Research Square

Preprints are preliminary reports that have not undergone peer review. They should not be considered conclusive, used to inform clinical practice, or referenced by the media as validated information.

# Systems biology in disease modeling and control, analysis and perspectives

**Rongting Yue** ( rongting.yue@uconn.edu )
University of Connecticut

**Abhishek Dutta**
University of Connecticut

**Research Article**

**Keywords:** disease modeling, omics, drug molecules

**Posted Date:** June 2nd, 2021

**DOI:** https://doi.org/10.21203/rs.3.rs-578017/v1

**License:** This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

# Systems biology in disease modeling and control, analysis and perspectives

**Rongting Yue**[1,*] **and Abhishek Dutta**[2]

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

## ABSTRACT

Omics and drug molecules become increasingly influential in identifying disease mechanisms and drug response. Because diseases and drug responses are co-expressed and regulated in the relevant omics interactions, the traditional way that grabbing molecular data from single isolated layers cannot always obtain valuable inference. Also, adverse effects exist in drugs that impair patients, and launching new medicines for diseases is costly. To resolve the above difficulties, systems biology is then applied to predict potential molecular interaction elements by integrating omics data from genomic, proteomic, transcriptional, and metabolic layers. Combined with known drug reactions, the resulting models improve medicines' therapeutical 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 omics and drug molecules for modeling disease mechanism and drug response. The therapeutical performance can be improved by combining the predictive and computational models with drug administration designed by control laws. The challenges are discussed for its clinic uses.

The high mortality of many diseases prohibits human longevity. Therapy design is needed to suppress disease progression and aid organisms to recover from abnormal states[1]. However, the cost of launching new drugs is increasing due to the long-term safety procedures in clinical trials[2]. Drug side effects and toxicity should also be considered in treatment design due to its damage to patients. Side effects occur when drugs bind with unexpected targets and disturb normal regulations. The side effects of new medications are unknown and may vary dramatically for different individuals[3]. For example, a new drug Torcetrapib has been designed for cardiovascular disease[4], but the drug's uncertainty might cause severe side-effects of hypertension[5]. At the same time, drug toxicity is considered in drug overdose, drug-drug interactions, as well as adverse effects at therapeutic doses[6].

Omics analyses contributes to the investigation of disease mechanism and drug response. It refers to genomics, proteomics, metabolomics, and transcriptomics. Analyses of each omics layer focus on a specific aspect with reduced complexity[7], but only using a single omics level has limitations for analyzing disease[8]. For instance, analyzing conventional single marker or haplotype association can not reveal the combined effects of single nucleotide polymorphisms (SNPs), which may potentially induce stroke[9]. The integration of multi-level data, e.g., transcriptional factors[10], gene and gene expression products[11], provides more comprehensive information about omics interactions than that of single-layer for predicting complex traits[12]. It can also capture potential interactions among the omics data[13] through latent interactions from other omics levels. Thus, the analysis of complex cellular networks requires integrating the omics dataset from large scales[14]. Similarly, multiple aspects of biological processes have to be integrated for the model predictions to get full-scale patient models[15]. A general figure of omics interactions is shown in Fig.1.

Besides, in pharmacology, most 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[3] are not sufficient for meeting clinical drug safety criteria or specifying the significant variances among individuals. The clinical data from patients should be integrated to improve the predictive performance of drugs[19]. In systems pharmacology, drug actions are rooted in molecular interactions between the drugs and their targets in the context of such targets interacting with and regulating other cellular components[20]. And the new drug targets can be found by exploring other members of the protein families of the existing drug targets[21]. The modeling of multi-layer interactions (e.g., signaling pathways) in cell biology provides novel insights for cellular toxicity and adaptation mechanisms such that the drug interactions are predictable[6]. As for drug infusion control, the system-level analysis of dynamic gene regulation provides a highly effective

