell-like traits through KDM2A-mediated H3K36 demethylation. EBioMedicine. 2020;53:102676.
8. López López V, Serrano A, Fuentes A, Ferrer Lores B, Chaves JF, Solano C, Terol MJ. Liquid biopsy and lymphoma monitoring in clinical practice. Blood. 2021;138:4483–4483.
9. Jia Q, Chu H, Jin Z, Long H, Zhu B. High-throughput single-cell sequencing in cancer research. Signal Transduct Target Ther. 2022;7:145.
10. Beekhof R, Bertotti A, Böttger F, Vurchio V, Cottino F, Zanella ER, Migliardi G, Viviani M, Grassi E, Lupo B, et al. Phosphoproteomics of patient-derived xenografts identifies targets and markers associated with sensitivity and resistance to EGFR blockade in colorectal cancer. Sci Transl Med. 2023;15:eabm3687.
11. Rudqvist N-P, Charpentier M, Lhuillier C, Wennerberg E, Spada S, Sheridan C, Zhou XK, Zhang T, Formenti SC, Sims JS, et al. Immunotherapy targeting different immune compartments in combination with radiation therapy induces regression of resistant tumors. Nat Commun. 2023;14:5146.
12. Bury D, Martin-Hirsch PL, Martin FL, Dawson TP. Are new technologies translatable to point-of-care testing? The Lancet. 2017;390:2765–6.
13. Bueschbell B, Caniceiro AB, Suzano PMS, Machuqueiro M, Rosário-Ferreira N, Moreira IS. Network biology and artificial intelligence drive the understanding of the multidrug resistance phenotype in cancer. Drug Resist Updates. 2022;60:100811.
14. Fraunhoffer N, Hammel P, Conroy T, Nicolle R, Bachet JB, Harlé A, Rebours V, Turpin A, Ben Abdelghani M, Mitry E, et al. Development and validation of AI-assisted transcriptomic signatures to personalize adjuvant chemotherapy in patients with pancreatic ductal adenocarcinoma. Ann Oncol. 2024;35:780–91.
15. Hajjo R, Sabbah DA, Bardaweel SK, Tropsha A. Identification of tumor-specific MRI biomarkers using Machine Learning (ML). Diagnostics. 2021;11:742.
16. Ballard Jenna L, Wang Z, Li W, Shen L, Long Q. Deep learning-based approaches for multi-omics data integration and analysis. BioData Mining. 2024;17:38.
17. Rathore S, Nasrallah M, Akbari H, Shukla G, Bagley S, Watt C, Min Ha S, Mamourian E, Sako C, Binder Z, et al. TMOD-40. in vivo evaluation of o6-methylguanine-dna-methyltransferase (mgmt) promoter methylation status for de novo glioblastoma patients using deep learning features. Neuro Oncol. 2019;21:vi271–2.
18. Salahi A, Honrado C, Moore J, Adair S, Bauer TW, Swami NS. Supervised learning on impedance cytometry data for label-free biophysical distinction of pancreatic cancer cells versus their associated fibroblasts under gemcitabine treatment. Biosens Bioelectron. 2023;231:115262.
19. Mu Q, Chai R, Yang Y, Jiang T, Wang J. EPCO-49. Myc amplification at diagnosis regulates therapy-induced hypermutation at recurrent glioma. Neuro Oncol. 2023;25:v135–v135.
20. Mu Q, Chai R, Pang B, Yang Y, Liu H, Zhao Z, Bao Z, Song D, Zhu Z, Yan M, et al. Identifying predictors of glioma evolution from longitudinal sequencing. Sci Transl Med. 2023;15:eadh4181.
21. Huang H, Li Q, Tu X, Yu D, Zhou Y, Ma L, Wei K, Gao Y, Zhao G, Han R, et al. DNA hypomethylation patterns and their impact on the tumor microenvironment in colorectal cancer. Cell Oncol. 2024;47:1375–89.
22. Thomas M, Theophanous S, Whittle I, Finetti M, Pollock S, Ahmed N, Tanner G, Chakrabarty A, Ismail A, Bulpitt A, Stead L. Predicting glioblastoma gene expression therapy response with machine learning. Neuro Oncol. 2023;25:iii13–4.
23. Suh Y-S, Lee J, George J, Seol D, Jeong K, Oh S-Y, Bang C, Jun Y, Kong S-H, Lee H-J, et al. RNA expression of 6 genes from metastatic mucosal gastric cancer serves as the global prognostic marker for gastric cancer with functional validation. Br J Cancer. 2024;130:1571–84.
24. Tang X, Thompson KJ, Kalari KR, Sinnwell JP, Suman VJ, Vedell PT, McLaughlin SA, Northfelt DW, Aspitia AM, Gray RJ, et al. Integration of multiomics data shows down regulation of mismatch repair and tubulin pathways in triple-negative chemotherapy-resistant breast tumors. Breast Cancer Res. 2023;25:57.
25. Lee J, Kim D, Kong J, Ha D, Kim I, Park M, Lee K, Im S-H, Kim S. Cell-cell communication network-based interpretable machine learning predicts cancer patient response to immune checkpoint inhibitors. Sci Adv. 2024;10:eadj0785.
