REVIEW ARTICLE   OPEN

Computational systems biology in disease modeling and control, review and perspectives

Rongting Yue [1] ✉ and Abhishek Dutta [1]

Omics-based approaches have become increasingly influential in identifying disease mechanisms and drug responses. Considering that diseases and drug responses are co-expressed and regulated in the relevant omics data interactions, the traditional way of grabbing omics data from single isolated layers cannot always obtain valuable inference. Also, drugs have adverse effects that may impair patients, and launching new medicines for diseases is costly. To resolve the above difficulties, systems biology is applied to predict potential molecular interactions by integrating omics data from genomic, proteomic, transcriptional, and metabolic layers. Combined with known drug reactions, the resulting models improve medicines' therapeutic performance by re-purposing the existing drugs and combining drug molecules without off-target effects. Based on the identified computational models, drug administration control laws are designed to balance toxicity and efficacy. This review introduces biomedical applications and analyses of interactions among gene, protein and drug molecules for modeling disease mechanisms and drug responses. The therapeutic performance can be improved by combining the predictive and computational models with drug administration designed by control laws. The challenges are also discussed for its clinical uses in this work.

INTRODUCTION

The high mortality of many diseases prohibits human longevity, and therapies need to be designed to suppress disease progression and aid organisms to recover from abnormal states[1]. However, the cost of launching new drugs is expensive and increasing, due to the long-term safety procedures in clinical trials[2] caused by drug overdose toxicity and off-target side effects[3,4]. The unknown drug targets in individuals may also cause problems. For example, the drug Torcetrapib has been designed for cardiovascular disease[5], but it may cause severe side effects of hypertension[6]. Analyses of omics, including genomics, proteomics, metabolomics and transcriptomics, contribute to the studies of disease mechanisms and drug responses. While a single omics layer focuses on a specific aspect with less complexity but limited information[7,8]. For instance, using only the conventional marker or the haplotype association cannot reveal the combined effects of Single Nucleotide Polymorphisms (SNPs), which potentially induces stroke[9]. The systemic view on dynamic gene regulation shows that genes work as part of complex networks instead of acting alone to perform cellular processes[10]. The integrative multi-layer omics data, such as transcriptional factors, genes and their expression products, provides a comprehensive map of metabolism and molecular regulation when analyzing and predicting based on complex cellular networks[11–15]. This leads to the prediction of potential molecular interactions through latent information of omics data. A general figure of omics data interactions is shown in Fig. 1.

In pharmacology, drug molecules act by binding to specific proteins, thereby changing their biochemical and biophysical activities[16]. Traditional treatment design based on physical parameters and external modalities[17,18] or simple ligand-protein interactions[4] are not sufficient for meeting clinical drug safety criteria or specifying variability among individuals. Modeling of the integrated clinical data and multi-layer molecular interactions makes the drug responses predictable[3,19].

With multi-layer omics data, a single disease can be studied across different clinical modalities simultaneously (i.e., the horizontal direction in Fig. 2), and different diseases can be explored from a single modality (i.e., the vertical direction in Fig. 2). Systems approach makes chemical molecules and biomolecules more likely to be linked to phenotypes for analyzing diseases and drugs and identifying their potential connections. Modeling of disease pathways and drug responses through different layers of regulation contributes to drug repurposing and drug combination based on known molecular interactions. This review classifies the models for interactions among gene, protein and drug molecules into two main classes: static network and dynamic modeling. Both frameworks integrate biological information. The modeling is served for studying disease mechanisms and drug responses. The two main tasks include (1) deriving potential molecular interactions from disease mechanism and drug response, and (2) designing drug dosages.

NETWORK STRUCTURE IN SYSTEMS BIOLOGY

A network structure visualizes a wide range of components such as genes or proteins and their interconnections. Network-based modeling can be established for systematic analysis based on omics data from various scales[20], which expands the use of bioinformatics beyond its original meaning by mining structural motifs for novel interaction prediction[21]. This agrees with the ideas that the networks with hierarchical bio-information consist of the metabolic, signal transduction, and gene regulation pathways all contribute to the analyses of interactions between protein inhibitors[22]. Diseases with overlapping network modules show significant co-expression patterns, symptom similarity and comorbidity[23], whereas diseases residing in separated network neighborhoods are phenotypically distinct[24].

A basic network is made up of nodes and edges. For molecular interactions, nodes can be genes[25], proteins[26], and drugs. Node

[1]Department of Electrical and Computer Engineering, University of Connecticut, 371 Fairfield Way, Storrs, CT 06269, USA. ✉email: rongting.yue@uconn.edu
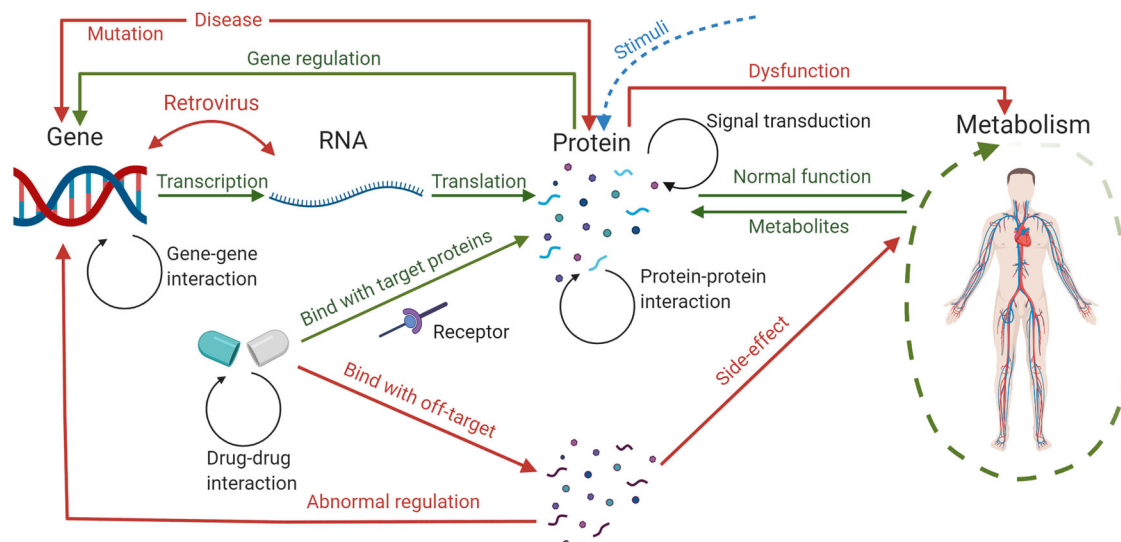
npj

Fig. 1   A systemic view of disease. Interactions among genomic, proteomic and transcriptomic levels reveal the regulatory process within organisms. Drug molecules intervene in this process by binding with specific target ligands. These omics data and chemical molecules are required to be analyzed simultaneously to study the entire disease mechanisms and drug reaction.
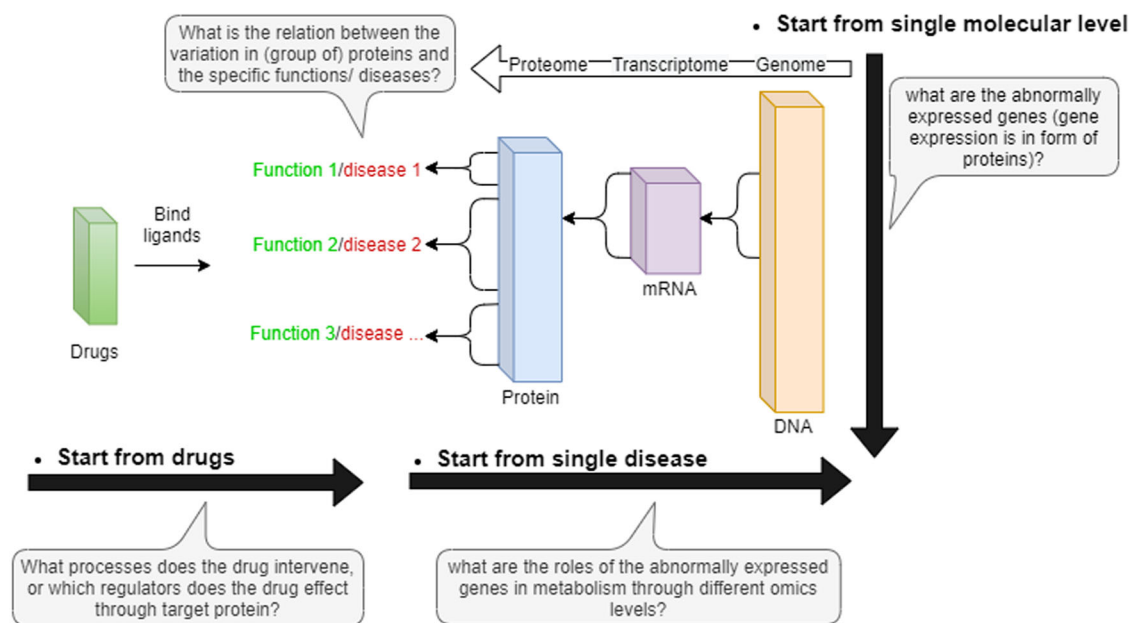

Fig. 2   Analyses of disease and drug effect through single and multiple layers of omics data. Systemic view enables scientists to establish disease models from a higher hierarchic level by multi-layer data integration. Through vertical analysis within single layers, the disease-related molecules are identified by abnormal gene expression values. Through horizontal analysis with different layers, the interactive information are used to track diseases or drug effects throughout the entire biological process.

annotations can be connective properties, including binding affinities[27], interactive directions[28], and the importance and confidence of the connections[29]. Edges link the nodes, and edge annotations can be functional interactions between nodes, including protein physical interactions, gene regulatory relations[30], mechanism of activation and inhibition[31,32], and disease associations[33]. Besides, network complexities reflect in its size. Large networks with high complexity can be iteratively divided into measurable subunits to reduce the complexity of analysis[34], and each subnetwork can be a set of functionally grouped nodes[28]. The patterns in the known annotations can be used to predict new annotation[35], and structural patterns can be obtained by network motif. Motif encodes regulatory behaviors and decreases internal cell noise[28], and it also helps

identify drug molecules with common reactions, discover unknown drug responses, and predict potential therapeutics[36]. The regulatory motifs in the gene regulatory network contribute to modeling the cell fate dynamics in the immune system[37]. The pathway motifs within gene regulatory networks help interpret genetic and epigenetic variation[38]. The topological motifs in the interaction network of drugs and targets help select target protein candidates for drug synergy[39].

STATIC NETWORK OF DISEASES AND DRUGS
A static network models the statically functional interactions from omics data. Network structure provides topological properties

from the presented interactions. It integrates intra- and extra-cellular information for identifying the modules' functional response by multiple network alignment. The overlapped multi-omics data integration is informative for reveal new molecular interactions[12,15].

The purpose of constructing a static network is to predict the potential interactions among drug molecules and target proteins through the shared components, as they can be the intermediaries to convey information to different network layers[4,40]. For example, the diseases can be associated based on the shared genetic associations, the gene-disease interactions, and the disease mechanism[23,26,30], such that disease connections can be built through the shared genes for drug repurposing. In a host-pathogen interaction network, the shared enzymes and regulatory components connect the metabolic reactions for predicting drugs for fungal infection[7]. Compared to a multiplex network, which only contains the same type of nodes and integrates the subnetworks from different layers, a heterogeneous network has the capability to include different types of nodes and edges. The multi-layer connections for the same nodes will result in a multiplex-heterogeneous network. More details about the different network structures can be found in reference[41]. Additionally, the new interacting pairs may account for variability in disease progression or drug response among individuals. To conclude, the shared components across layers may reveal new findings through multi-layer omics data modeled by a heterogeneous (or multiplex-heterogeneous) network structure.

The absent interactions don't guarantee the negative interacting relations, since the available binding profiles are limited. One obstacle preventing expanding the database is that the clinical experiments are costly. To avoid the expensive clinic experiments, based on the static network models, machine learning-based methods are used to predict possible interactions using known interaction data.

Interactions from omics data

Due to the expression relations, genes and proteins are always analyzed together for genetic analysis of complex diseases. For genome-wide association, proteins and gene interactions can identify densely connected modules in the human protein interactome[9]. The protein-protein interaction (PPI) networks encode the information of proteins (nodes) and their interactions (edges) into the network structure. PPI networks help predict the potential disease-related proteins, based on the assumption that shared components in disease-related PPI networks may cause similar disease phenotypes[7,33,42,43]. For example, PPIs can be used with gene co-expression networks to assess the host-pathogen response for clinical treatment of Covid infections[44]. To be specific, the HCoV-host interactome was used to predict SARS-CoV-2 pathogenesis and provide a theoretical host-pathogen interaction model for HCoV infections.

The aim of modeling static molecular interactions is to use the interaction profiles to find out the potential interacting pairs. The modeling starts from identifying disease-related regulators using omics data. Consider that the RNA-sequencing data on a disease-related microarray is available. The disease-related genes can be selected from the differentially expressed genes (DEGs) based on the moderated t-statistics analyses and empirical Bayes using Limma in R[45]. The genes with large variations in expression data can be chosen based on fold-change and p-value, and a PPI network can be mapped. Limma focuses on the statistical meaning of the gene expression level, and its performance is affected by the number of samples.