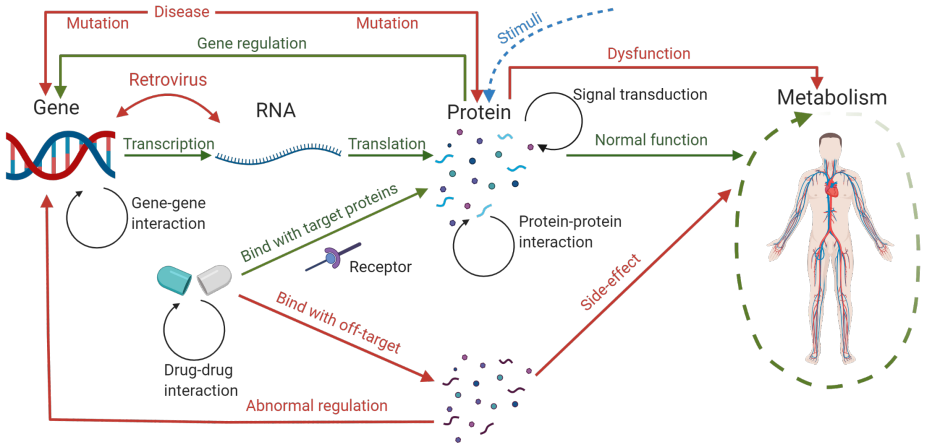



**Figure 1.** A systemic view of disease. Interactions among genomic, proteomic and transcriptomic levels reveal the regulatory process within organisms. Drug molecules intervene the process by binding with specific target ligands. These omics and chemical molecules should be analyzed simultaneously for studying the entire disease mechanisms and drug reaction.

drug administration to balance between side-effects and efficacy[15].

As in Fig.2, with omics data, researchers can investigate a single disease simultaneously across different clinical modalities (horizontal) or explore diseases from a single modality (vertical). The analysis of systems approaches for disease or medicines may give molecules (chemical molecules in drugs and biological molecules in organisms) more chance to be linked to a particular function or process by identifying their hidden connections. The disease progression and drug effect can then be tracked throughout different layers of the regulatory process. Thus, it is feasible to predict drug repurposing and drug combination for disease by searching the molecular interactions from existing studies. This review classifies the models for interactions among omics and drug molecules into two main classes: static network and dynamic modeling. Both the frameworks integrate biological information for drug combination and drug repurposing. We also discuss the techniques that 1) derive potential omics information from disease mechanism and drug response; and 2) design the drug infusion strategies. The commonly used tools and databases are listed in supplementary materials.

## Network structure in systems biology

A network structure is able to represent a wide range of components such as genes or proteins and their interconnections. With the network model, researchers can extend the use of omics beyond its original meaning. Network-based modeling can be established for systematic analysis based on the omics data from various scales[22]. This agrees with the previous ideas that the networks with hierarchical bio-information consist of the metabolic, signal transduction, and gene regulation pathways that contribute to the analyses of interactions between disease inhibitors[23]. The networks consist of interactions from tons of biological data that would make the bioinformatics more useful by mining structural motifs for novel interaction prediction[24]. Diseases with overlapping network modules show significant co-expression patterns, symptom similarity and comorbidity[25], whereas diseases residing in separated network neighborhoods are phenotypically distinct[26].

The network is composed of nodes and edges. Nodes can represent molecules of interest, e.g., genes[27], proteins[21] and chemical components of drugs. Edges link the nodes in the network. The annotations on edges can be functional interactions between molecules, including protein physical interactions, gene regulatory relations[28], mechanism of activation and inhibition[29,30], disease associations[31]. The annotations can be connective properties between nodes, including binding affinities[32], interactive directions[33], and the importance and confidence of the connections[34]. 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[35], and each subnetwork can be a set of functionally grouped molecules[33].

Network motif records structural patterns of particular functions or properties. Motif structures encode a regulatory behavior that enables network motifs to rapidly respond to internal and external stimuli and decrease internal cell noise[33]. The combination of motifs with similar functions can enrich the hierarchy and throughput of the associated nodes[36]. As for applications, the network motif helps identify drug molecules with common reactions, discover unknown drug responses, and predict potential therapeutics[37]. The patterns in the known annotations from databases can be used to predict new gene function annotation[38]. The regulatory motif in the gene regulatory network contributes to modeling the cell fate dynamics in the immune system[39]. The topological structure motif helps select target protein candidates for drug synergy[40]. And the pathway