26. Hu D, Shen X, Gao P, Mao T, Chen Y, Li X, Shen W, Zhuang Y, Ding J. Multi-omic profiling reveals potential biomarkers of hepatocellular carcinoma prognosis and therapy response among mitochondria-associated cell death genes in the context of 3P medicine. EPMA Journal. 2024;15:321–43.
27. Harel M, Lahav C, Jacob E, Issler E, Bar H, Dicker A, Sharon O, Bacchiocchi A, Halaban R, Sznol M, Shaked Y. 21 Plasma-based proteomic profiling as a tool for predicting response to immunotherapy in melanoma patients. J Immunother Cancer. 2020;8:A11–2.
28. Hwangbo S, Kim SI, Kim J-H, Eoh KJ, Lee C, Kim YT, Suh D-S, Park T, Song YS. Development of machine learning models to predict platinum sensitivity of high-grade serous ovarian carcinoma. Cancers. 1875;2021:13.
29. Volinsky-Fremond S, Horeweg N, Andani S, Barkey Wolf J, Lafarge MW, de Kroon CD, Ørtoft G, Høgdall E, Dijkstra J, Jobsen JJ, et al. Prediction of recurrence risk in endometrial cancer with multimodal deep learning. Nat Med. 2024;30:1962–73.
30. Pour AF, Wu TC, Martinek J, Kumar S, Anzeneder T, Torkler S, Sommermeyer F, Palucka K, Chuang J: Abstract P1–02–01: Deep learning identifies morphological changes in whole slide images of treatment-resistant TNBC. Cancer Research 2022, 82:P1–02–01-P01–02–01.
31. Barrera C, Corredor G, Viswanathan VS, Ding R, Toro P, Fu P, Buzzy C, Lu C, Velu P, Zens P, et al. Deep computational image analysis of immune cell niches reveals treatment-specific outcome associations in lung cancer. npj Precis Oncol. 2023;7:52.
32. Jiang P, Sinha S, Aldape K, Hannenhalli S, Sahinalp C, Ruppin E. Big data in basic and translational cancer research. Nat Rev Cancer. 2022;22:625–39.
33. Kann BH, Hosny A, Aerts HJWL. Artificial intelligence for clinical oncology. Cancer Cell. 2021;39:916–27.
34. Liu M, Li S, Yuan H, Ong MEH, Ning Y, Xie F, Saffari SE, Shang Y, Volovici V, Chakraborty B, Liu N. Handling missing values in healthcare data: a systematic review of deep learning-based imputation techniques. Artif Intell Med. 2023;142:102587.
35. Boehm KM, Khosravi P, Vanguri R, Gao J, Shah SP. Harnessing multimodal data integration to advance precision oncology. Nat Rev Cancer. 2022;22:114–26.
36. Albahra S, Gorbett T, Robertson S, D'Aleo G, Kumar SVS, Ockunzzi S, Lallo D, Hu B, Rashidi HH. Artificial intelligence and machine learning overview in pathology & laboratory medicine: a general review of data preprocessing and basic supervised concepts. Semin Diagn Pathol. 2023;40:71–87.
37. Firoozbakht F, Yousefi B, Schwikowski B. An overview of machine learning methods for monotherapy drug response prediction. Brief Bioinform. 2021;23:bbab408.
38. Yao Y, Lv Y, Tong L, Liang Y, Xi S, Ji B, Zhang G, Li L, Tian G, Tang M, et al. ICSDA: a multi-modal deep learning model to predict breast cancer recurrence and metastasis risk by integrating pathological, clinical and gene expression data. Brief Bioinform. 2022;23:bbab408.
39. Álvez MB, Edfors F, von Feilitzen K, Zwahlen M, Mardinoglu A, Edqvist P-H, Sjöblom T, Lundin E, Rameika N, Enblad G, et al. Next generation pan-cancer blood proteome profiling using proximity extension assay. Nat Commun. 2023;14:4308.
40. Xia X, Zhu C, Zhong F, Liu L. TransCDR: a deep learning model for enhancing the generalizability of drug activity prediction through transfer learning and multimodal data fusion. BMC Biol. 2024;22:227.
41. Ahn B, Moon D, Kim HS, Lee C, Cho NH, Choi HK, Kim D, Lee JY, Nam EJ, Won D, et al. Histopathologic image–based deep learning classifier for predicting platinum-based treatment responses in high-grade serous ovarian cancer. Nat Commun. 2024;15:4253.
42. Gunning D, Stefik M, Choi J, Miller T, Stumpf S, Yang GZ. XAI—Explainable artificial intelligence. Sci Robot. 2019;4:eaay7120.

43. Tang S, Zhang H, Liang J, Tang S, Li L, Li Y, Xu Y, Wang D, Zhou Y. Prostate cancer treatment recommendation study based on machine learning and SHAP interpreter. Cancer Sci. 2024;115:3755–66.
44. Guo Q-H, Xie F-C, Zhong F-M, Wen W, Zhang X-R, Yu X-J, Wang X-L, Huang B, Li L-P, Wang X-Z. Application of interpretable machine learning algorithms to predict distant metastasis in ovarian clear cell carcinoma. Cancer Med. 2024;13:e7161.