For gene co-expression analyses based on microarray data, Pearson Correlation Coefficient (PCC) is frequently used. For example, a gene co-expression network for the Z. mays and A. flavus genes can be mapped directly from pairwise PCC masked by a customized cutoff[46]. WGCNA[47] constructs an approximately scale-free network for detecting functional gene clusters based on PCC of gene co-expressions, under the assumption that proteins work together to perform metabolic functions. The disease-related hub genes/proteins with high connectivity are selected from the clusters. However, R^2 value and connectivity of the identified gene network are sensitive to gene quantity, and different parameter settings (i.e., soft threshold) will result in different co-expression modules. In the frequent gene co-expression network[48], gene pairs with high PCC, which are collected from different cancer and normal microarray dataset, are selected to build subnetworks of tightly co-expressed gene clusters using an iterative greedy algorithm "Quasi-Clique Merger". The edges in subnetworks are weighted by the frequency of these genes, and similar subnetworks are merged into larger networks that are identified for specific diseases. The researchers noted that compared to differential expression analysis, where normal samples are necessary for comparison, this approach integrates multiple microarray datasets that even make the use of data without normal samples, which makes the constructed network more informative. However, the size of the datasets has to be large enough to ensure a high level of significance for PCC. A decision tree-based method Randomforest GENIE3[49] can be used to infer gene co-expression network by solving p (i.e., the number of genes) regression subproblems of identifying gene expression patterns and then grouping the genes. It can fast detect gene networks from large gene datasets that have multifactorial expression data. However, this method assumes knowing the transcription factors in the gene dataset of the experimentally confirmed gene interactions. Note that PCC assumes the gene expressions are linearly correlated, which may not be true for biology systems[48]. In a gene co-expression network for identifying cross-species interactions, mutual information and Z-scores of gene pairs are calculated using Context Likelihood of Relatedness algorithm[50], which are used to infer edges in the network. As described by the authors, this algorithm can cope with nonlinear changes of gene expression, and it shows higher accuracy compared to PCC. However, PCC is still needed to discriminate the (positive or negative) directions of correlations of gene pairs. See Table 1 for comparison.

Eventually, the target proteins are obtained based on the gene clusters in the gene co-expression networks. Drugs that potentially intervene in disease progression can then be predicted based on these proteins. The question that remains is how to detect the disease mechanism relevant to small gene expression variations, since small changes in some genes may have more essential contributions to the overall process. Note that gene expression level study excludes the non-transcriptional interactions[48], which can be a potential limitation for predicting molecular interactions.

Drugs and targets interaction

Similarly, a static network can model the interactions among drugs and targets. Drug-target interaction (DTI) networks have been applied to study the prediction of drug response[51], interaction profiles of new drug-target pairs[16,26,27,52], and side effects of unknown drug combination[22,53,54]. A DTI-based target inhibition model has been proposed to identify the disease-specific target set with possible drug-target combinations by mapping drug inhibition profiles with the use of protein candidates, which were relevant to cancer survival with known drug binding profiles[52]. Notably, the drug target proteins in a DTI network may have larger degrees (i.e., more interactive molecules) than those proteins in a PPI network[16].

Nodes in DTI networks include drugs, drug targets, and off-target proteins. For predicting drug side effects, the off-targets

Table 1. Methods for constructing gene co-expression networks.

| Algorithms and applications	Advantages	Potential limitations
Quasi-Clique Merger algorithm for finding co-expressed gene clusters[48].	Integrates multiple microarray datasets, even including the data without normal samples[48].	Requires large datasets to ensure a high level of significance for correlations of gene expressions.
Context Likelihood of Relatedness algorithm for inferring edges in the network to identify cross-species gene interactions[50].	Captures nonlinear changes in gene expressions[50].	Can't discriminate the direction of correlations of gene pairs without Pearson Correlation Coefficient.
GENIE3 for inferring gene co-expression network[49].	Fast detects gene networks from large multifactorial gene expression data.	Requires prior knowledge of the transcription factors.
WGCNA for detecting functional gene clusters[47].	The approximately scale-free network structure reserves connectivity when randomly removing nodes.	Sensitive to the number of genes and the choices of parameters (i.e., soft threshold).
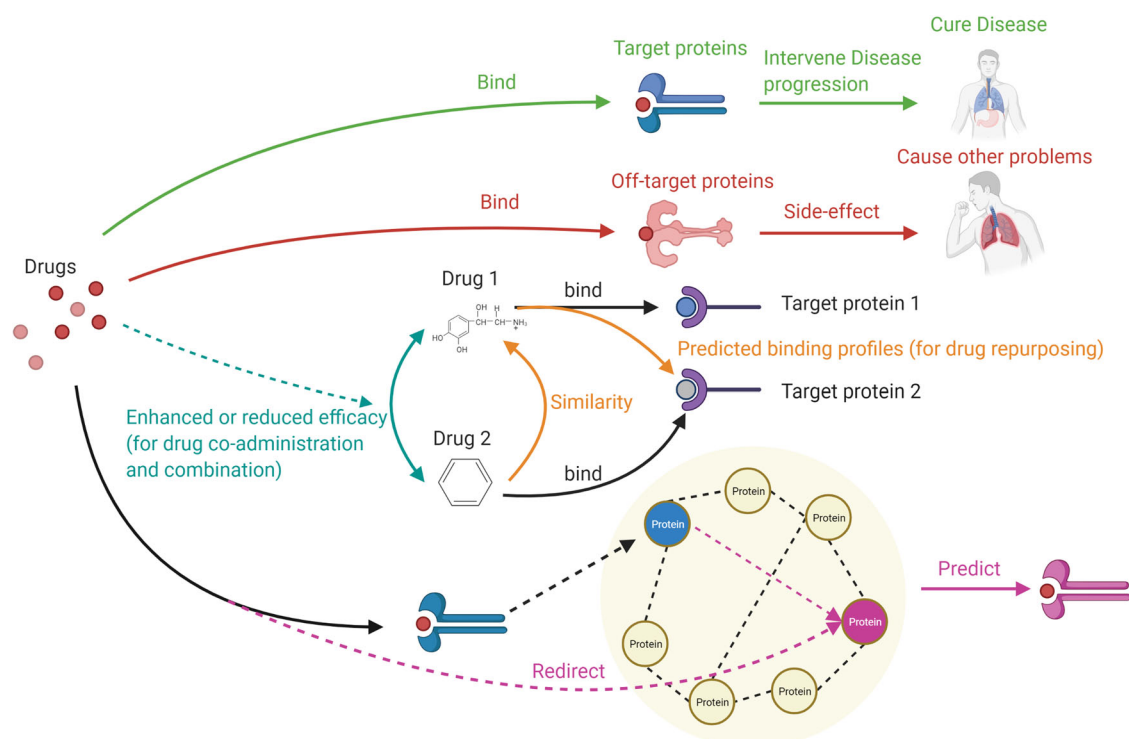
Fig. 3 Interactions among drugs and target proteins offer chances for drug combination, co-administration, and repurposing. Drug molecules may bind with off-target proteins that induce side effects (labeled by red arrows), which should be avoided. Though, the therapeutic drug molecules should be reserved, as they function as desired to cure diseases (labeled in green arrows). Interactions between drugs provide an opportunity for enhanced therapeutic performance through drug co-administration and combination (labeled by cyan arrow). Drug similarity conveyed by drug-target interactions provides a chance for drug repurposing (labeled by orange arrows). Drug targets may be expanded to similar proteins using protein-protein interactions for drug repurposing (labeled in purple arrows). Drug molecules in this figure include epinephrine (for "Drug 1") and benzene (for "Drug 2") as examples, using the icons from Biorender.

can be linked based on drug clinical relevance[54]. For drug combinations, different drugs may have interactions that induce side effects in patients or reduce the drug efficacy[55,56]. This requires us to identify potential drug-drug interactions (DDI), which can be explored by expanding DDIs through the shared targets of the drugs. The enriched DDIs identify potential targets and find new therapeutic uses (that the drugs do not initially aim at) and combination with surprising efficacy[4]. Figure 3 delineates a general view of the abridged drug and target interactions. Besides, the text mining-based predictions of molecular interactions rely heavily on the existing reports and references[57], and insufficient clues would result in inaccurate predictions. Some of the frequently used databases for gene and protein interaction analyses are listed in Table 2.

An example of constructing a static network

Here is an example of constructing a heterogeneous network that integrates interactomics data from different subnetworks. The task is to predict potential drugs for a disease using RNA-sequencing data. Data can be obtained from online databases such as Gene Expression Omnibus database[58]. DEGs can be identified through statistical analyses (e.g., empirical Bayes using "Limma" in R[45]) on these gene expression data. The resulting genes are mapped into the signaling pathways database (e.g., KEGG database[59]) such that the highly perturbed disease-related pathways are selected. Target proteins are selected from the pathways, since we aim at intervening in the disease's progress by blocking the relevant pathways using drugs. Interactions of target proteins are parsed from database (e.g., STRING database[60]) for constructing PPI networks and embedding the internal connecting information. DTI

Table 2. Databases for gene and protein interaction analyses.

Application	Database	Description	Reference
DTI, PPI	Comparative Toxicogenomics Database	Information of chemicals, pathways, disease, organisms, genes, drug-gene interactions. Data are mainly collected from references.	[39,53]
Gene regulation and interaction	GEO (Gene Expression Omnibus)	One of NCBI databases. Gene expression data (eg. RNA, genome methylation and proteins) that comes from data submissions such as microarray or other researches.	[13]
Genomics of cancer	Cancer Genome Atlas (TCGA)	Cancer molecular data including genome, epigenome, transcriptome and proteome.	[38]
Biological pathway	GeneGo MetaBase	Bioinformatics including signaling and metabolic pathways, interactions among drugs and proteins as well as kinetic information of drugs.	[54,65]
PPI, DTI, signaling pathway	KEGG (Kyoto Encyclopedia of Genes and Genomes)	Information of pathways, genome, chemicals and diseases based on diagrams of interaction and reaction. It is complementary to the majority of the existing molecular biology databases that contain information on individual molecules or individual genes.	[12,40,44,72]
PPI	STRING (Search Tool for the Retrieval of Interacting Genes)	Functional links in PPI based on experimental data. Interactions are predicted by comparative genomics and text mining based on the scoring system.	[9,44,85]
Gene regulation and interaction	CCLE (the Cancer Cell Line Encyclopedia)	Gene expression data for human cancer analysis, including information of mutation, Gene Methylation and the associations between cell line and genomics.	[51,52,170]
DTI, DDI	PubChem (NCBI)	Characteristics of chemical molecules and activities from experimental results or literature. For drug analysis, it provides information on the chemical structure for each drug and the validated chemical depiction information.	[51,85,95]
Gene regulation and interaction	GO (Gene Ontology)	Biological annotations including structure, function and dynamics in pathways, molecules and organism level for a variety of species.	[19,35,44]
DTI	STITCH (search tool for interactions of chemicals)	Profiles of chemicals and proteins interactions. The data source includes experimental results and text mining. More than 9 million proteins come from almost 2,000 organisms in this database.	[39]
DTI	ChEMBL	Biological activities and characteristics of molecules such as chemicals and proteins that contribute to the study of drug target and drug discovery.	[54,54,86]
Gene regualtion and transcription	UniGene (NCBI)	Gene sequences from animals and plants. The well-characterized sequences are driven from algorithm-based classification which helps to identify uniqueness among genes. The Source of intact gene sequences is GenBank.	[13,74]
DTI, DDI	Drug Bank	Drug-target and drug-drug interacting information such as chemical sequence, three-dimensional structure and pharmacological pathway involvement.	[16,40,71,76]
Genomics of breast cancer	METABRIC (Molecular Taxonomy of Breast Cancer International Consortium)	Clinical and expression data for breast tumors. The collected breast cancer specimens are grouped for discovery and validation. It helps to assess the survival prediction of cancer patients.	[173]
PPI	HPRD (Human Protein Reference Database)	Bioinformatics of human protein-protein interactions from literature and data are manually curated.	[33,42]
Signaling pathway	Reactome	Bioinformatics information including pathway, proteins and drugs for model visualization and analysis.	[9,29,44]
DDI	Online Mendelian Inheritance in Man (OMIM)	Disease data including disease loci, known disease genes and the known disorder-gene associations such as the molecular relationship between genetic variation and phenotypic expression.	[16,28]
PPI	Human Protein Reference Database (HPRD)	Protein information based on interactions described in published reports. The interaction set is expected to be biased toward known disease genes.	[15,33,42]