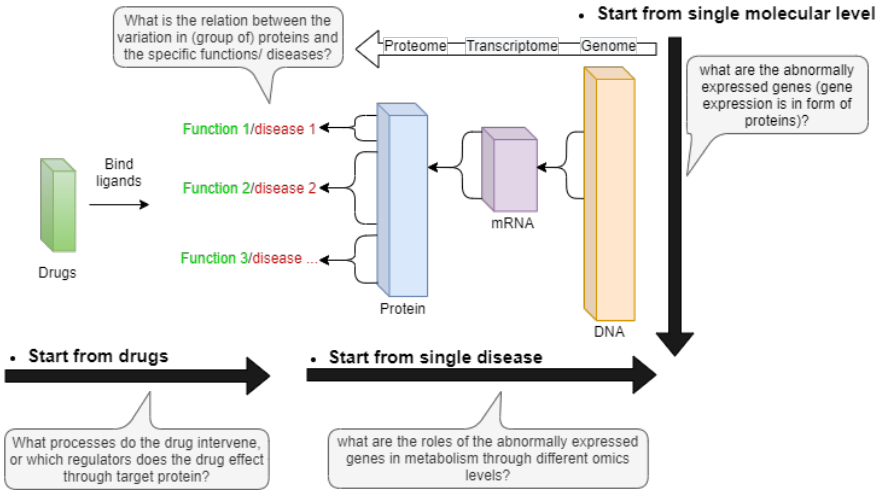



**Figure 2.** Analyses of disease and drug effect through single and multiple layers of omics information. Systemic view enables scientists to establish disease models from a higher hierarchic level by multi-layer data integration. Through vertical analysis within single layers, we can find out disease-related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

motifs within regulatory networks of genome-wide association studies give interpretations of genetic and epigenetic variation[41]. These motifs provide extra information which can not be derived from original nodes and edges since they separately spread in different scales. And it can self-organize and construct the network by shared components among different motifs. To derive the network motifs, one can decompose the biological networks into basic repeated modular components that consist of only a few genes and represent specific functions[27].

## Static network of diseases and drugs

Static network models molecular interactions that are statically functional. Network structure in the static network provides the topological properties from the presented interactions. It integrates intra- and extra-cellular information for identifying the modules' functional response by multiple network alignment. With omics integration, the overlayed multi-omics data can be more informative that provides unexpected biological interactions and reveals new findings[14]. For instance, collapsing multi-omics data into the gene-based interactions helps identify the disease-related genes[12]. Besides, the shared components can be the intermediaries for connecting different networks. For example, in disease–disease networks[28], the diseases are linked based on the shared genetic associations, the protein interactions, or the gene-disease interactions. The shared enzymes and regulatory components can be used to connect the metabolic reactions in a model of host-pathogen interaction networks[7]. For network medicine, the shared disease mechanism contributes to drug repurposing, and the shared gene associations can offer mechanistic insights into the common pathophysiology of diseases[25]. In a word, network analyses may generate new findings by integrating molecular data from the individual level with the shared components.

**Model structure in static network** The network frame has topological descriptors for different properties. The degree of a node refers to the number of connected edges[28], which is relevant to the total number of edges connected to that node. Nodes' degree has no limitation in the scale-free network. It enables us to add all information related to the particular node. Besides, hub nodes would be a key component in the biological interaction network. In protein interaction networks, hub refers to the protein with a central position in cellular network[33]. It is the focal node shared by many other cellular functions and has become information exchanging and processing center. The two types of hubs in the human interactome are the intramodular hubs and the intermodular hubs. They are co-expressed with the interactive partners in most tissues or a tissue-restricted manner, respectively[19]. Hub is not limited to proteins. It can also be the genes as global transcription factors in gene regulatory network[33]. The betweenness mathematically measures the shortest path that describes the centrality of the given nodes[42]. It will change when removing intermodular hubs compared to the intramodular[19]. The distance measures the shortest summarized edges between two target nodes[43]. By selecting these topological descriptors properly, we can identify the pivots in disease mechanisms.