45. Sahu A, Wang X, Munson P, Klomp JPG, Wang X, Gu SS, Han Y, Qian G, Nicol P, Zeng Z, et al. Discovery of targets for immune-metabolic antitumor drugs identifies estrogen-related receptor alpha. Cancer Discov. 2023;13:672–701.
46. Sui Q, Chen Z, Hu Z, Huang Y, Liang J, Bi G, Bian Y, Zhao M, Zhan C, Lin Z, et al. Cisplatin resistance-related multi-omics differences and the establishment of machine learning models. J Transl Med. 2022;20:171.
47. Cai T, Feng T, Li G, Wang J, Jin S, Ye D, Zhu Y. Deciphering the prognostic features of bladder cancer through gemcitabine resistance and immune-related gene analysis and identifying potential small molecular drug PIK-75. Cancer Cell Int. 2024;24:125.
48. Yujia X, Zhangsheng Y. Thorny but rosy: prosperities and difficulties in â□¯AI plus medicineâ□™ concerning data collection, model construction and clinical deployment. Gen Psychiatry. 2024;37:e101436.
49. Kadioglu O, Efferth T. A machine learning-based prediction platform for P-glycoprotein modulators and its validation by molecular docking. Cells. 2019;8:1286.
50. Xie F, Zhang J, Wang J, Reuben A, Xu W, Yi X, Varn FS, Ye Y, Cheng J, Yu M, et al. Multifactorial deep learning reveals pan-cancer genomic tumor clusters with distinct immunogenomic landscape and response to immunotherapy. Clin Cancer Res. 2020;26:2908–20.
51. Ren S, Cooper GF, Chen L, Lu X. An interpretable deep learning framework for genome-informed precision oncology. Nat Machine Intellig. 2024;6:864–75.
52. Corridore S, Verreault M, Martin H, Delobel T, Carrère C, Idbaih A, Ballesta A: Circumventing glioblastoma resistance to temozolomide through optimal drug combinations designed by quantitative systems pharmacology and machine learning. bioRxiv 2024:2024.2005.2031.596811.
53. Munson BP, Chen M, Bogosian A, Kreisberg JF, Licon K, Abagyan R, Kuenzi BM, Ideker T. De novo generation of multi-target compounds using deep generative chemistry. Nat Commun. 2024;15:3636.
54. Wen T, Li L, Li Y, Wang J, Gao P, Xie G, Ma F. A deep learning approach to discover cyclin-dependent kinases 12 (CDK12) inhibitors in breast cancer. J Clin Oncol. 2022;40:e15086–e15086.
55. Gerratana L, Reduzzi C, Davis AA, Velimirovic M, Clifton K, Hensing WL, Shah AN, Dai CS, D'Amico P, Donahue J, et al. Defining resistance mechanisms to CDK4/6 inhibition in hormone receptor-positive HER2-negative metastatic breast cancer (MBC) through a machine learning approach applied to circulating tumor DNA (ctDNA). J Clin Oncol. 2022;40:3055–3055.
56. Misek SA, Appleton KM, Dexheimer TS, Lisabeth EM, Lo RS, Larsen SD, Gallo KA, Neubig RR. Rho-mediated signaling promotes BRAF inhibitor resistance in de-differentiated melanoma cells. Oncogene. 2020;39:1466–83.
57. Chapple RH, Liu X, Natarajan S, Alexander MIM, Kim Y, Patel AG, LaFlamme CW, Pan M, Wright WC, Lee H-M, et al. An integrated single-cell RNA-seq map of human neuroblastoma tumors and preclinical models uncovers divergent mesenchymal-like gene expression programs. Genome Biol. 2024;25:161.
58. Wiecek AJ, Cutty SJ, Kornai D, Parreno-Centeno M, Gourmet LE, Tagliazucchi GM, Jacobson DH, Zhang P, Xiong L, Bond GL, et al. Genomic hallmarks and therapeutic implications of G0 cell cycle arrest in cancer. Genome Biol. 2023;24:128.
59. Zhang Z, Wang Z-X, Chen Y-X, Wu H-X, Yin L, Zhao Q, Luo H-Y, Zeng Z-L, Qiu M-Z, Xu R-H. Integrated analysis of single-cell and bulk RNA sequencing data reveals a pan-cancer stemness signature predicting immunotherapy response. Genome Med. 2022;14:45.
60. Yang ZY, Ye ZF, Xiao YJ, Hsieh CY, Zhang SY. SPLDExtraTrees: robust machine learning approach for predicting kinase inhibitor resistance. Brief Bioinform. 2022;23:bbac050.
61. Zhao X, Singhal A, Park S, Kong J, Bachelder R, Ideker T. Cancer mutations converge on a collection of protein assemblies to predict resistance to replication stress. Cancer Discov. 2024;14:508–23.
62. Park S, Silva E, Singhal A, Kelly MR, Licon K, Panagiotou I, Fogg C, Fong S, Lee JJY, Zhao X, et al. A deep learning model of tumor cell architecture elucidates response and resistance to CDK4/6 inhibitors. Nat Cancer. 2024;5:996–1009.