data is parsed from the drug database (e.g., DrugBank database[61]) by inputting the name of target proteins. Drugs that bind with at least one target protein are obtained to construct the DTI network. The DDI subnetwork is constructed by expanding the drugs from DTIs to their interactive partners, using interaction data from the drug database. And the internal connecting information among the expanded drug set is embedded into the subnetwork. The molecules from the three types of subnetworks are then integrated using a heterogeneous network, whose nodes are the drugs and target proteins, and edges are the interactions among proteins and targets. More recent works have focused on graph-based representation of molecular interaction network $G(V, E)$, where V is the set of molecules and E is the set of molecular interactions[62–64]. And information is propagated from nodes to nodes through graph edges. The overview of the delineated process is visualized in Fig. 4.
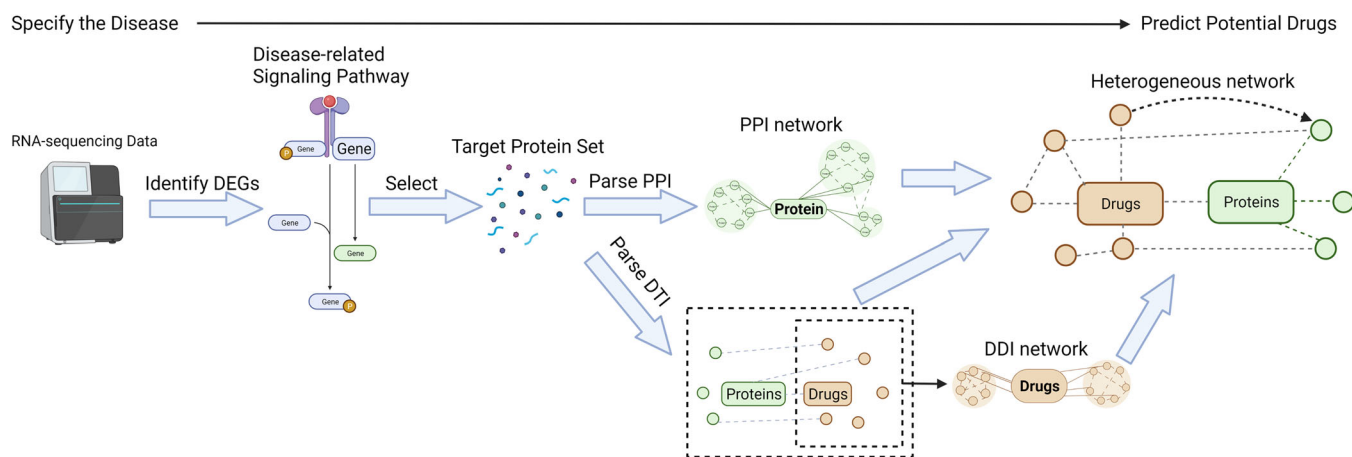
Fig. 4 Statistic analyses on gene expression Values from RNA-sequencing data identify DEGs. The signaling pathways that contain the highest ratio of DEGs are regarded as disease-related. All or part of proteins in these pathways are selected to form the target protein set. By parsing PPI and DTI data from databases, the PPI network and the DTI network are constructed. The drug molecules in the DTI network are also used to parse and construct the DDI network. Finally, the DDI and the PPI are connected by the DTI, and thus the heterogeneous network is constructed. The task of predicting the potential drugs for a given disease is now transferred to the prediction of the interactions between the proteins and the drug molecules. Abbreviation: DEGs Differentially Expressed Genes, PPI Protein-Protein Interaction, DDI Drug-Drug Interaction, DTI Drug-Target Interaction.

Limitation of static network modeling

The dynamic metabolic behaviors in patients result in changing expressions of genes. The interactomics in static modeling may become invalid, since molecule expression levels deviate a lot from the points that the model is built based on, thus leading to the failure of static modeling. This makes the dynamic modeling in Section V necessary. The major purpose of dynamic modeling is to map the regulatory relations among molecules, such that drugs can be used to intervene in disease progression by binding with target proteins, which drive the expression levels of genes to the normal range. Once gene expressions are corrected, the predictive static models will be valid again.

ANALYSIS OF STATIC MODELING

Although the models of diseases or drugs have been studied for decades, the actual biosystems are far more complicated than complete modeling[65], which indicates the potential of exploring more comprehensively models. This section reviews recent techniques that have been used to predict more information based on static models.

Importance quantification

The importance of interactions in the network requires measuring and ranking for reducing network complexity and generating the weights, such that the simplified networks include only the most disease-related molecules[65], which improves the efficacy of the learning process. The topological descriptors, such as degree, betweenness, and closeness, are frequently used to quantify the node importance in the network and embed spacial information into the modeling[16,66,67]. Node degree D, which is the number of connected edges[30], measures the connectivity of nodes. In molecular interaction networks, the highly connected nodes (i.e., the hub nodes) usually provide more biological insights, compare to nodes with low degrees[45,47]. However, when identify disease-specific regulators, the importance will be penalized on hubs, since they don't have much information[66]. The unconnected nodes will be discarded, since there is no path available to convey messages. Betweenness that describes the centrality of the given nodes is measured by the shortest path[68], and it has more change when removing intermodular hubs compared to the intramodular[19]. Importance quantification can also be done based on

statistics. For example, Z-score (i.e., the harmonic mean of precision and recall) measures the variability of the observations, and it has been used to quantify the importance of shortest paths between drug targets and the cardiovascular disease-related proteins[53]. The eigenvalues can be used to quantify the importance of data projection basis. In WGCNA[47], the significance of gene co-expression clusters is quantified using the "eigengenes" to reserve the most important genes and modules when modeling.

Similarity analysis

Similarity characterizes how elements are similar to each other in a static network. The assumption is that similar nodes have similar interaction profiles. For example, similar chemical structures of drugs show similar therapeutic effects for diseases[33,40,51]. Different types of similarities can be used based on the type of molecules. Protein similarity can be obtained based on protein sequence using the Smith-Waterman alignment algorithm[27,69]. Similarities between small molecules drugs are often calculated based on Jaccard coefficient of chemical structure notated by Simplified Molecular Input Line Entry System (SMILES)[70]. Similarity between cell lines can be obtained using similarity of gene expression profiles[51]. Other types of similarities can be the drug phenotypic side-effect similarity[71], pathological similarities[33], and so on.

Learning-based methods for clustering and classification

The evolution in learning-based networks has shown its ability to efficiently learn from massive datasets[4]. The learning-based methods can be supervised (with labels for training data), unsupervised (without labels) or semi-supervised (with partially labeled data).

A supervised learning approach enables the model to predict unknown parts (links) of the network based on the known interacting molecules[72–74]. Support Vector Machine (SVM), as a supervised learning model for classification, can refine topological information from network structure. SVM has been used for predicting DTIs based on DDIs, drug chemical structures, and side-effect information[75]. And the result shows that AUC values reach 0.76 in predicting the interaction between 261 drugs and 2,140 proteins. The multi-class SVM has been used to predict the therapeutic class of FDA-approved compounds using drug similarities, and it shows 78% classification accuracy of level 2
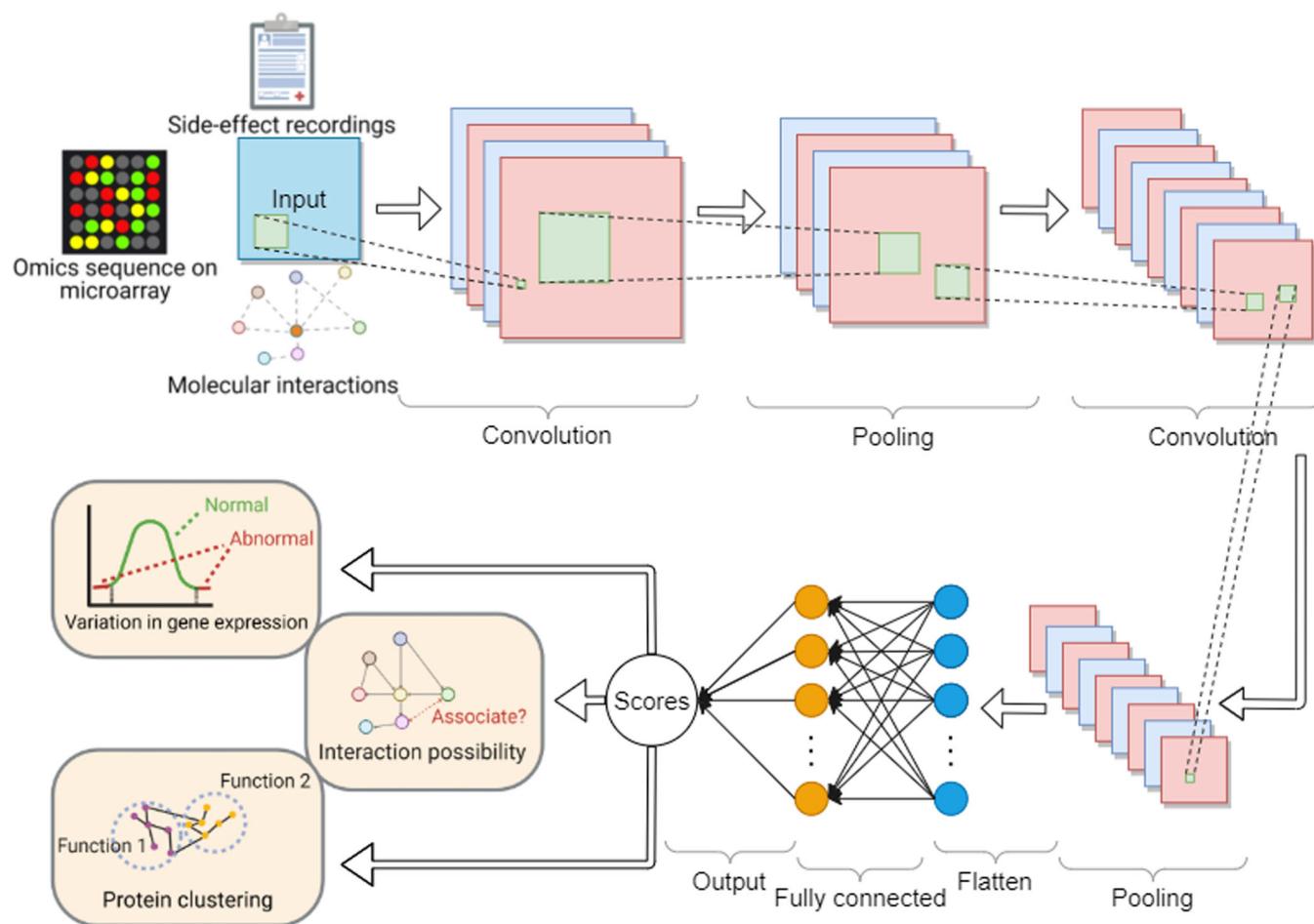
Fig. 5 A convolutional neural network. This neural network is comprised of an input layer, convolutional layers (that extract features from hyperplanes of input data by projection/convolution), pooling layers (that reduce the spatial size and mitigate the locational sensitivity), flatten layer (that flattens the features and feeds them into the artificial neural network), fully connected layer (that learns nonlinear function of the extracted features) and an output layer. The input is the raw features such as molecules sequence patterns, gene regulation annotations and patterns, molecular interaction network motifs, molecular structures and structural associations, drug chemical structures, drug side-effect reports, and so on. The output can be the predicted classification (e.g., molecular binding profiles) and regression (e.g., quantified molecular binding affinities) when obtaining new samples.

ATC codes among 410 drugs[76]. Kernels in SVM measure the features between gene pairs to train the classifier[77], and the classifier makes the binary prediction for the interactions between the existing molecules and the incoming components[72]. Random Forest (RF), as an ensemble learning algorithm, can be used to performance classification based on decision trees[78], such as the prediction of the contact probability between protein coevolved residues[79]. RF is robust to noise and it is capable to handle small sample size[80]. However, RF is less interpretable, and the computational load of RF will increase exponentially as the size of data increases[81]. The convolutional neural network (CNN), which uses the convolutional kernels and refines molecular features from arbitrary network frames of different sizes and shapes[82], has been applied to refine DDI features using text mining on the biomedical information[56]. In the study, the prediction performance reached a F-score of 70% when evaluating about 900 drug documents. While a single-layer neural network may not have good predicting performance, a deep neural network can be deployed and obtain better results. Deep learning (DL) methods, which are built by multiple layers of neural networks, are gaining more and more attention because of the structural flexibility[83] and their capability of extracting molecular patterns by mining latent information from the network structures[82,84–86]. For example, a deep CNN learning architecture[87]

shows its high concordance index (large value is better) in predicting drug-target binding affinity. While using more layers of neural networks results in more parametric settings, which could be potentially time-consuming. One efficient method to search these network parameters can be Particle swarm optimization[88]. A scheme of a convolutional neural network is shown in Fig. 5.

However, supervised learning heavily relies on the size of labeled samples during the training. Consider n_d drugs and n_t targets that include $n_d \cdot n_t$ drug-target pairs, then the number of available DTIs $n_a << n_d \cdot n_t$. Moreover, when predicting DDIs base on drug side effect, the positive samples (i.e., drug pairs with known interactions) can be obtained from database, but the negative samples (i.e., drug pairs with with clinically validated safe co-prescriptions) are almost unavailable[89]. The lack of labeled data will decrease the predicting performance when using supervised learning methods. Though data preprocessing partly copes with the missing data problem, it causes certain information loss[90].