**Omics interactions** Omics molecules interact with each other during the functional regulation of an organism. Protein interaction networks encode the information of proteins (nodes) and their interactions (edges) into the network structure, and it also refers to the protein-protein interaction (PPI). PPI network is capable to predict the potential disease-related omics molecules[7,44,45] based on the assumption that: the interactive proteins may cause similar disease phenotypes with the shared components between the PPI networks (of a specific disease) and the disease-related pathways. The assumption is consistent with that the relatedness in PPI can identify the variants that may act jointly on an intermediate phenotype such as biologic pathway[9]. By combining high-throughput bioinformatics, PPI network reveals disease connection[31] and expand the scale of existing cellular networks[33]. The common interactive neighbors or the shared protein interactions connect different protein networks during the network alignment[21]. Besides, genes and proteins are always analyzed simultaneously due to their expression relations for gene function discovery and the genetic analysis of complex diseases. As for genome-wide association, proteins and gene interactions can be applied in identifying densely connected modules in the human protein interactome[9]. Covid-19 pandemic has been devastating in terms of human lives and economy[46]. A model based on PPI and gene co-expression network[47] has been proposed to search the virus-host protein interactions and assess the host-pathogen response for clinical treatment of human coronavirus (HCoV) infections. The nodes in the proposed proximity network were assigned a proximity value as seed protein. Functional analysis of HCoV–host interactome was carried out to provide a theoretical host-pathogen interaction model for HCoV infections and translated the results to predict SARS-CoV-2 pathogenesis. Some other molecular mechanisms on genomic level include the interactions between genotype-phenotype[41], protein-ligand[23] and gene-gene[33]. Besides, human transcriptomics data have been analyzed for revealing the pathways related to longevity[48] by identifying and rating the variety of factors with the known effects on the transcriptomes at different ages. All possible activators and repressors' gene expression levels in this model were used to quantify disturbance in the signal pathway. Drugs that minimized the difference of metabolic states between the old and the young patients were identified. In addition to gene and protein network, the metabolic networks enable researchers to investigate protein functions in the network separately and identify their steady states[7].

One way to identify the disease-related genes (that account for the disease mechanisms) is to screen out the differentially expressed genes (DEG) by selecting the most variant genes. With 'Limma' package[49] in R, statistically analyses can be performed on the collected microarray data, e.g., gene expression profiles. This tool compares the expression levels in disease and control states. The gene expression profiles with large variation are selected out by setting certain statistical threshold (i.e., fold-change and p-value). These representative genes can be mapped into PPI netowrk for the functional enrichment analysis of the given disease through the correlated pathways[50] or other molecular mechanisms. This method mathematically identifies the genes with large variations that account for the disease mechanism. 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 contribution to the overall process. In another frame work, the "Weighted Gene Co-Expression Network Analysis" (WGCNA)[51], has been widely used to detect the functional gene clusters based on the weighted correlations. It constructs the omics co-expression networks that are close to scale-free distribution. This framework has been implemented as a package[52] in R. Based on the adjacency of gene expression profiles, the gene clusters make more sense for functional co-expressions in organisms. Each cluster' significance is quantified with the eigengenes derived by singular vector decomposition[52]. The software Pathway Tools[53] is another tool that provides genomics and metabolomics level analyses. Its web-based querying function finds out data of interest and visualizes the resulting omics annotations. Pathways of metabolites can be searched and modeled to predict functional molecules in gene regulation.

**Drugs and targets interaction** The interactions between different drugs or between drugs and targets can also be put into network structures for analyses. Drug-target interaction (DTI) networks have been applied to study the prediction of drug response[54], new drug-target pairs[32], the identification of the unknown drug combination for side effects[23,55,56] and cancers[57], and the examination of the global relationships between drug-target proteins[16].

A DTI-based target inhibition model has been proposed to identify the disease-specific target set with possible drug-target combination and functional significance in disease meachanism[57]. The model mapped cellular targets of drug into drug-drug inhibition profile and computed the target pairs' synthetic lethality effect score. This model's input was the potential target set that was relevant to cancer survival and would bind at least one of different drugs. Notably, the drug target proteins in DTI network have larger degrees (which means more interactive molecules) than proteins in the PPI network[16]. Besides, the "pharmacological space" has been proposed to map and predict the interactions between drugs and target proteins[58] based on the chemical structure information of drug molecules and sequence information of proteins.

In DTI networks, nodes can be drugs, drug molecules, and drug target/off-target proteins. Node sizes can reflect the number of target proteins that the specific drug is correlated with[16]. Links between drugs and proteins can be established if the protein has been studied for the drug[58]. Two drugs will be linked if the shared target protein exists. Two target proteins will be linked if they are both targeted by the same drugs. Similarly, link widths can be proportional to the number of common drugs targeting both connected proteins. For predicting off-target molecules and side effects, the off-targets can also be linked based on the