63. Zhou Q-n, Lei R-e, Liang Y-x, Li S-q, Guo X-w, Hu B-l. Oxaliplatin related lncRNAs prognostic models predict the prognosis of patients given oxaliplatin-based chemotherapy. Cancer Cell Int. 2023;23:103.
64. Chen D, Liu J, Zang L, Xiao T, Zhang X, Li Z, Zhu H, Gao W, Yu X. Integrated machine learning and bioinformatic analyses constructed a novel stemness-related classifier to predict prognosis and immunotherapy responses for hepatocellular carcinoma patients. Int J Biol Sci. 2022;18:360–73.
65. Yeh S-J, Paithankar S, Chen R, Xing J, Sun M, Liu K, Zhou J, Chen B: TransCell: In silico characterization of genomic landscape and cellular responses by deep transfer learning. Gen Proteo Bioinform 2024, 22.
66. Nasimian A, Ahmed M, Hedenfalk I, Kazi JU. A deep tabular data learning model predicting cisplatin sensitivity identifies BCL2L1 dependency in cancer. Comput Struct Biotechnol J. 2023;21:956–64.
67. Kong J, Lee H, Kim D, Han SK, Ha D, Shin K, Kim S. Network-based machine learning in colorectal and bladder organoid models predicts anti-cancer drug efficacy in patients. Nat Commun. 2020;11:5485.
68. Zeng Z, Gu SS, Wong CJ, Yang L, Ouardaoui N, Li D, Zhang W, Brown M, Liu XS. Machine learning on syngeneic mouse tumor profiles to model clinical immunotherapy response. Sci Adv. 2022;8:eabm8564.
69. Cotler MJ, Ramadi KB, Hou X, Christodoulopoulos E, Ahn S, Bashyam A, Ding H, Larson M, Oberg AL, Whittaker C, et al. Machine-learning aided in situ drug sensitivity screening predicts treatment outcomes in ovarian PDX tumors. Transl Oncol. 2022;21:101427.
70. Ji G, Liu J, Zhao Z, Lan J, Yang Y, Wang Z, Feng H, Ji K, Jiang X, Xia H, et al: Polyamine Anabolism Promotes Chemotherapy-Induced Breast Cancer Stem Cell Enrichment. Advanced Science, n/a:2404853.
71. Davis N, McKinney MS, Reddy A, Love C, Smith E, Happ L, Dave S. Novel mechanisms for resistance to targeted therapy identified through machine learning approaches in 1167 RNA-seq drug exposure profiles in lymphoma. Blood. 2018;132:1370–1370.
72. He L, Bulanova D, Oikkonen J, Häkkinen A, Zhang K, Zheng S, Wang W, Erkan EP, Carpén O, Joutsiniemi T, et al. Network-guided identification of cancer-selective combinatorial therapies in ovarian cancer. Brief Bioinform. 2021;22:bbab272.
73. Zhou J-B, Tang D, He L, Lin S, Lei JH, Sun H, Xu X, Deng C-X. Machine learning model for anti-cancer drug combinations: analysis, prediction, and validation. Pharmacol Res. 2023;194:106830.
74. Lamballe F, Ahmad F, Vinik Y, Castellanet O, Daian F, Müller A-K, Köhler UA, Bailly A-L, Josselin E, Castellano R, et al. Modeling heterogeneity of triple-negative breast cancer uncovers a novel combinatorial treatment overcoming primary drug resistance. Adv Sci. 2021;8:2003049.
75. Fang X, Yu WY, Zhu CM, Zhao N, Zhao W, Xie TT, Wei LJ, Sun XR, Xie J, Zhao Y. Chromosome instability functions as a potential therapeutic reference by enhancing chemosensitivity to BCL-XL inhibitors in colorectal carcinoma. Acta Pharmacol Sinic. 2024;45(11):2420–31.
76. Asada K, Takasawa K, Machino H, Takahashi S, Shinkai N, Bolatkan A, Kobayashi K, Komatsu M, Kaneko S, Okamoto K, Hamamoto R. Single-cell analysis using machine learning techniques and its application to medical research. Biomedicines. 2021;9:1513.
77. Zou Y, Xie J, Zheng S, Liu W, Tang Y, Tian W, Deng X, Wu L, Zhang Y, Wong C-W, et al. Leveraging diverse cell-death patterns to predict the prognosis and drug sensitivity of triple-negative breast cancer patients after surgery. Int J Surg. 2022;107:106936.
78. Gu H, Yin X, Peng T, Pan Y, Cui H, Li Z, Sun W, Ding B, Hu X, Zhang Z, Liu Z. Geographical origin identification and chemical markers screening of Chinese green tea using two-dimensional fingerprints technique coupled with multivariate chemometric methods. Food Control. 2022;135:108795.
79. Ran D, Moharil J, Lu J, Gustafson H, Culm-Merdek K, Strand-Tibbitts K, Benjamin L, Navratil M. Platform comparison of HTG EdgeSeq and RNA-Seq for gene expression profiling of tumor tissue specimens. J Clin Oncol. 2020;38:3566–3566.