Unsupervised learning aims at clustering data based on features, without the use of data labels. Examples are given as follows. When identifying relevance of disease phenotype and treatment response between patients, the number of patient clusters is unknown, and a hierarchical clustering (HC) algorithm with multiple linkage methods has been used for clustering patients based on genome-wise similarity and variability[91]. The

Table 3. Learning-based Methods for Predicting Molecular Interactions.

Type	Advantages	Disadvantages	Applications
Supervised learning	Use full label information of omics data.	Rely heavily on size of labeled data. Data preprocessing for noise and features may be needed, but this causes information loss.	Logistic regression for genome-wise prediction on relevant functions, disease and trait[104]; Multiple kernel learning for predicting drug response of cancer cell lines using omics profiles and pathways[90]; Support vector machine for drug-target interaction[75]; Convolutional neural network for identifying drug-drug interaction from document[56]; Random Forests for predicting protein contact[79].
Unsupervised learning	No need for data labels. Suitable for the case where the labeled data is few and expensive to obrain.	Lose the informative features brought by labels.	Autoencoder for denoising a single-cell RNA-sequencing model[93] and extracting representative features from drug molecular structure and protein sequences[95]; Hierarchical clustering algorithm for clustering patients based on genome-wise similarity and variability[91].
Semi-supervised learning	Combine the benefits of feature extraction brought by unsupervised learning, and also make full use of the informative label data.	Algorithms work under proper assumptions. The trained model will loss generalization on testing data if assumptions don't hold.	Autoencoder-based semi-supervised learning for predicting DTI[99], DDI[100] and PPI[101]; Label propagation for predicting DDIs[89].

benchmark test on 191 Multiple Sclerosis patient samples reaches Rand index moer than 0.85, and it also shows the capability of reducing feature dimension of Single Nucleotide Polymorphisms of 191 patients from more than 25,000 to about 1500. Advantages of HC includes that it is visualizable and user can customize the granularity by cutting the clustering at the desired level[92]. As a widely used unsupervised learning method, the autoencoder (AE) learns and refines the low-dimensional features from data to reconstruct the input, so that a minimalistic description can be derived for differentiability among samples[84]. This encoding-decoding frame can be used to denoise a single-cell RNA-sequencing model[93], transform molecules directly into a numerical representation[86], and compress computational dimensionality[4] when detecting the cell type-specific clusters in an ensemble clustering way[94]. The stacked AE has been used to extract highly representative features from drug molecular structure and protein sequences[95], which help identify the potential DTIs[95]. Instead of capturing the deterministic latent features, the Variational Autoencoder (VAE) aims at capturing the distribution of the latent variables, with the help of variational inference[96,97]. Though VAE is capable of generating artificial samples for sparsely labeled molecules data, its heavy computational load due to the optimization of hyperparameters may limit the usage[92]. Note that AE and VAE are not classifiers.

Transductive learning learns to predict labels of unlabeled data in the existing network by training on entire dataset, while inductive learning learns a classifier that make prediction on testing data out of the current network. We focus on transductive learning, because the static network structure is deterministic for predicting potential interactions among existing nodes.

Semi-supervised learning (SSL) combines supervised and unsupervised learning. SSL is widely used for predicting molecular interactions using a small number of known interactions and many unknown interactions. Label propagation is a classical transductive SSL method that utilizes a small labeled dataset and predicts the label of the unlabeled data iteratively[98]. This method has been used to predict DDIs based on side effect similarity among 569 drugs with 52,416 DDI pairs[89]. The result showed that when the size of the training dataset is small, instead of the training data, the output of the proposed predictive model mainly relies on the geometric structure of the entire dataset. Also, the Area Under the Precision-Recall Curve value ranged from 0.650 to 0.729 for different ratios of testing and training dataset, which indicates the predictive model is stable. The autoencoder-based semi-supervised learning has been applied for predicting DTI[99], DDI[100] and PPI[101]. The unsupervised AEs/VAEs with a supervised deep neural network form a semi-supervised neural network, which shows higher AUC (over 0.8) compared to other learning methods[99–101]. Though, SSL cannot work without proper assumptions, as it will loss generalization from a finite training set to other test cases[102,103]. Besides, SSL is not always superior to supervised learning. An example is that a supervised logistic regression outperformed a semi-supervised label propagation when evaluating the prediction accuracy of gene functions, relevant disease and gene traits of on the full network connectivity, since the former algorithm efficiently extracts local network patterns, while the latter focuses on network topology[104]. The comparison of learning-based methods for predicting molecular interactions can be found in Table 3.

Graph-based learning for predicting molecular interactions

Graph modeling is gaining increasing attention, since it encodes the structural and spatial information of data into models. The basic graph components include the sets of nodes, edges and an adjacency matrix that stores node connection information. A heterogeneous network of molecular interactions can be easily represented by a graph model. Graph Embedding (GE) encodes the spatial and topological information of graphs into low-dimensional feature vectors using a parametric function[105,106]. GE in transductive learning is deterministic because of a fixed graph. While for inductive learning, GE is generated from graph input features[105].

Several graph learning methods have been proposed. Similar to CNN, graph convolution extracts features from graphs. Graph Convolutional Network (GCN), as a spatial convolution approach and the first-order approximation of Chebyshev polynomials of the graph spectral filter for semi-supervised classification tasks, aggregates the weighted node features from neighborhoods to the current node being visited using locally convolutional computation[107]. However, GCN has a relatively shallow structure. In a deep GCN, the oversmoothing property gives similar embedding to nodes from different labels/classes, resulting in mislabeling/misclassification issues[108]. The spectral graph convolution approaches partition/diffuse graph in Fourier domain

Table 4. Graph learning methods.

Methods	Advantages	Limitations
Graph Covolutional Network[107]	Aggregate graph information and make the use of structural information of graphs.	Less computationally efficient for large graphs. Lack of generalization to graphs with different structure[110]. Oversmoothing of graph embeddings[108].
Graph Attention Network[110]	Computationally efficient for node-level parallel processing. Don't need the knowledge of the entire graph structure.	Can't tell the differences between local and global structures well[111].
(Variational) Graph Autoencoder[112]	Reduce data dimensionality and speed up the training process.	Captures more information from dataset, rather than the relevant information to the problem. And the reconstruction process loses information.
Graph Generative Adversarial Nets[113]	Can augment dataset and impute missing values.	Instability of gradient updates, and the vanished gradients of generator[114], etc.

using Eigen-decomposition of the graph Laplacian[109], which could be computationally expensive when dealing with large graphs, and the graph convolution relies on specific graph structures that are not generalizable to other graphs with different structures[110]. Graph Attention network (GAT) utilizes the self-attention mechanism to obtain the normalized attention score (i.e., the relative importance) of each node from its neighborhood[110], instead of the average aggregation functions in GCN. The attention mechanism solves the oversmoothing problem[108]. Node level attention is computed in parallel which is time-efficient. And it alleviates the effect brought by the lack of knowledge of the entire graph structure. However, GAT can't tell the differences between local and global structures well because the aggregators lack of cardinality preservation mechanism[111]. Graph Autoencoder (GAE) and Variational GAE (VGAE) finds the (distribution of) latent variables of embeddings (which can be encoded by GCN), using a (variational) inference model and a Generative Model (GM), in low-dimensional space to recover the adjacency matrix of a graph in an unsupervised manner[112]. Loss functions in GAE and VGAE can be the reconstruction loss of the adjacency matrix (i.e., graph connectivity) and the variational lower bound, respectively[112]. The reduced dimensionality of data speeds up the training/back-propagation processes. However, the autoencoder-based methods focus on capturing most of the information from dataset based on a lossy reconstruction, which may not be relevant to the problem, and a relatively large dataset is needed to train an autoencoder. The inadequate bioinformatic data can be augmented by learning strategies. Graph Generative Adversarial Nets (GraphGAN)[113] composed of a GM and a Discriminative Model (DM) can be used for this task. GM in both of GraphGAN and VGAE captures the distribution of the graph connectivity. However, rather than predicting the interacting pairs of nodes in each training epoch in VGAE, the GM in GraphGAN generates fake samples to deceive the discriminiator, such that DM learns to discriminate the true samples (that come from training data) from the generated samples. A Nash equilibrium that balances between two models is desired for convergence. Some limitations of GAN, including the unstable gradient updates and the vanished gradients of generator[114], may potentially cause problems when being applied to graphs. See Table 4 for comparison of these graph learning methods.

Additionally, the negative interacting pairs are rarely available in bioinformatics data, which results in an unbalanced dataset. The negative sampling[115] can be applied that randomly assigns negative labels to the unlabeled data during the training process. This step aims at training the classifier to tell the difference between positive samples and the pseudo-negative samples drawn from training data. Lastly, a powerful package "PyTorch Geometric"[116] (python), which contains multiple recent graph learning algorithms and uses fast tensor operation in GPU, can be used for implementation.

Compare to a binary classification problem (e.g., predict if the interaction exists or not), the multi-class classification problem is more attractive since more than two classes are involved. Examples of multi-class prediction can be found in predicting the interaction type of protein pairs[117], the gene phenotype[118], etc. The loss functions are designed based on class or labels. Each sample in a multi-class problem is assigned to one of the classes, and the labels for multi-label problems are usually obtained using categorical one-hot encoding[119], which encodes the label as a vector with binary values. The entropy-based loss will be different for multi-class and multi-label classification. The Softmax function is often used to output probabilities of classes for the multi-class problems[120], and Sigmoid function is used for multi-label problems.

The integrative static modeling conveys information through networks, which potentially results in a more comprehensive disease model by prediction. Similarities of nodes and edges among the nodes are calculated, and the quantified information is embedded in nodes and edges. Then the learnable models find representative features to predict potential molecular interactions. The benefits of static modeling-based prediction for disease treatment include: (1) the predicted disease-related regulators can be the new drug targets, which potentially improves treatment efficacy by finding new drugs for the same diseases; (2) the known drug target proteins redirect the drug molecules to bind with similar proteins through PPIs for drug re-purposing; (3) proteins/genes identified from the predictive model may explain variability among patients for specific disease phenotype or specific drug responses, which contributes to precision medicine; (4) the predicted off-target activities help avoid side effects; (5) the potential drug interactions contribute to designing novel therapies by drug combination, since the combinations may either reduce or enhance drug therapeutic efficacy. Though hypotheses and predictions can be made in a static network, the clinical use of potential drugs, drug combinations, or drug repurposing is highly concerned with patients' safety, which means the accuracy and generalization of current predictive algorithms could be problematic. Besides, the commonly acknowledged problem in analyzing interactions between biological and chemical molecules based on database is data scarcity[20], and how to use limited data to reliably produce more data for learning and obtain highly accurate predictive models remains an open problem.

DYNAMIC MODELING OF REGULATION IN ORGANISMS

The term dynamic disease refers to that disease pathogenesis is mainly caused by the appearance of new dynamic behaviors of organism, independently of the underlying pathogenesis[121]. The dynamics of diseases and drug reactions are informative for treatment design. For example, the viral reservoir dynamics are significant to understand the natural properties of ongoing viral
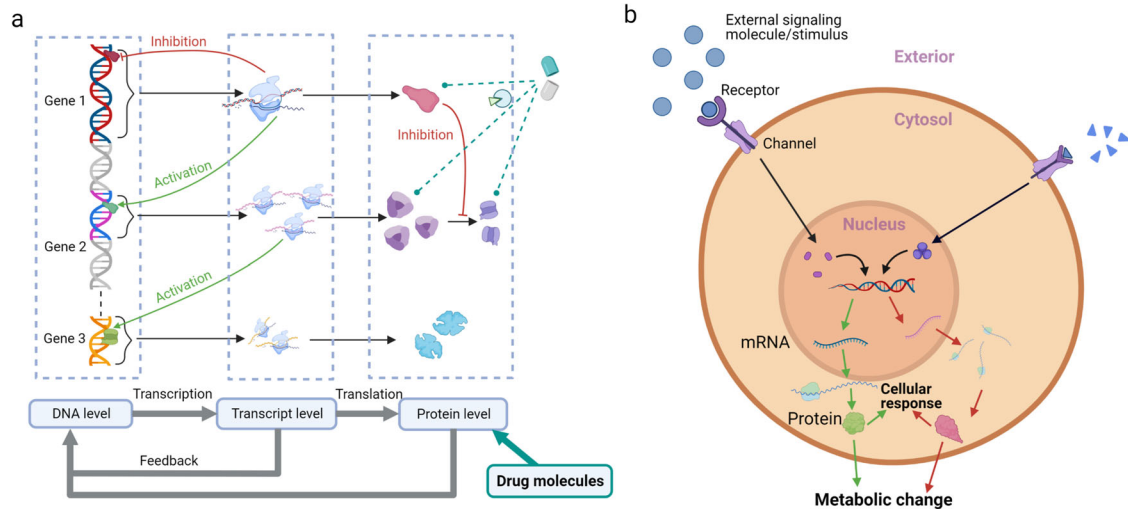
Fig. 6 Schematic diagrams in dynamic modeling. a Gene regulatory network adapted from the work[25]. Gene 1, 2, and 3 are coding genes. Gene 1 regulates its own expression and those of Gene 2. The protein produced by Gene 1 regulates Gene 3 expression through a signaling factor/protein (that is produced from the protein expressed by Gene 2.) Drugs can intervene in the regulation by binding with proteins that change the gene expressions. **b** Diagram of signaling transduction. Signals are received and enter the nucleus to change gene expression. Proteins are synthesized to regulate phenotypic behaviors of cells or tissue. Errors (e.g., dysregulation) in signaling pathways (e.g., dysregulation) may cause the ceasing of cell apoptotic that results in unlimited growth and division.

progression with treatment[122], which help researchers find the cures for diseases. A kinetic model predicts disease outcome by modeling the time-course behaviors of patients' interactome[19,25], and model performance can be assessed by its ability to generate testable predictions[123].