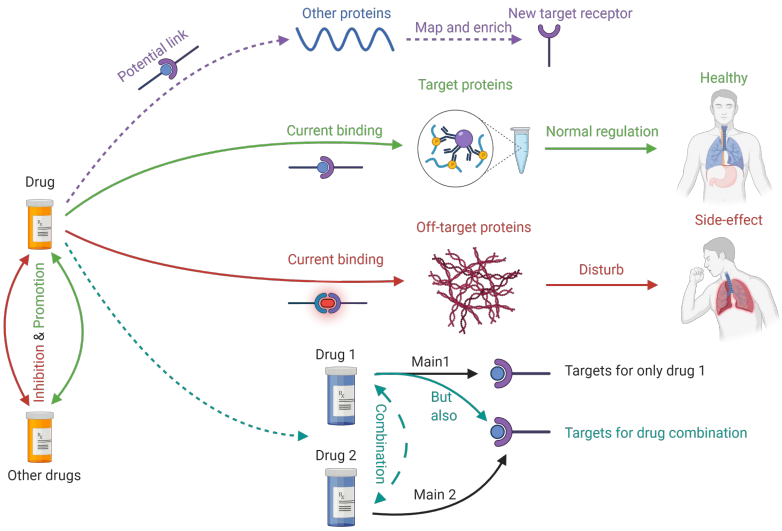



**Figure 3.** Drug and target interactions. Drugs may bind with off-target proteins that induce side-effect. Drug molecule interactions may improve or inhibit the efficacy, which offers chances for drug repositioning and drug combination.

quantified drugs clinical relevance[56]. Besides, when applying drug combinations, different drug agents may have interactions that induce unexpected side effect to patients or reduce the drugs efficacy[59,60]. It requires researchers to predict potential drug-drug interactions (DDI), which can be explored by expanding DDIs with the shared target components of the drugs. The enriched DDIs can then identify potential targets of the drugs being tested, which reveals new therapeutic use and combination with surprising efficacy or for the diseases that the drugs do not initially aim at[3]. An abridged general view of drug and target interactions is shown in Fig.3.

The text mining-based predictions of molecular interactions rely heavily on the existing reports and references[61], which indicates that insufficient clues would result in inaccurate predictions for the combination of omics and drug molecules. The predicting performance also depends on the setting of the cutoffs. Instead of the hard cutoffs, soft cutoffs (which may be similar to the soft thresholds[52]) may be used when considering the generality for different data sets. Some of the frequently used databases for omics interaction analyses are listed in Table.1.

## Analysis of static modelling

**Importance quantification** Although the models of diseases or drugs have been updated and well-studied for decades, the actual biosystems are far more complicated from complete modelling[63]. Though, not all the multi-layer information will be used after quantifying the response measure for efficiency. The importance of the interactions requires measuring and ranking for reducing the complexity of the networks or generating the corresponding weights. The simplified network will include only those most relevant molecules from a biological standpoint and directly associated with questions we concern about[75]. Thus, the initialization of static network analyses can be the importance quantification and ranking of its components, such as genes, interaction importance[63], treatment or drug efficacy[48]. Z-score quantification have been used to measure the importance of the shortest path lengths between targets of drugs and proteins correlated with the cardiovascular module, which reduces the bias brought by the highly connected nodes[55]. Network proximity was applied to measure the importance of network distance between a drug and a given disease. In PPI networks, the edge weights quantify the confidence of interactions[50]. The importance of cell physiology has been measured by the observable concentration levels, the change rate, and the initial concentration of ligand proteins in a cell signaling models[76]. For example, in a protein interaction network[63], the top 20 ranked cardiac septal defect candidate genes have been selected to represent most of the functional interactions. The known disease genes were seeds in the network, and the prioritization defined the rank of the importance of all the other nodes towards the seed proteins. Rank of the same proteins would be unique since no overlapping happens within the disease network. All the other proteins contributed to the vectorized interaction measurement toward each of the specific seed proteins. In DTI network, the shared drug-substructures prioritize DTI pairs to build the links between the new chemical entities and known drugs[61]. The 'resource-spreading processes' that utilized adjacency of these interaction pairs were proposed to quantify the interaction probabilities and make predictions. Also, drugs can be repurposed with quantified drug-target binding affinity. For instance, Fabahistin is designed as a symptomatic antihistamine, while it can be used for Alzheimer disease since it has higher off-target affinity to the disease-related ligands[32].