80. Hafemeister C, Satija R. Normalization and variance stabilization of single-cell RNA-seq data using regularized negative binomial regression. Genome Biol. 2019;20:296.
81. Garcia PL, Miller AL, Yoon KJ. Patient-derived xenograft models of pancreatic cancer: overview and comparison with other types of models. Cancers. 2020;12:1327.
82. Liu Y, Wu W, Cai C, Zhang H, Shen H, Han Y. Patient-derived xenograft models in cancer therapy: technologies and applications. Signal Transduct Target Ther. 2023;8:160.
83. Lee J, Lee G, Park HS, Jeong B-K, Gong G, Jeong JH, Lee HJ. Factors associated with engraftment success of patient-derived xenografts of breast cancer. Breast Cancer Res. 2024;26:49.
84. Pellecchia S, Viscido G, Franchini M, Gambardella G. Predicting drug response from single-cell expression profiles of tumours. BMC Med. 2023;21:476.
85. Yang C, Yang C, Yarden Y, To KKW, Fu L. The prospects of tumor chemosensitivity testing at the single-cell level. Drug Resist Updates. 2021;54:100741.
86. Mitra AK, Mukherjee U, Harding T, Stessman H, Li Y, Jin J, Kumar SK, Rajkumar SV, Van Ness BG. Scattome: a single-cell analysis of targeted transcriptome program to predict drug sensitivity of single cells within human myeloma tumors. Blood. 2015;126:4249–4249.
87. Zhao Z, Chen X, Dowbaj AM, Sljukic A, Bratlie K, Lin L, Fong ELS, Balachander GM, Chen Z, Soragni A, et al. Organoids. Nat Rev Methods Prim. 2022;2:94.
88. Bai L, Wu Y, Li G, Zhang W, Zhang H, Su J. AI-enabled organoids: Construction, analysis, and application. Bioactive Mater. 2024;31:525–48.
89. Güvenç Paltun B, Kaski S, Mamitsuka H. Machine learning approaches for drug combination therapies. Brief Bioinform. 2021;22:bbab293.
90. Ianevski A, Nader K, Driva K, Senkowski W, Bulanova D, Moyano-Galceran L, Ruokoranta T, Kuusanmäki H, Ikonen N, Sergeev P, et al. Single-cell transcriptomes identify patient-tailored therapies for selective co-inhibition of cancer clones. Nat Commun. 2024;15:8579.
91. Ianevski A, Lahtela J, Javarappa KK, Sergeev P, Ghimire BR, Gautam P, Vähä-Koskela M, Turunen L, Linnavirta N, Kuusanmäki H, et al. Patient-tailored design for selective co-inhibition of leukemic cell subpopulations. Sci Adv. 2021;7:eabe4038.
92. Wang H, Wang R, Luo L, Hong J, Chen X, Shen K, Wang Y, Huang R, Wang Z. An exosome-based specific transcriptomic signature for profiling regulation patterns and modifying tumor immune microenvironment infiltration in triple-negative breast cancer. Front Immunol. 2023;14:1295558.
93. Wang Z, Boretto M, Millen R, Natesh N, Reckzeh ES, Hsu C, Negrete M, Yao H, Quayle W, Heaton BE, et al. Rapid tissue prototyping with micro-organospheres. Stem Cell Reports. 2022;17:1959–75.
94. Goldstein Y, Cohen OT, Wald O, Bavli D, Kaplan T, Benny O. Particle uptake in cancer cells can predict malignancy and drug resistance using machine learning. Sci Adv. 2024;10:eadj4370.
95. Hostallero DE, Wei L, Wang L, Cairns J, Emad A. Preclinical-to-clinical anti-cancer drug response prediction and biomarker identification using TINDL. Gen Proteom Bioinform. 2023;21:535–50.
96. Mitra AK, Mukherjee UK, Harding T, Jang JS, Stessman H, Li Y, Abyzov A, Jen J, Kumar S, Rajkumar V, Van Ness B. Single-cell analysis of targeted transcriptome predicts drug sensitivity of single cells within human myeloma tumors. Leukemia. 2016;30:1094–102.
97. Rashid MM, Selvarajoo K. Advancing drug-response prediction using multi-modal and -omics machine learning integration (MOMLIN): a case study on breast cancer clinical data. Brief Bioinform. 2024;25:bbae300.
98. Che G, Yin J, Wang W, Luo Y, Chen Y, Yu X, Wang H, Liu X, Chen Z, Wang X, et al. Circumventing drug resistance in gastric cancer: a spatial multi-omics exploration of chemo and immuno-therapeutic response dynamics. Drug Resist Updates. 2024;74:101080.
99. Blatti C, de la Fuente Js, Gao H, Marín-Goñi I, Chen Z, Zhao SD, Tan W, Weinshilboum R, Kalari KR, Wang L, Hernaez M: Bayesian machine learning enables identification of transcriptional network disruptions associated with drug-resistant prostate cancer. Cancer Res 2023, 83:1361-1380.
100. Zhong F, Zhang X, Wang Z, Li X, Huang B, Kong G, Wang X. The therapeutic and biomarker significance of ferroptosis in chronic myeloid leukemia. Front Immunol. 2024;15:1402669.