Gene regulatory network

In organisms, the changing gene expression results in phenotypic changes. Genes interact with each other through RNA and protein expression products, thereby governing the rates at which genes in the network are transcribed into messenger RNA[121]. The causal information in gene expression data, e.g., key drivers of complex traits in phenotype-related gene interaction network[13], can be identified by variations in DNA and gene expression-related traits[25,37]. Gene regulatory networks (GRN) are thus used to detect potential mechanisms based on dynamic behaviors of epigenomic activity between normal and disease state[10]. For applications, GRNs have been used to analyze infectious diseases by detecting gene regulations related to the infectious and viral mechanism[15]. A combination of gene regulatory components including myeloid and lymphoid has been studied to identify cell fate specification[11]. Analyzing GRN is regarded as the reverse engineering for investigating gene regulatory relations by going backward with the observed gene expression[28]. The causal regulatory relations of genes are desired to be found from genome-wide expression data. A diagram of GRNs is shown in Fig. 6a. A GRN model contains genes and regulators as the nodes, and directional edges as the regulatory relations between the nodes[31]. To model and reconstruct GRN, the time-course microarray data of gene expression product (e.g., gene expression values, chromatin expression profiles) are required[25,28,124], and the quantitative regulation can be obtained from computational models of GRN.

Signaling pathway and transcription

Signal transduction pathway contains regulators between molecules in an organism. It attributes to changes in both gene expression and gene connectivity[29]. Errors in signal transduction lead to altered development and incorrect behavioral decisions in organisms, whose dysfunction may result in uncontrolled cell growth or tumorigenesis[28,125,126]. At the protein level, signaling

pathways are comprised of protein interactions covering the biological functions in living cells, which captures the inter- and intracellular regulatory mechanisms of gene transcription and protein synthesis[28,31]. The modeling of dynamic signaling pathways measures disease progression. For example, pathway analyses contribute essentially to the systematic profiling of the transcriptome in heart failures[29]. The fungal signaling pathways model the regulatory behaviors in fungal pathogen infections[7]. Notably, the pathway-wide association can even extract valuable information from background noise and the context-specific logic of GRN[38].

Transcription Factors (TFs) are the keys in modeling gene regulatory relation. For example, TF regulates the development of innate and adaptive cells of the immune system[11]. The dynamic transcriptional and translational subnetworks have been used to model the trigger mechanism of the innate response regulated by intercellular and intracellular heterogeneity[15]. A diagram of signal transduction is shown in Fig. 6b.

ANALYSIS OF DYNAMIC MODELING
Mathematical modeling for dynamic regulation

The dynamics of omics data reflects organism's response to the changes of interior milieu or environmental factor. Variability in metabolisms and phenotypes can be huge even if most of the corresponding genes are the same[127]. Dynamic modeling exerts mathematical tools that quantify the rate of state change in gene regulation in different conditions and time sequences[25]. More logic models and kinetic models using Hill function or piecewise linear differential equations for quantifying the dynamic behaviors of gene network have been reviewed in the references[128]. This section reviews the computational modeling of diseases using Differential Equations (DEs), which can be used for drug administration by control theory.

DEs capture how the system reacts to the variations caused by disease or drugs. From a systemic view, DEs integrate information from multi-layer omics data into a unified form[34]. DEs model the dynamic behaviors of organisms, such as transcription[31], gene regulation[25], metabolite concentrations[129], reaction rate[34], and factors in signal transduction pathway[127]. DEs have also been used to quantify signal flow in pathways and explores the effect of
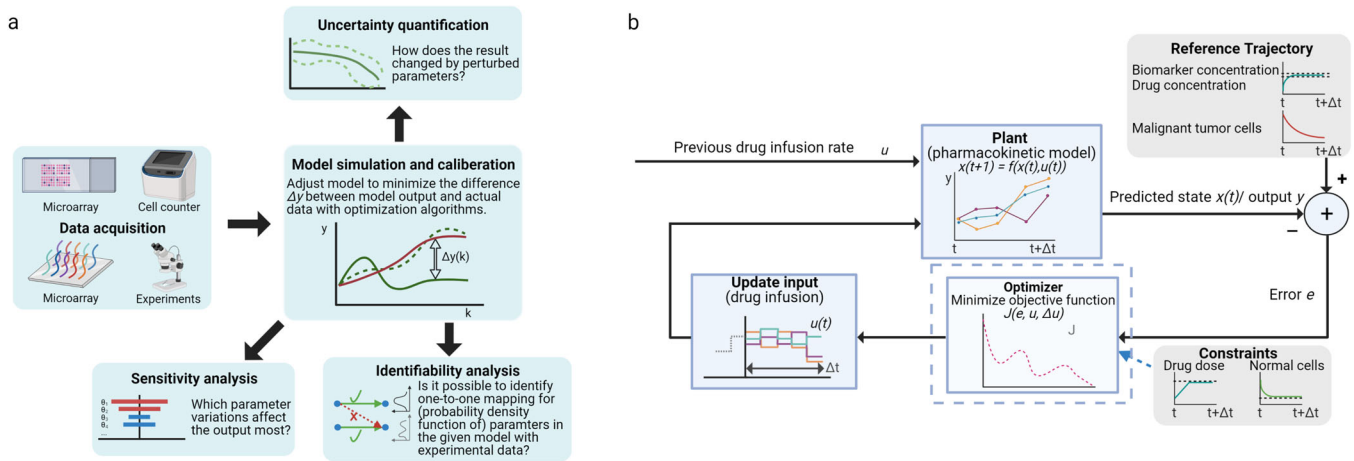
Fig. 7 Dynamic modeling and analyses. a Scheme of parameter estimation. After data acquisition, parameters are fitted in models by minimizing the difference between experimental data and model output. Sensitivity analysis, uncertainty quantification and identifiability analysis help assess the performance and robustness of the fit. **b** Loop of Model Predictive Control. Based on the output of model prediction, this control strategy updates control input (e.g., drug administration) to make the system dynamics track reference trajectory (e.g., desired tumor cell decrement) during each time interval. Its essence is to handle the constrained optimization problem (Constraints can be maximal drug doses and minimal normal cell populations).

oncogenic mutations on dynamics of ligands[123]. For disease modeling and treatment, DEs have been used to capture the dynamics of HIV viral infection[130], drug response of an irradiation-induced cellular senescence[131], and cancer cell population[132]. The components in DEs can be the metabolite concentration in nonlinear biochemical models[129], the signal transduction molecules in dynamic cell compartment models[34], and transcription factors and regulatory site[31]. A two-step approach, including the estimation of cell expression velocity using finite difference, and the estimation of a numerical square matrix that depicts the gene regulatory influence using sparse regression, can be used to computationally model a GRN from time-course single-cell RNA-sequencing data[133]. The accuracy of the approximated model depends on the user-defined sampling rate in finite difference. A robust dynamic model design makes it less sensitive to biological noise and disturbances, allowing us to track abnormal variations[21,134]. An example of the robust design of a physical system can be found in [134]. Modeling noise is usually assumed to be Gaussian. Non-Gaussian stochastic noise can be modeled by solving the Fokker-Planck equation[135].

Parameter estimation in dynamic models

The main task of parameter estimation is to find the best-fit parameters to characterize physical processes and reproduce experiments[136–138]. To develop data-driven models, one can optimize the fit of a collection of parameters iteratively to a given dataset with random disturbance[137,139], which explores the parameter space extensively and limits the number of non-convergent solutions[131]. The least-square estimation (LSE) that minimizes quadratic errors between the predicted model values and experimental data[129,140,141]), and the maximum likelihood estimation[142], are widely used for this task. A Kalman Filter (KF) can recursively generate the maximum likelihood estimates for a linear dynamic system from a series of noisy measurements[143,144]. It handles the approximate modeling of high-dimensional noisy data with small sample sizes[145]. The Extended KF deals with nonlinear models in biology[146]. Note that the Gaussian noise is assumed in KF, and Particle Filters (PFs) are more appropriate when dealing with non-Gaussian processes. By randomly drawing samples from numerical simulation, the Monte Carlo method estimates parameter values[122] and quantifies their uncertainties[147]. PFs in sequential Monte Carlo method obtain weighted samples from the non-Gaussian posterior probability of the state

in nonlinear systems[148]. Study shows that a PF with an orthogonal basis (used for approximating the posterior by an orthogonal series expansion) outperformed Extended KF when estimating parameters of a Wiener anesthesia delivery model[146]. The process of parameter estimation is shown in Fig. 7a.

Sensitivity analyses assess how sensitive the models' outputs are to the fitted models' parameters changes. Sensitivity can then quantify model uncertainty through finite differencing or variational equations[139]. Identifiability checks the reliability of the estimates and assesses how well the model explains experimental data[144,149]. The global optima for parameter estimation are always desired, leading the multiple approaches to converge to the same solution ultimately[129]. Time of searching parameters is upper bounded to prevent an endless search[139]. Calling a single data cluster recursively may cause the absence of global parameters[34] and get stuck in the misleading local optimum that should be avoided[139] due to the generality. Precision medicine requires the patient-specific parameters for individual variability, resulting in a more complicated model. Note that a robust model design[134] is desired after adding new parameters.

Drug administration with control theory

Drug dosage is designed for patients' recovery by control algorithms. Though the dynamics of drug concentration in plasma and drug response are usually nonlinear in real world, a linearized model is usually used to approximate the nonlinear behaviors of the original system due to the reduced complexity[150–153]. Several methods can be applied for the linear approximation. For example, when the pharmacokinetic data is available, drug model can be fitted using regression methods[4]. Linear models can be approximated from a nonlinear model by truncating Taylor expansion at the first-order term. Or, the linear descriptors can be obtained using Koopman operator, which projects the infinite-dimensional time-evolving observables (e.g., time-series data of drug concentration and cell population) into finite-dimensional states, using dynamic mode decomposition[154]. The Koopman method is driven by data, which can be generated from the nonlinear model[155]. Then the linear controllers for drug dosages can be designed based on a linear model.

The Proportional, Integral and Derivative (PID) control is a classic control algorithm that takes the error $e(t)$ between closed-loop feedback signal $y(t)$ and the setpoint $r(t)$ (i.e., $e(t) = r(t) - y(t)$) as the controller input to calculate input needed, such that the

model states can be driven to the setpoints[156]. PID controller has been used to design the drug administration for the chemotherapy treatment in a cancer model[18] and the anesthesia in the neuromuscular blockade models[146,151]. The clinical evaluation and simulation results show that drug concentration levels can reach and be maintained at certain levels in acceptable time horizon. A PID controller copes with the uncertainty in the system's dynamics caused by interpatient variability[157] and time variations[151]. The concise scheme of PID[158] makes it flexible for functional expansion, such as an I-PD controller constructed by cascading[18]. However, the performance and the robustness of PID controllers depend heavily on tuning[146,159], and one of the frequently used tuning approaches is the Ziegler-Nichols method[160].

Besides maintaining the drug concentration at certain levels, drug dosage should also balance the therapeutic performance and the toxicity/side effects. Optimal control law minimizes control errors (e.g., drive the systems along trajectories) and control efforts (e.g., less energy consumption)[161] subject to system dynamics. The linear quadratic regulator (LQR), as one of the optimal control strategies, can be used for deriving an optimal control sequence (e.g., a sequence of drug infusion rates) when dealing with a linear model in the form $\dot{x}(t) = A(t)x(t) + B(t)u(t)$, with quadratic objective functions in form of $l2$-norm $|| \cdot ||_2$) from time t_0 to t_f with the initial state x_0, as shown in Eqn. (1)[162].

$$J = \frac{1}{2}||x(t_f)||^2_{S_f} + \frac{1}{2}\int_{t_0}^{t_f}(||x(t)||^2_{Q(t)} + ||u(t)||^2_{R(t)}), \text{subject to } x(t_0) = x_0 \quad (1)$$

where S_f is the terminal weighting matrix. $Q(t)$ and $R(t)$ are user-defined positive semi-definite and positive definite weighting matrix, respectively. Large $Q(t)$ results in aggressive drug doses, and large $R(t)$ leads to medical conservatism. The optimal feedback control law $u(t) = -K(t)x(t)$ is derived by solving the algebraic Riccati equation and Hamiltonian. Compared to PID control, optimal control balances the drug toxicity and dosage. For chemotherapy treatment, the control problem can be to minimize the kinetic energies of all the cancerous cells with low drug toxicity[18,163–165] by searching the optimal sequence of drug dosages. In a HIV model, the goal can be to maximize the benefit based on levels of healthy CD4+ T cells and immune response cells by reducing the systemic cost of anti-HIV drugs[130], and the cost function consists of beneficiary T Cell population and systemic costs of therapy[166]. For multi-drug treatment, the ratio of co-administration between different drugs is also considered[164,167]. More examples of optimal control in cancer treatment can be found in the book[168]. The control law calculated by minimizing the cost can be piecewise constants or linear in the finite time horizon[169].