**Table 1.** Databases for omics 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. | 40,55 |
| 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. | 11,43,50 |
| Genomics of cancer | Cancer Genome Atlas (TCGA) | Cancer molecular data including genome, epigenome, transcriptome and proteome. | 41,62 |
| Biological pathway | GeneGo MetaBase | Bioinformatics including signaling and metabolic pathways, interactions among drugs and proteins as well as kinetic information of drugs. | 56,63 |
| 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. | 12,47,64,65 |
| 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. | 9,47,50,66 |
| 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. | 54,57,67 |
| 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. | 54,66,68 |
| 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. | 19,38,47 |
| 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. | 40 |
| DTI | ChEMBL | Biological activities and characteristics of molecules such as chemicals and proteins that contribute to the study of drug target and drug discovery. | 56,56,69 |
| 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. | 11,70 |

| Application | Database | Description | Reference |
|---|---|---|---|
| DTI, DTI | Drug Bank | Drug-target and drug-drug interacting information such as chemical sequence, three-dimensional structure and pharmacological pathway involvement. | 16,65,71,72 |
| 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 evaluate the effectiveness of survival prediction with network regularization. | 62,73 |
| PPI | HPRD (Human Protein Reference Database) | Bioinformatics of human protein-protein interactions from literature and data are manually curated. | 31,44,74 |
| Signaling pathway | Reactome | Bioinformatics information including pathway, proteins and drugs for model visualization and analysis. | 9,34,47 |
| 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. | 16,33,74 |
| PPI | Human Protein Reference Database (HPRD) | Protein information based on interactions described in published reports. This interaction set is expected to be biased toward known disease genes. | 14,31,44 |

**Similarity analysis** Similarities describe the closeness between the components in static network. It helps convey the knowledge from the well-studied components to the unexplored ones, and it has the potential for novel clinical findings when two components share the same neighbor[3]. From the view of DTI and PPI, drug target proteins in the significantly differentially expressed signature modules can share the same protein family for the given diseases[31], which indicates that similar drugs may be suitable to treat a cluster of similar diseases[65]. Drug pairs with more similar structures have significantly higher drug sensitivity correlations by activity area, which suggests that drugs with similar chemical structures show similar inhibitory effects across the cell lines tested[54].

There are different types of similarities in molecular interactions. Cell lines with similar gene expression profiles exhibit similar drug responses, which can be used to predict the drug response of cancer patients to therapeutic agents[54]. The sequence-similarity ranking has shown good performance for prediction based on ligand similarity[32]. Similarity in phenotypic side-effect contributes to drug targets identification. This has been validated experimentally that the antiulcer drug rabeprazole binded the serotonin receptor HTR1D and inhibited the dopamine receptor DRD3 through similarity in drug-target connections[71]. The global structure similarity has been used to build connections between the off-targets[23]. In DDI, the similarities of positive DDI pairs were significantly higher than those of random DDI pairs[77]. With similarity inherence and the collective inference approach, the Probabilistic Soft Logic language[65] conveys the similarity of drugs (targets) to new targets (drugs) through the known DTIs.

The similarity in a network-based inference model of drug and target interactions can be derived by the drug-drug structure, the target-target genomic sequence and the drug-target topology network[78]. When identifying off-target effect in drug activity[56], the two-dimensional structural similarity of a drug to the target's ligand set was quantified as an E value using the similarity ensemble approach. This method computes whether a molecule will bind to a target based on the common chemical features with known ligands. The pathological similarities used for DTI prediction can be quantified by the Partial Spearman correlation calculation[31] with module response scores. The binarized fingerprint of side effects and the Jaccard index can be used to compute similarities of drug-target pairs[79]. Rules of how the shared drugs and targets convey similarity have been compared[65].

**Data fusion** Combination of high throughput data contributes to the prediction of biomedical outcome, which means data synthesis of specific module functions and network motifs makes data more informative. Useful information will be maintained after fusion. For omics information, the data fusion is often used by incorporating multiple data types and sources to expand the level of omic and pharmacological data from molecule to organism[3]. Similarity contributes to the fusion of the separated