101. Shi Y, Li C, Zhang X, Peng C, Sun P, Zhang Q, Wu L, Ding Y, Xie D, Xu Z, Zhu W. D3EGFR: a webserver for deep learning-guided drug sensitivity prediction and drug response information retrieval for EGFR mutation-driven lung cancer. Brief Bioinform. 2024;25:bbae121.
102. Zhou Y, Liu Z, Gong C, Zhang J, Zhao J, Zhang X, Liu X, Li B, Li R, Shi Z, et al. Targeting treatment resistance: unveiling the potential of RNA methylation regulators and TG-101,209 in pan-cancer neoadjuvant therapy. J Exp Clin Cancer Res. 2024;43:232.
103. Guan J, Xu X, Qiu G, He C, Lu X, Wang K, Liu X, Li Y, Ling Z, Tang X, et al. Cellular hierarchy framework based on single-cell/multi-patient sample sequencing reveals metabolic biomarker PYGL as a therapeutic target for HNSCC. J Exp Clin Cancer Res. 2023;42:162.
104. Cardoso MR, Silva AAR, Talarico MCR, Sanches PHG, Sforça ML, Rocco SA, Rezende LM, Quintero M, Costa TBBC, Viana LR, et al. Metabolomics by NMR combined with machine learning to predict neoadjuvant chemotherapy response for breast cancer. Cancers. 2022;14:5055.
105. Chen Q, Gao F, Wu J, Zhang K, Du T, Chen Y, Cai R, Zhao D, Deng R, Tang J. Comprehensive pan-cancer analysis of mitochondrial outer membrane permeabilisation activity reveals positive immunomodulation and assists in identifying potential therapeutic targets for immunotherapy resistance. Clin Transl Med. 2024;14:e1735.
106. Shi R, Sun J, Zhou H, Hu T, Gao Z, Wang X, Li M, Zhou Z, Shu Y. Hypoxia within tumor microenvironment characterizes distinct genomic patterns and aids molecular subtyping for guiding individualized immunotherapy. J Big Data. 2024;11:81.
107. Qian L, Zhu J, Xue Z, Zhou Y, Xiang N, Xu H, Sun R, Gong W, Cai X, Sun L, et al. Proteomic landscape of epithelial ovarian cancer. Nat Commun. 2024;15:6462.
108. Bonstingl L, Zinnegger M, Sallinger K, Pankratz K, Pritz E, Odar C, Skofler C, Ulz C, Oberauner-Wappis L, Borrás-Cherrier A, et al: Circulating tumor cell characterization and classification by novel combinatorial dual-color (CoDuCo) in situ hybridization and supervised machine learning. bioRxiv 2024:2024.2005.2008.592946.
109. Wang Z, Wang Y, Yang T, Xing H, Wang Y, Gao L, Guo X, Xing B, Wang Y, Ma W. Machine learning revealed stemness features and a novel stemness-based classification with appealing implications in discriminating the prognosis, immunotherapy and temozolomide responses of 906 glioblastoma patients. Brief Bioinform. 2021;22:bbab032.
110. Shaked Y, Harel M, Lahav C, Jacob E, Sela I, Yahalom G, Elon Y, Sharon O, Kamer I, Bar H, et al. Integration of proteomic and clinical data for the prediction of response to immune checkpoint inhibitor therapy in non-small cell lung cancer. J Clin Oncol. 2021;39:e21110–e21110.
111. Saghand PG, Naqa IE, Tan AC, Xie M, Dai D, Chen JL, Ratan A, McCarter M, Carpten JD, Shah H, et al. A deep learning approach utilizing clinical and molecular data for identifying prognostic biomarkers in patients treated with immune checkpoint inhibitors: An ORIEN pan-cancer study. J Clin Oncol. 2022;40:2619–2619.
112. Harel M, Lahav C, Jacob E, Dahan N, Sela I, Elon Y, Raveh Shoval S, Yahalom G, Kamer I, Zer A, et al. Longitudinal plasma proteomic profiling of patients with non-small cell lung cancer undergoing immune checkpoint blockade. J Immunother Cancer. 2022;10:e004582.
113. Scanlon E, Lavery A, Stevenson L, Kennedy C, Byrne R, Walker A, Eatock M, Middleton M, Thomas A, Turkington R: P-OGC08 Translational Insights from the Dual ErbB Inhibition in Oesophago-gastric Cancer (DEBIOC) Clinical Trial - a bioinformatic analysis. British J Surg 2021, 108.
114. Alnuhait MA, Shahbar AN, Alrumaih I, Alzahrani T, Alzahrani A, alanizi A, Alrashed MA, Elrggal M, Alhuthali A, Alsuhebany N: Advancing cancer care: How artificial intelligence is transforming oncology pharmacy. Inform Med Unlocked 2024, 50:101529.
115. Lipkova J, Chen RJ, Chen B, Lu MY, Barbieri M, Shao D, Vaidya AJ, Chen C, Zhuang L, Williamson DFK, et al. Artificial intelligence for multimodal data integration in oncology. Cancer Cell. 2022;40:1095–110.