Drug concentration in patients is delicate. The dangerous therapeutic window of drug concentration in plasma determines whether the level is tolerable and the drug is effective for patients[17,170]. Thus, state and input constraints are necessary to a drug delivery system. Usually, constraints include the upper boundary of toxicity level and the certain therapeutic window for drug concentration and disease progression[18,167]. The constrained optimization problems require more complicated control algorithms. In a multi-objective genetic algorithm, the Pareto optimal sets have been used to search for lower values of the objectives simultaneously[18]. Nonlinear optimization can be solved by Bock's direct multiple shooting method with a numerical solution on a fixed control discretization grid[167]. The steepest descent method can be used to search numerical solutions iteratively to minimize or maximize the cost function[164], and the numerical solutions can be derived with Miser3/Matlab[132]. Model Predictive Control (MPC) handles the constraints in optimization problems and the mismatch between nominal and actual processes[169], compared to optimal control. MPC solves a finite horizon open-loop optimal control problem to obtain control actions with predicted states from models. The local asymptotic stability of the control law is

guaranteed when time horizon is sufficiently long[169]. The loop of model predictive control is shown in Fig. 7b. Besides, controllers can also "learn" from complex situations through iterative learning-based strategies to obtain optimal parameterized control signals[171]. The equilibrium in a chemotherapy model refer to the elimination or the stop of cancer cell proliferation, so that no more treatment will be needed[164]. While using MPC, the feasibility and stability should be carefully considered. Compared to optimal control, MPC handles the constrained optimization problems, which makes it more suitable for the drug administration design subject to patients' physical constraints.

The advanced control has good performance on drug dose adjustment. When eliminating cancer cell population, it has been found that giving bursts of high-dose abiraterone reduces tumor burden more than 10 times, compared to giving a constant dose[163]. For robustness, when model parameter error is 25%, MPC reaches 98% success rate (among 100 simulations) of stabilizing HIV infection in 2 years[165]. When designing the dosage of remdesivir for SARS-CoV-2, an optimal control sequence has been obtained by solving a constrained optimization problem, and simulation shows that the proposed control scheme reduces the treatment horizon from 10 days to 5 days and it also reduces more than 50% of the drug dose, compared to the recommended treatment regimes from FDA and WHO[172]. See Table 5 for comparing these methods used for drug dosage design. By using the advanced control algorithms, fewer drugs can be used to obtain the same or better treatment efficacy with less toxicity. There are some other control strategies that use feedback control laws based on nonlinear models. For example, a positive semi-definite Lyapunov function, whose gradient requires to be negative semi-definite, can be designed to calculate the controlled vaccination rate that asymptotically stabilizes Covid infection[138].

Optimal drug dosage designed by control algorithms is used to efficiently intervene in disease progression, based on dynamic models with parameters estimated from sample data. Drug administration is transformed into an optimization problem with or without model constraints. The design of personalized treatments requires the patient-specific parameters for individual variability, which makes control scheme more complicated. Notably, the impulsivity of the model should be considered during the design process, and discrete drug therapy should be considered to allow the normal cells to rebuild in clinical treatment[166].

CONCLUSION AND OUTLOOK

Launching new drugs could be costly compared to using existing drugs for new therapeutic performance, and drug predictions based on computational systems biology have shown its potential in precision medicine, drug combinations, and repurposing. A static network composed of molecular interactions is able to predict the potential interaction pairs based on known omics data by conveying information through nodes and edges. The new participants in the map of molecular interactions help obtain a more comprehensive view of disease progression and drug response. This results in using drug molecules with better therapeutic performance while avoiding off-target effects. Also, the potential patient-specific regulators can be identified to explain individual variability for personalized treatment. However, the predictive models may fail due to in vivo changing behaviors, which makes dynamic modeling necessary. Dynamic modeling aims at building math models to predict disease progression and drug response. Model parameters are estimated from clinic data. The potential participants identified from static modeling can be the new elements in dynamic models. By applying optimal drug dosages designed by control algorithms, disease progression can be intervened efficiently, which also indicates the predictive

Table 5. Control algorithms for drug administration.

Control Objective	Algorithm	Advantages	Disadvantages	Applications
Maintain drug concentration at certain levels	PID control	Easy to implement; flexible structure for functional expansion.	Design of PID controllers usually needs linear models. The linear approximation of the nonlinear model leads to information loss that eventually decreases control performance.	For anesthesia delivery[146]. For designing anti-retroviral therapy for HIV[165].
Balance between drug toxicity, drug cost and therapeutic performance	Optimal control	Balance drug cost, toxicity and therapeutic performance.	More computations are required. Weighting variables Q and R are sensitive and they should be adjusted carefully.	For optimal imatinib treatment for leukemia model[174]. For eliminating cancer cells using abiraterone[163].
Handle constraints in the model; Balance between drug toxicity, drug cost and therapeutic performance	Model predictive control	Consider physical constraints based on drug toxicity and patient conditions.	Solving the constrained optimization problem needs additional work to ensure stability, optimality and feasibility.	For stabilizing HIV infection[165]. For anesthesia delivery with uncertainty[175].

model in static modeling will be valid again. The combination of static and dynamic modeling makes it a powerful tool for disease analysis and therapy design. SSL outperforms other learning methods for making static modeling predictive, while the underlying assumptions may not hold in real cases, which makes the model loss generality. The learning algorithms that have an accurate prediction of other testing data are always desired. For drug administration, the simple control algorithms cannot meet the complicated design objectives (e.g., control with constraints), while a complicated control algorithm may not be time-efficient, though it aims at more control objectives. A control algorithm that handles multi-objectives and computes drug dosage needed efficiently is desired.

The modeling of DTIs by expanding drugs and targets based on DDIs and PPIs offer opportunities for (1) finding new targets for the same drug, (2) exploring new drugs for the same disease, and (3) minimizing off-target side effects for safe therapies. Though side effects should be avoided for safety issues, it does not mean all drugs that cause side effects should be abandoned. Treatments like chemotherapy harm normal cells, so we shall choose the drug agents that specifically target on cancer cells, rather than healthy cells. Combined with drug administration for changing conditions of patients using control theory, the treatment with better therapeutic performance and lower impairment can be realized simultaneously.

REFERENCES

1. Yang, K., Bai, H., Ouyang, Q., Lai, L. & Tang, C. Finding multiple target optimal intervention in disease-related molecular network. *Mol. Syst. Biol.* **4**, 228 (2008).
2. Dickson, M. & Gagnon, J. P. Key factors in the rising cost of new drug discovery and development. *Nat. Rev. Drug Discov.* **3**, 417–429 (2004).
3. Liebler, D. C. & Guengerich, F. P. Elucidating mechanisms of drug-induced toxicity. *Nat. Rev. Drug Discov.* **4**, 410–420 (2005).
4. Lo, Y. C., Rensi, S. E., Torng, W. & Altman, R. B. Machine learning in chemoinformatics and drug discovery. *Drug Discov. Today* **23**, 1538–1546 (2018).
5. Tanne, J. H. Pfizer stops clinical trials of heart drug. *BMJ* **333**, 1237 (2006).
6. Forrest, M. J. Torcetrapib-induced blood pressure elevation is independent of CETP inhibition and is accompanied by increased circulating levels of aldosterone. *Br. J. Pharmacol.* **154**, 1465–1473 (2008).
7. Horn, F. et al. Systems biology of fungal infection. *Front. Microbiol.* **3**, 108 (2012).
8. Hasin, Y., Seldin, M. & Lusis, A. Multi-omics approaches to disease. *Genome Biol.* **18**, 1–15 (2017).
9. Arning, A. et al. A genome-wide association study identifies a gene network of ADAMTS genes in the predisposition to pediatric stroke. *Blood* **120**, 5231–5236 (2012).
10. Grechkin, M., Logsdon, B. A., Gentles, A. J. & Lee, S. I. Identifying network perturbation in cancer. *PLoS Comput. Biol.* **12**, e1004888 (2016).
11. Laslo, P., Pongubala, J. M., Lancki, D. W. & Singh, H. Gene regulatory networks directing myeloid and lymphoid cell fates within the immune system. *Semin. Immunol.* **20**, 228–235 (2008).
12. Wang, B. et al. Integrative omics approach to identifying genes associated with atrial fibrillation. *Circ. Res.* **126**, 350–360 (2020).
13. Schadt, E. E. et al. An integrative genomics approach to infer causal associations between gene expression and disease. *Nat. Genet.* **37**, 710–717 (2005).
14. Kitano, H. Computational systems biology. *Nature* **420**, 206–210 (2002).
15. Subramanian, N., Torabi-Parizi, P., Gottschalk, R. A., Germain, R. N. & Dutta, B. Network representations of immune system complexity. *Wiley Interdiscip. Rev.: Syst. Biol. Med.* **7**, 13–38 (2015).
16. AY, M., Goh, K. I., Cusick, M. E., Barabasi, A. L. & Vidal, M. Drug–target network. *Nat. Biotechnol.* **25**, 1119–1127 (2007).
17. Mage, P. et al. Closed-loop control of circulating drug levels in live animals. *Nat. Biomed. Eng.* **1**, 1–10 (2017).
18. Algoul, S., Alam, M. S., Hossain, M. A. & Majumder, M. Multi-objective optimal chemotherapy control model for cancer treatment. *Med. Biol. Eng. Comput.* **49**, 51–65 (2011).
19. Taylor, I. W. et al. Dynamic modularity in protein interaction networks predicts breast cancer outcome. *Nature biotechnology* **27**, 199–204 (2009).
20. Durmuş, S., Çakir, T., Özgür, A. & Guthke, R. A review on computational systems biology of pathogen–host interactions. *Front. Microbiol.* **6**, 235 (2015).
21. Albert, R. Network inference, analysis, and modeling in systems biology. *Plant Cell* **19**, 3327–3338 (2007).
22. Xie, L., Li, J., Xie, L. & Bourne, P. E. Drug discovery using chemical systems biology: identification of the protein-ligand binding network to explain the side effects of CETP inhibitors. *PLoS Comput. Biol.* **5**, e1000387 (2009).
23. Langhauser, F. et al. A diseasome cluster-based drug repurposing of soluble guanylate cyclase activators from smooth muscle relaxation to direct neuroprotection. *NPJ Syst. Biol. Appl.* **4**, 1–13 (2018).
24. Menche, J. et al. Uncovering disease-disease relationships through the incomplete interactome. *Science* **347**, 1257601 (2015)
25. Hecker, M., Lambeck, S., Toepfer, S., Van Someren, E. & Guthke, R. Gene regulatory network inference: data integration in dynamic models-a review. *Biosystems* **96**, 86–103 (2009).
26. Sharan, R. et al. Conserved patterns of protein interaction in multiple species. *Proc. Natl Acad. Sci.* **102**, 1974–1979 (2005).
27. Keiser, M. J. et al. Predicting new molecular targets for known drugs. *Nature* **462**, 175–181 (2009).
28. Wang, E., Lenferink, A. & O'Connor-McCourt, M. Genetic studies of diseases. *Cell. Mol. Life Sci.* **64**, 1752–1762 (2007).
29. Ma, X. Revealing pathway dynamics in heart diseases by analyzing multiple differential networks. *PLoS Comput. Biol.* **11**, e1004332 (2015).
30. Goh, K. I. et al. The human disease network. *Proc. Natl Acad. Sci.* **104**, 8685–8690 (2007).
31. Husmeier, D. Sensitivity and specificity of inferring genetic regulatory interactions from microarray experiments with dynamic Bayesian networks. *Bioinformatics* **19**, 2271–2282 (2003).
32. Meyer-Hermann, M., Figge, M. T. & Straub, R. H. Mathematical modeling of the circadian rhythm of key neuroendocrine–immune system players in rheumatoid arthritis: a systems biology approach. *Arthritis Rheumatism* **60**, 2585–2594 (2009).

33. Suthram, S. et al. Network-based elucidation of human disease similarities reveals common functional modules enriched for pluripotent drug targets. *PLoS Comput. Biol.* **6**, e1000662 (2010).

34. Bentele, M. et al. Mathematical modeling reveals threshold mechanism in CD95-induced apoptosis. *J. Cell Biol.* **166**, 839–851 (2004).

35. Chicco, D., Sadowski, P. & Baldi, P. Deep autoencoder neural networks for gene ontology annotation predictions. In *Proc. 5th ACM Conference on Bioinformatics, Computational Biology, and Health Informatics* 533–540 (Association for Computing Machinery, 2014). https://dl.acm.org/doi/proceedings/10.1145/2649387.

36. Lamb, J. et al. The Connectivity Map: using gene-expression signatures to connect small molecules, genes, and disease. *Science* **313**, 1929–1935 (2006).

37. Singh, H., Khan, A. A. & Dinner, A. R. Gene regulatory networks in the immune system. *Trends Immunol.* **35**, 211–218 (2014).

38. Califano, A., Butte, A. J., Friend, S., Ideker, T. & Schadt, E. Leveraging models of cell regulation and GWAS data in integrative network-based association studies. *Nat. Genet.* **44**, 841–847 (2012).

39. Vitali, F. et al. A network-based data integration approach to support drug repurposing and multi-target therapies in triple negative breast cancer. *PloS ONE* **11**, e0162407 (2016).

40. Fakhraei, S., Huang, B., Raschid, L. & Getoor, L. Network-based drug-target interaction prediction with probabilistic soft logic. *IEEE/ACM Trans. Computat. Biol. Bioinform.* **11**, 775–787 (2014).

41. Valdeolivas, A. et al. Random walk with restart on multiplex and heterogeneous biological networks. *Bioinformatics* **35**, 497–505 (2019).

42. Oti, M., Snel, B., Huynen, M. A. & Brunner, H. G. Predicting disease genes using protein–protein interactions. *J. Med. Genet.* **43**, 691–698 (2006).

43. Cui, T., Zhang, L., Wang, X. & He, Z. G. Uncovering new signaling proteins and potential drug targets through the interactome analysis of Mycobacterium tuberculosis. *BMC Genomics* **10**, 118 (2009).

44. Messina, F. et al. COVID-19: viral–host interactome analyzed by network based-approach model to study pathogenesis of SARS-CoV-2 infection. *J.Transl. Med.* **18**, 1–10 (2020).

45. Smyth, G. K. *Bioinformatics and Computational Biology Solutions Using R and Bioconductor* 397–420 (Springer, 2005).

46. Musungu, B. M. et al. A network approach of gene co-expression in the Zea mays/Aspergillus flavus pathosystem to map host/pathogen interaction pathways. *Front. Genet.* **7**, 206 (2016).

47. Langfelder, P. & Horvath, S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinform.* **9**, 1–13 (2008).

48. Zhang, J. et al. Weighted frequent gene co-expression network mining to identify genes involved in genome stability. *PLoS Comput. Biol.* **8**, e1002656 (2012)

49. Huynh-Thu, V. A., Irrthum, A., Wehenkel, L. & Geurts, P. Inferring regulatory networks from expression data using tree-based methods. *PLoS ONE* **5**, e12776 (2010).

50. McClure, R. S. et al. Species-specific transcriptomic network inference of interspecies interactions. *ISME J.* **12**, 2011–2023 (2018).

51. Zhang, N. et al. Predicting anticancer drug responses using a dual-layer integrated cell line-drug network model. *PLoS Comput. Biol.* **11**, e1004498 (2015).

52. Tang, J. et al. Target inhibition networks: predicting selective combinations of druggable targets to block cancer survival pathways. *PLoS Comput. Biol.* **9**, e1003226 (2013).

53. Cheng, F. et al. Network-based approach to prediction and population-based validation of in silico drug repurposing. *Nat. Commun.* **9**, 1–12 (2018).

54. Lounkine, E. et al. Large-scale prediction and testing of drug activity on side-effect targets. *Nature* **486**, 361–367 (2012).

55. Tatonetti, N. P., Fernald, G. H. & Altman, R. B. A novel signal detection algorithm for identifying hidden drug-drug interactions in adverse event reports. *J. Am. Med. Inform. Assoc.* **19**, 79–85 (2012).

56. Liu, S., Tang, B., Chen, Q. & Wang, X. Drug-drug interaction extraction via convolutional neural networks. *Comput. Math. Methods Med.* **2016**, 6918381 (2016).

57. Wu, Z. et al. SDTNBI: an integrated network and chemoinformatics tool for systematic prediction of drug–target interactions and drug repositioning. *Briefings Bioinform.* **18**, 333–347 (2017).

58. Clough, E. & Barrett, T. *Statistical genomics* 93–110 (Springer, 2016).

59. Kanehisa, M., Sato, Y. & Kawashima, M. KEGG mapping tools for uncovering hidden features in biological data *Protein Sci.* **31**, 47–53 (2021)

60. Szklarczyk, D. et al. The STRING database in 2017: quality-controlled protein–protein association networks, made broadly accessible *Nucleic Acids Res.* **45**, D362–D368 (2016)

61. Wishart, D. S. et al. DrugBank 5.0: a major update to the DrugBank database for 2018. *Nucleic Acids Res.* **46**, D1074–D1082 (2018).

62. Luo, J., Ding, P., Liang, C. & Chen, X. Semi-supervised prediction of human miRNA-disease association based on graph regularization framework in heterogeneous networks. *Neurocomputing* **294**, 29–38 (2018).

63. Lee, B., Zhang, S., Poleksic, A. & Xie, L. Heterogeneous multi-layered network model for omics data integration and analysis. *Front. Genet.* **10**, 1381 (2020).

64. Zhang, F., Wang, M., Xi, J., Yang, J. & Li, A. A novel heterogeneous network-based method for drug response prediction in cancer cell lines. *Sci. Rep.* **8**, 1–9 (2018).

65. Morris, M. K., Saez-Rodriguez, J., Clarke, D. C., Sorger, P. K. & Lauffenburger, D. A. Training signaling pathway maps to biochemical data with constrained fuzzy logic: quantitative analysis of liver cell responses to inflammatory stimuli. *PLoS Comput. Biol.* **7**, e1001099 (2011).

66. Sadegh, S. et al. Exploring the SARS-CoV-2 virus-host-drug interactome for drug repurposing. *Nat. Commun.* **11**, 1–9 (2020).

67. Ashtiani, M. et al. A systematic survey of centrality measures for protein-protein interaction networks. *BMC Syst. Biol.* **12**, 1–17 (2018).

68. Joy, M. P., Brock, A., Ingber, D. E. & Huang, S. High-betweenness proteins in the yeast protein interaction network. *J. Biomed. Biotechnol.* **2005**, 96 (2005).

69. Smith, T. F. & Waterman, M. S. Identification of common molecular subsequences. *J. Mol. Biol.* **147**, 195–197 (1981).

70. Weininger, D. SMILES, a chemical language and information system. 1. Introduction to methodology and encoding rules. *J. Chem. Inform. Comput. Sci.* **28**, 31–36 (1988).

71. Campillos, M., Kuhn, M., Gavin, A. C., Jensen, L. J. & Bork, P. Drug target identification using side-effect similarity. *Science* **321**, 263–266 (2008).

72. Bleakley, K., Biau, G. & Vert, J. P. Supervised reconstruction of biological networks with local models. *Bioinformatics* **23**, i57–i65 (2007).

73. Ding, H., Takigawa, I., Mamitsuka, H. & Zhu, S. Similarity-based machine learning methods for predicting drug–target interactions: a brief review. *Briefings Bioinform.* **15**, 734–747 (2014).

74. Nakaya, H. I. et al. Systems biology of vaccination for seasonal influenza in humans. *Nat. Immunol.* **12**, 786 (2011).

75. Kim, S., Jin, D. & Lee, H. Predicting drug-target interactions using drug-drug interactions. *PLoS ONE* **8**, e80129 (2013).

76. Napolitano, F. et al. Drug repositioning: a machine-learning approach through data integration. *J. Cheminform.* **5**, 30 (2013).

77. Mordelet, F. & Vert, J. P. SIRENE: supervised inference of regulatory networks. *Bioinformatics* **24**, i76–i82 (2008).

78. Breiman, L., Friedman, J. H., Olshen, R. A. & Stone, C. J. *Classification and Regression Trees* (Routledge, 2017).

79. Ma, J., Wang, S., Wang, Z. & Xu, J. Protein contact prediction by integrating joint evolutionary coupling analysis and supervised learning. *Bioinformatics* **31**, 3506–3513 (2015).

80. Qi, Y. *Ensemble Machine Learning* 307–323 (Springer, 2012).

81. Hengl, T., Nussbaum, M., Wright, M. N., Heuvelink, G. B. & Gräler, B. Random forest as a generic framework for predictive modeling of spatial and spatio-temporal variables. *PeerJ* **6**, e5518 (2018).

82. Duvenaud, D. K. et al. Convolutional networks on graphs for learning molecular fingerprints. *Adv. Neural Inform. Process. Syst.* **2015-January**, 2224–2232 (2015).

83. Chen, H., Engkvist, O., Wang, Y., Olivecrona, M. & Blaschke, T. The rise of deep learning in drug discovery. *Drug Discov. Today* **23**, 1241–1250 (2018).

84. Gawehn, E., Hiss, J. A., Brown, J. B. & Schneider, G. Advancing drug discovery via GPU-based deep learning. *Expert Opin. Drug Discov.* **13**, 579–582 (2018).

85. Kadurin, A., Nikolenko, S., Khrabrov, K., Aliper, A. & Zhavoronkov, A. druGAN: an advanced generative adversarial autoencoder model for de novo generation of new molecules with desired molecular properties in silico. *Mol. Pharm.* **14**, 3098–3104 (2017).

86. Blaschke, T., Olivecrona, M., Engkvist, O., Bajorath, J. & Chen, H. Application of generative autoencoder in de novo molecular design. *Mol. Informatics* **37**, 1700123 (2018).

87. Öztürk, H., Özgür, A. & Ozkirimli, E. DeepDTA: deep drug–target binding affinity prediction. *Bioinformatics* **34**, i821–i829 (2018).

88. Qolomany, B., Maabreh, M., Al-Fuqaha, A., Gupta, A. & Benhaddou, D. Parameters optimization of deep learning models using particle swarm optimization. In *13th International Wireless Communications and Mobile Computing Conference (IWCMC)* (eds Gerla, G. & Mauri, J. L) 1285–1290 (IEEE, 2017). https://ieeexplore.ieee.org/xpl/conhome/7975134/proceeding.

89. Zhang, P., Wang, F., Hu, J. & Sorrentino, R. Label propagation prediction of drug-drug interactions based on clinical side effects. *Sci. Rep.* **5**, 1–10 (2015).

90. Ali, M. & Aittokallio, T. Machine learning and feature selection for drug response prediction in precision oncology applications. *Biophys. Rev.* **11**, 31–39 (2019).

91. Lopez, C., Tucker, S., Salameh, T. & Tucker, C. An unsupervised machine learning method for discovering patient clusters based on genetic signatures. *J. Biomed. Inform.* **85**, 30–39 (2018).

92. Karim, M. R. et al. Deep learning-based clustering approaches for bioinformatics. *Briefings Bioinform.* **22**, 393–415 (2021).

93. Eraslan, G., Simon, L. M., Mircea, M., Mueller, N. S. & Theis, F. J. Single-cell RNA-seq denoising using a deep count autoencoder. *Nat. Commun.* **10**, 1–14 (2019).

94. Geddes, T. A. et al. Autoencoder-based cluster ensembles for single-cell RNA-seq data analysis. *BMC Bioinform.* **20**, 660 (2019).

95. Wang, L. et al. A computational-based method for predicting drug–target interactions by using stacked autoencoder deep neural network. *J. Comput. Biol.* **25**, 361–373 (2018).

96. Ding, Y., Tian, L. P., Lei, X., Liao, B. & Wu, F. X. Variational graph auto-encoders for miRNA-disease association prediction. *Methods* **192**, 25–34 (2021).

97. Cao, S., Lu, W. & Xu, Q. Deep neural networks for learning graph representations. In: *Proc. AAAI Conference on Artificial Intelligence* 30 (eds. Schuurmans, D & Wellman, M) (AAAI Press, 2016). https://ojs.aaai.org/index.php/AAAI/issue/view/303.

98. Raghavan, U. N., Albert, R. & Kumara, S. Near linear time algorithm to detect community structures in large-scale networks. *Physical review E* **76**, 036106 (2007).

99. Bahi, M. & Batouche, M. Drug-target interaction prediction in drug repositioning based on deep semi-supervised learning. In *IFIP International Conference on Computational Intelligence and Its Applications* (eds Amine, A, Mouhoub, M, Mohamed, O. A. & Djebbar, B) 302–313 (Springer, 2018).

100. Liu, N., Chen, C. B. & Kumara, S. Semi-supervised learning algorithm for identifying high-priority drug–drug interactions through adverse event reports. *IEEE J. Biomed. Health Informatics* **24**, 57–68 (2019).

101. Zhang, Y. & Lu, Z. Exploring semi-supervised variational autoencoders for biomedical relation extraction. *Methods* **166**, 112–119 (2019).

102. Chapelle, O., Scholkopf, B. & Zien, A. Semi-supervised learning (chapelle, o. et al., eds.; 2006)[book reviews]. *IEEE Trans. Neural Networks* **20**, 542–542 (2009).

103. Ouali, Y., Hudelot, C. & Tami, M. An overview of deep semi-supervised learning Preprint at https://arxiv.org/abs/2006.05278 (2020).

104. Liu, R., Mancuso, C. A., Yannakopoulos, A., Johnson, K. A. & Krishnan, A. Supervised learning is an accurate method for network-based gene classification. *Bioinformatics* **36**, 3457–3465 (2020).