116. Sinha S, Vegesna R, Mukherjee S, Kammula AV, Dhruba SR, Wu W, Kerr DL, Nair NU, Jones MG, Yosef N, et al. PERCEPTION predicts patient response and resistance to treatment using single-cell transcriptomics of their tumors. Nature Cancer. 2024;5:938–52.
117. Shaked Y, Harel M, Lahav C, Yellini B, Tepper E, Wolf I, Harkovsky T, Leibowitz R, Gottfried M, Abu-Amana M, et al. Personalized approach for response prediction and treatment management for non-small cell lung cancer patients based on a liquid biopsy. J Clin Oncol. 2022;40:e21132–e21132.

118. Zhou Y, Tao L, Qiu J, Xu J, Yang X, Zhang Y, Tian X, Guan X, Cen X, Zhao Y. Tumor biomarkers for diagnosis, prognosis and targeted therapy. Signal Transduct Target Ther. 2024;9:132.
119. Lou E, Vogel RI, Hoostal S, Klein M, Linden MA, Teoh D, Geller MA. Tumor-stroma proportion as a predictive biomarker of resistance to platinum-based chemotherapy in patients with ovarian cancer. JAMA Oncol. 2019;5:1222–4.
120. Long L, Assaraf YG, Lei Z-N, Peng H, Yang L, Chen Z-S, Ren S. Genetic biomarkers of drug resistance: a compass of prognosis and targeted therapy in acute myeloid leukemia. Drug Resist Updates. 2020;52:100703.
121. Aaltonen LA, Abascal F, Abeshouse A, Aburatani H, Adams DJ, Agrawal N, Ahn KS, Ahn S-M, Aikata H, Akbani R, et al. Pan-cancer analysis of whole genomes. Nature. 2020;578:82–93.
122. McLendon R, Friedman A, Bigner D, Van Meir EG, Brat DJ, M. Mastrogianakis G, Olson JJ, Mikkelsen T, Lehman N, Aldape K, et al. Comprehensive genomic characterization defines human glioblastoma genes and core pathways. Nature. 2008;455:1061–8.
123. Yang W, Soares J, Greninger P, Edelman EJ, Lightfoot H, Forbes S, Bindal N, Beare D, Smith JA, Thompson IR, et al. Genomics of Drug Sensitivity in Cancer (GDSC): a resource for therapeutic biomarker discovery in cancer cells. Nucleic Acids Res. 2012;41:D955–61.
124. Liu X, Yi J, Li T, Wen J, Huang K, Liu J, Wang G, Kim P, Song Q, Zhou X. DRMref: comprehensive reference map of drug resistance mechanisms in human cancer. Nucleic Acids Res. 2023;52:D1253–64.
125. Seashore-Ludlow B, Rees MG, Cheah JH, Cokol M, Price EV, Coletti ME, Jones V, Bodycombe NE, Soule CK, Gould J, et al. Harnessing connectivity in a large-scale small-molecule sensitivity dataset. Cancer Discov. 2015;5:1210–23.
126. Cao X, Zhou X, Hou F, Huang YE, Yuan M, Long M, Chen S, Lei W, Zhu J, Chen J, et al. ncRNADrug: a database for validated and predicted ncRNAs associated with drug resistance and targeted by drugs. Nuc Acids Res. 2023;52:D1393–9.
127. Xu Y, Zhang Y, Song K, Liu J, Zhao R, Zhang X, Pei L, Li M, Chen Z, Zhang C, et al. ScDrugAct: a comprehensive database to dissect tumor microenvironment cell heterogeneity contributing to drug action and resistance across human cancers. Nucleic Acids Res. 2024;53(D1):D1536–46.
128. Sun X, Zhang Y, Li H, Zhou Y, Shi S, Chen Z, He X, Zhang H, Li F, Yin J, et al. DRESIS: the first comprehensive landscape of drug resistance information. Nucleic Acids Res. 2023;51:D1263-d1275.
129. Jiang J, Ma Y, Yang L, Ma S, Yu Z, Ren X, Kong X, Zhang X, Li D, Liu Z. CTR-DB 2.0: an updated cancer clinical transcriptome resource, expanding primary drug resistance and newly adding acquired resistance datasets and enhancing the discovery and validation of predictive biomarkers. Nuc Acids Res. 2024;53(D1):D1335–47.
130. Kumar R, Chaudhary K, Gupta S, Singh H, Kumar S, Gautam A, Kapoor P, Raghava GPS. CancerDR: cancer drug resistance database. Sci Rep. 2013;3:1445.
131. Yang Z, Ye Z, Qiu J, Feng R, Li D, Hsieh C, Allcock J, Zhang S. A mutation-induced drug resistance database (MdrDB). Commun Chemist. 2023;6:123.
132. Barretina J, Caponigro G, Stransky N, Venkatesan K, Margolin AA, Kim S, Wilson CJ, Lehár J, Kryukov GV, Sonkin D, et al. The cancer cell line encyclopedia enables predictive modelling of anticancer drug sensitivity. Nature. 2012;483:603–7.