105. Yang, Z., Cohen, W. & Salakhudinov, R. Revisiting semi-supervised learning with graph embeddings. In *International Conference on Machine Learning* (ed. Lawrence, N) 40–48 (JMLR, Inc. and Microtome Publishing, 2016).

106. Hamilton, W. L., Ying, R. & Leskovec, J. Representation learning on graphs: methods and applications. *IEEE Data Engineering Bulletin* **40**, 52–74 (2017).

107. Kipf, T. N. & Welling, M., Semi-supervised classification with graph convolutional networks. *5th International Conference on Learning Representations* (2017).

108. Min, Y., Wenkel, F. & Wolf, G. Scattering gcn: overcoming oversmoothness in graph convolutional networks. *Adv. Neural Inform. Process. Syst.* **33**, 14498–14508 (2020).

109. Bruna, J., Zaremba, W., Szlam, A. & LeCun, Y. Spectral networks and locally connected networks on graphs. Preprint at https://arxiv.org/abs/1312.6203 (2013).

110. Velickovic, P. et al. Graph attention networks. *6th International Conference on Learning Representations* (2018).

111. Zhang, S., Xie, L. Improving attention mechanism in graph neural networks via cardinality preservation. In *IJCAI: Proceedings of the Conference 2020* (ed. Bessiere, C) 1395 (International Joint Conferences on Artificial Intelligence, 2020).

112. Kipf, T.N. & Welling, M. Variational graph auto-encoders. *Bayesian Deep Learning Workshop (NIPS 2016)* (2016).

113. Wang, H. et al. Learning graph representation with generative adversarial nets. *IEEE Trans. Knowledge Data Eng.* **33**, 3090–3103 (2019).

114. Arjovsky, M. & Bottou, L. Towards principled methods for training generative adversarial networks. *5th International Conference on Learning Representations* (2017).

115. Mikolov, T., Sutskever, I., Chen, K., Corrado, G. S. & Dean, J. Distributed representations of words and phrases and their compositionality. *Adv. Neural Inform. Process. Syst.* **26**, (2013).

116. Fey, M. & Lenssen, J. E. Fast graph representation learning with PyTorch Geometric. *7th International Conference on Learning Representations* (2019).

117. Chen, M. et al. Multifaceted protein–protein interaction prediction based on Siamese residual RCNN. *Bioinformatics* **35**, i305–i314 (2019).

118. Chen, L. et al. Predicting gene phenotype by multi-label multi-class model based on essential functional features. *Mol. Genet. Genomics* **296**, 905–918 (2021).

119. Asteriou, D. & Hall, S. G. *Applied Econometrics* (Macmillan International Higher Education, 2015).

120. Grandini, M., Bagli, E. & Visani, G. Metrics for multi-class classification: an overview Preprint at https://arxiv.org/abs/2008.05756 (2020).

121. Villoslada, P., Steinman, L. & Baranzini, S. E. Systems biology and its application to the understanding of neurological diseases. *Annals Neurol.* **65**, 124–139 (2009).

122. Luo, R., Piovoso, M. J., Martinez-Picado, J. & Zurakowski, R. HIV model parameter estimates from interruption trial data including drug efficacy and reservoir dynamics. *PLoS ONE* **7**, e40198 (2012).

123. Chen, W. W. et al. Input–output behavior of ErbB signaling pathways as revealed by a mass action model trained against dynamic data. *Mol. Syst. Biol.* **5**, 239 (2009).

124. Ramirez, R. N. et al. Dynamic gene regulatory networks of human myeloid differentiation. *Cell Syst.* **4**, 416–429 (2017).

125. Kolch, W., Halasz, M., Granovskaya, M. & Kholodenko, B. N. The dynamic control of signal transduction networks in cancer cells. *Nat. Rev. Cancer* **15**, 515–527 (2015).

126. Sever, R. & Brugge, J. S. Signal transduction in cancer. *Cold Spring Harbor Perspect. Med.* **5**, a006098 (2015).

127. Cho, K. H. & Wolkenhauer, O. Analysis and modelling of signal transduction pathways in systems biology. *Biochem. Soc. Trans.* **31**, 1503–1509 (2003).

128. Le Novere, N. Quantitative and logic modelling of molecular and gene networks. *Nat. Rev. Genet.* **16**, 146–158 (2015).

129. Moles, C. G., Mendes, P. & Banga, J. R. Parameter estimation in biochemical pathways: a comparison of global optimization methods. *Genome Res.* **13**, 2467–2474 (2003).

130. Culshaw, R. V., Ruan, S. & Spiteri, R. J. Optimal HIV treatment by maximising immune response. *J. Math. Biol.* **48**, 545–562 (2004).

131. Dalle Pezze, P. et al. Dynamic modelling of pathways to cellular senescence reveals strategies for targeted interventions. *PLoS Comput. Biol.* **10**, e1003728 (2014).

132. Pillis, L. G. et al. Chemotherapy for tumors: an analysis of the dynamics and a study of quadratic and linear optimal controls. *Math. Biosci.* **209**, 292–315 (2007).

133. Aubin-Frankowski, P. C. & Vert, J. P. Gene regulation inference from single-cell RNA-seq data with linear differential equations and velocity inference. *Bioinformatics* **36**, 4774–4780 (2020).

134. Dutta, A. Robust design of a multirotor aerial vehicle. *Sci. Rep.* **11**, 1–13 (2021).

135. Chen, X., Wu, F., Duan, J., Kurths, J. & Li, X. Most probable dynamics of a genetic regulatory network under stable Lévy noise. *Appl. Math. Comput.* **348**, 425–436 (2019).

136. Zi, Z. & Klipp, E. SBML-PET: a Systems Biology Markup Language-based parameter estimation tool. *Bioinformatics* **22**, 2704–2705 (2006).

137. Aster, R. C., Borchers, B. & Thurber, C. H. *Parameter Estimation and Inverse Problems* (Elsevier, 2018)

138. Dutta, A. Covid-19 waves: variant dynamics and control. *Sci. Rep.* **12**, 1–9 (2022).

139. Ashyraliyev, M., Fomekong-Nanfack, Y., Kaandorp, J. A. & Blom, J. G. Systems biology: parameter estimation for biochemical models. *FEBS J.* **276**, 886–902 (2009).

140. Ding, F. Combined state and least squares parameter estimation algorithms for dynamic systems. *Appl. Math. Modelling* **38**, 403–412 (2014).

141. Dutta, A. Stabilizing COVID-19 infections in US by feedback control based test and quarantine. In *2020 IEEE Global Humanitarian Technology Conference (GHTC)* (ed. Cunningham, P. M) 1–6 (IEEE, 2020). https://ieeexplore.ieee.org/xpl/conhome/9342745/proceeding.

142. Myung, I. J. Tutorial on maximum likelihood estimation. *J. Math. Psychol.* **47**, 90–100 (2003).

143. Bavdekar, V. A., Deshpande, A. P. & Patwardhan, S. C. Identification of process and measurement noise covariance for state and parameter estimation using extended Kalman filter. *J. Process Control* **21**, 585–601 (2011).

144. Lillacci, G. & Khammash, M. Parameter estimation and model selection in computational biology. *PLoS Comput. Biol.* **6**, e1000696 (2010).

145. Pirgazi, J. & Khanteymoori, A. R. A robust gene regulatory network inference method base on Kalman filter and linear regression. *PLoS ONE* **13**, e0200094 (2018).

146. Medvedev, A., Zhusubaliyev, Z. T., Rosén, O. & Silva, M. M. Oscillations-free PID control of anesthetic drug delivery in neuromuscular blockade. *Comput. Methods Programs Biomed.* **171**, 119–131 (2019).

147. Tennøe, S., Halnes, G. & Einevoll, G. T. Uncertainpy: a python toolbox for uncertainty quantification and sensitivity analysis in computational neuroscience. *Front. Neuroinform.* **12**, 49 (2018).

148. Chatzi, E. N. & Smyth, A. W. The unscented Kalman filter and particle filter methods for nonlinear structural system identification with non-collocated heterogeneous sensing. *Struct. Control Health Monitoring* **16**, 99–123 (2009).

149. Raue, A. et al. Structural and practical identifiability analysis of partially observed dynamical models by exploiting the profile likelihood. *Bioinformatics* **25**, 1923–1929 (2009).

150. Cobelli, C. & Carson, E. *Introduction to Modeling in Physiology and Medicine* (Academic Press, 2019).

151. Mendonça, T., Lemos, J. M., Magalhaes, H., Rocha, P. & Esteves, S. Drug delivery for neuromuscular blockade with supervised multimodel adaptive control. *IEEE Trans. Control Syst. Technol.* **17**, 1237–1244 (2009).

152. Orsini, N., Li, R., Wolk, A., Khudyakov, P. & Spiegelman, D. Meta-analysis for linear and nonlinear dose-response relations: examples, an evaluation of approximations, and software. *Am. J. Epidemiol.* **175**, 66–73 (2012).

153. Ionescu, C., Machado, J. T., De Keyser, R., Decruyenaere, J. & Struys, M. M. Nonlinear dynamics of the patient's response to drug effect during general anesthesia. *Commun. Nonlinear Sci. Numer. Simul.* **20**, 914–926 (2015).

154. Arbabi, H. & Mezic, I. Ergodic theory, dynamic mode decomposition, and computation of spectral properties of the Koopman operator. *SIAM J. Appl. Dyn. Syst.* **16**, 2096–2126 (2017).

155. Korda, M. & Mezić, I. Linear predictors for nonlinear dynamical systems: Koopman operator meets model predictive control. *Automatica* **93**, 149–160 (2018).

156. Johnson, M. A.& Moradi, M. H. *PID control* (Springer, 2005).

157. Van Heusden, K. et al. Design and clinical evaluation of robust PID control of propofol anesthesia in children. *IEEE Trans. Control Syst. Technol.* **22**, 491–501 (2013).

158. Hägglund, T. *PID Controllers: Theory, Design, and Tuning* (ISA: The Instrumentation, Systems, and Automation Society, 1995).

159. Tan, W., Liu, J., Chen, T. & Marquez, H. J. Comparison of some well-known PID tuning formulas. *Comput. Chem. Eng.* **30**, 1416–1423 (2006).

160. Ziegler, J. G. & Nichols, N. B. Optimum settings for automatic controllers. *Trans. ASME* **64** (1942).

161. Lewis, F. L., Vrabie, D. & Syrmos, V. L. *Optimal Control* (John Wiley & Sons, 2012).

162. Franklin, G. F., Powell, J. D., Emami-Naeini, A. & Powell, J. D. *Feedback Control of Dynamic Systems* (Prentice-Hall, 2002).

163. Cunningham, J. J., Brown, J. S., Gatenby, R. A. & Staňková, K. Optimal control to develop therapeutic strategies for metastatic castrate resistant prostate cancer. *J. Theoretical Biol.* **459**, 67–78 (2018).

164. Khalili, P. & Vatankhah, R. Derivation of an optimal trajectory and nonlinear adaptive controller design for drug delivery in cancerous tumor chemotherapy. *Comput. Biol. Med.* **109**, 195–206 (2019).

165. Zurakowski, R. & Teel, A. R. A model predictive control based scheduling method for HIV therapy. *J. Theor. Biol.* **238**, 368–382 (2006).

166. Ali, N., Zaman, G. & Alshomrani, A. S. Optimal control strategy of HIV-1 epidemic model for recombinant virus. *Cogent Math.* **4**, 1293468 (2017).

167. Engelhart, M., Lebiedz, D. & Sager, S. Optimal control for selected cancer chemotherapy ODE models: a view on the potential of optimal schedules and choice of objective function. *Math. Biosci.* **229**, 123–134 (2011).

168. Schättler, H. & Ledzewicz, U. *Optimal Control for Mathematical Models of Cancer Therapies* (Springer, 2015).

169. Chen, T., Kirkby, N. F. & Jena, R. Optimal dosing of cancer chemotherapy using model predictive control and moving horizon state/parameter estimation. *Comput. Methods Programs Biomed.* **108**, 973–983 (2012).

170. Yadav, B. et al. Quantitative scoring of differential drug sensitivity for individually optimized anticancer therapies. *Sci. Rep.* **4**, 5193 (2014).

171. Dutta, A. et al. Model-based and model-free learning strategies for wet clutch control. *Mechatronics* **24**, 1008–1020 (2014).

172. Dutta, A. Optimizing antiviral therapy for COVID-19 with learned pathogenic model. *Sci. Rep.* **12**, 1–9 (2022).

173. Wang, B. et al. Similarity network fusion for aggregating data types on a genomic scale. *Nat. Methods* **11**, 333 (2014).

174. Rădulescu, I., Candea, D. & Halanay, A. Optimal control analysis of a leukemia model under imatinib treatment. *Math. Comput. Simul.* **121**, 1–11 (2016).

175. Krieger, A. & Pistikopoulos, E. N. Model predictive control of anesthesia under uncer- tainty. *Comput. Chem. Eng.* **71**, 699–707 (2014).

AUTHOR CONTRIBUTIONS

All authors contributed to the discussion of the contents, and reviewed and edited the manuscript.

COMPETING INTERESTS

The authors declare no competing interests.

ADDITIONAL INFORMATION

Correspondence and requests for materials should be addressed to Rongting Yue.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.