133. Whirl-Carrillo M, Huddart R, Gong L, Sangkuhl K, Thorn CF, Whaley R, Klein TE. An evidence-based framework for evaluating pharmacogenomics knowledge for personalized medicine. Clin Pharmacol Ther. 2021;110:563–72.
134. Sondka Z, Dhir NB, Carvalho-Silva D, Jupe S, Madhumita, McLaren K, Starkey M, Ward S, Wilding J, Ahmed M, et al: COSMIC: a curated database of somatic variants and clinical data for cancer. Nucleic Acids Research 2023, 52:D1210-D1217.
135. Suehnholz SP, Nissan MH, Zhang H, Kundra R, Nandakumar S, Lu C, Carrero S, Dhaneshwar A, Fernandez N, Xu BW, et al. Quantifying the expanding landscape of clinical actionability for patients with cancer. Cancer Discov. 2024;14:49–65.
136. Wang C, Yang J, Luo H, Wang K, Wang Y, Xiao Z-X, Tao X, Jiang H, Cai H. CancerTracer: a curated database for intrapatient tumor heterogeneity. Nucleic Acids Res. 2019;48:D797–806.
137. Arango-Argoty G, Kipkogei E, Stewart R, Sun GJ, Patra A, Kagiampakis I, Jacob E. Pretrained transformers applied to clinical studies improve predictions of treatment efficacy and associated biomarkers. Nat Commun. 2025;16:2101.
138. Zhang H, Yang F, Xu Y, Zhao S, Jiang YZ, Shao ZM, Xiao Y. Multimodal integration using a machine learning approach facilitates risk stratification in HR+/HER2− breast cancer. Cell Rep Med. 2025;6:101924.
139. Ogunleye AZ, Piyawajanusorn C, Gonçalves A, Ghislat G, Ballester PJ. Interpretable machine learning models to predict the resistance of breast cancer patients to doxorubicin from their microRNA profiles. Adv Sci. 2022;9:2201501.
140. Rivera D, Lee JJ, Royce ME, Kluetz PG. FDA oncology center of excellence landscape analysis of real-world data submissions for oncology drugs. J Clin Oncol. 2021;39:e18787–e18787.
141. Tavolara TE, Su Z, Gurcan MN, Niazi MKK. One label is all you need: interpretable AI-enhanced histopathology for oncology. Semin Cancer Biol. 2023;97:70–85.
142. Allen GM, Lim WA. Rethinking cancer targeting strategies in the era of smart cell therapeutics. Nat Rev Cancer. 2022;22:693–702.
143. Tebon PJ, Wang B, Markowitz AL, Davarifar A, Tsai BL, Krawczuk P, Gonzalez AE, Sartini S, Murray GF, Nguyen HTL, et al. Drug screening at single-organoid resolution via bioprinting and interferometry. Nat Commun. 2023;14:3168.
144. Chan JY. Special issue "cancer immunotherapy: tumor microenvironment, biomarker discovery and immune resistance." Int J Mol Sci. 2024;25:5113.
145. Viswanathan VS, Parmar V, Madabhushi A. Towards equitable AI in oncology. Nat Rev Clin Oncol. 2024;21:628–37.
146. Yang J, Li A, Li Y, Guo X, Wang M. A novel approach for drug response prediction in cancer cell lines via network representation learning. Bioinformatics. 2018;35:1527–35.
147. You SC, Krumholz HM. the evolution of evidence-based medicine: when the magic of the randomized clinical trial meets real-world data. Circulation. 2022;145:107–9.
148. Elkhader J, Elemento O. Artificial intelligence in oncology: from bench to clinic. Semin Cancer Biol. 2022;84:113–28.
149. Zhang Z, Hernandez K, Savage J, Li S, Miller D, Agrawal S, Ortuno F, Staudt LM, Heath A, Grossman RL. Uniform genomic data analysis in the NCI genomic data commons. Nat Commun. 2021;12:1226.
150. Karim MR, Islam T, Shajalal M, Beyan O, Lange C, Cochez M, Rebholz-Schuhmann D, Decker S. Explainable ai for bioinformatics: methods, tools and applications. Brief Bioinform. 2023;24:bbad236.
151. Folstad M, Augustine A, Hollnagel F, Carroll C, Saner A, Pier J, Lavitschke M, Kleinschmidt P, Twedt H, Tevaarwerk AJ, Emamekhoo H. Improving quality of oncology (onc) documentation and enhancing structured data collection using a standardized onc note template. J Clin Oncol. 2024;42:11128–11128.
152. Le D, Brain A, Shenkier TN, Ingledew P-A. Virtual health in cancer care: results from a semi-structured interview-survey of oncology health care providers. J Clin Oncol. 2021;39:e13618–e13618.
153. Eldridge EH, Trowbridge E, Zandt MV, Adam A, Mack CD. Unstructured EMR data hold the key to oncology studies: methods for validating NLP-extracted phenotypes from EMR notes. J Clin Oncol. 2024;42:11133–11133.
154. Acosta JN, Falcone GJ, Rajpurkar P, Topol EJ. Multimodal biomedical AI. Nat Med. 2022;28:1773–84.

Publisher's Note