networks which consist of data from large scales such as genomics and its transcriptional products[62]. Data fusion has been used to refine essential and common information in cancer type identification[62]. The subnetworks, which were constructed for data of each scale by similarity measurements with non-linear fusion, eventually converged to one single network by fusing iteratively. Higher similarities from various levels were fused together in the resulting network. Lower similarities were discarded since they were regarded as noise. The fusion of different data and knowledge sources has been used to study drug repurposing by applying the tri-factorization algorithm[40]. Data fusion by matrix factorization in drug synergy was applied to identify the proposed method's potential target protein combinations. Besides, kernels or kernel functions that extract data features by matrix multiplication can be powerful data fusion tools. For drug side-effect prediction, the weighted kernels of the drug spaces and the side-effect spaces have been fused for information integration[80]. A multi-kernel based method, which refers to a composite kernel with a weighted average of each single kernel, has been utilized for omics integration in predicting the genomic risk[13]. The robustness of the proposed method for different disease analyses was retained by choosing the appropriate kernels' predictive genomic regions. Additionally, whether the fused information contains enough features from the original data depends partly on the kernels' selection or design. Some recent progresses of omics integration by kernel learning have been summarized[81]. While the integrated information generalizes the scope of predicting model, the weights that are used to overlay multi-layers should be carefully assigned, which can improve the accuracy of the predictions.

**Learning-based methods for clustering and classification** The evolution in learning-based networks has proven to be a promising architecture for efficient learning from massive datasets[3]. Learning-based approaches efficiently utilize computer power for extracting more complex features from high-throughput data in hyper-dimensions. It can identify global pattern of network from multi-layer omics data by combining them into fewer layers, which retains the maximal covariation from each original level. The learning-based methods can be supervised (with labels for training data) or unsupervised (without labels). A supervised learning approach enables the model to predict unknown part (links) of the network based on the known interacting molecules[58,64,70,82].

Support Vector Machine (SVM) is a classical supervised learning model for classification problem that can be applied to refine topological information from network structure. The multi-class SVM has been used to predict the therapeutic class of FDA-approved compounds and the classifier is trained with integrated drug similarities[72]. Kernels in SVM measure the features between gene pairs to train the classifier[83], and the classifier makes the binary prediction for the interactions between the existing molecules and the incoming components[64]. Moreover, the convolutional neural network, which outperforms SVM since it does not require manually defined features, has been applied to extract DDIs with text mining on biomedical information[60]. Besides, a supervised bipartite inference can predict unknown target proteins of drugs in drug-target interaction[84].

Deep learning (DL) methods outperform the previous learning-based algorithms as for the property of structure flexibility[85]. It has recently gained considerable attention as it allows the model to 'learn' to extract relevant features of the task at hand[86]. DL is also capable to extract molecular patterns by mining latent information from the network structures[66,69,87,88]. Given network structure, DL helps predict annotations[38] and bingding affinity[86] of the edges. A convolutional neural network can be used to refine molecular features from arbitrary network frame of different sizes and shapes[88]. This approach transformed 2-dimensional neural graphs into state matrixes with informative nodes and edges. Besides, auto-encoding automatically and dynamically extracts features from molecular structure, so that a minimalistic description can be derived for differentiability among samples[87]. An autoencoder (AE) can transform molecules directly into a numerical representation[69]. AEs can be used to denoise a single-cell RNA sequencing model[89] and to compresses computational dimensionality[3] in detecting the cell type-specific clusters in an ensemble clustering study[90]. The stacked AE has been used to identify the potential DTIs with information from drug molecular structure and protein sequences that generates highly representative features from multiple layers[68]. The feature descriptors were constructed by combining the molecular substructure fingerprint information and fed into the rotation forest for accurate prediction. The experimental cross-validation showed that it achieved superior performance on gold standard data sets with high accuracy. Additionally, the accuracy of the prediction can be assessed by area-under-curve derived from receiver-operating-characteristics curves[3,91]. Consequently, the classifiers generated from the clustered model can be utilized to predict the unknown functions of the nodes in PPI network[73], unknown interactions between drug and targets[84], potential interactions between proteins and the regulation between disease and genes[74,91]. A scheme of a convolutional neural network is shown in Fig.4.

The omics interactions identified in static framework reveal the possible connections within organisms. By selecting the disease-related omics from the resulting predictions, one can search the dynamic pathways in the regulation and identify key regulators for disease progression. With potential drug targets, drugs may be repurposed for more relevant disease. Interactions between drug molecules may either reduce or enhance drug efficacy, which may help design novel schemes for drug combination in disease treatment, such as drug cocktail[92]. Additionally, the clinical use of potential drug combinations or drug repurposing requires high confidence in the molecular interactions for patients' safety, so the accuracy of current algorithms can be improved, since it may not have good performance when testing on different dataset. While the commonly acknowledged problem in analyzing interactions between omics and chemical molecules based on database is data scarcity[22].

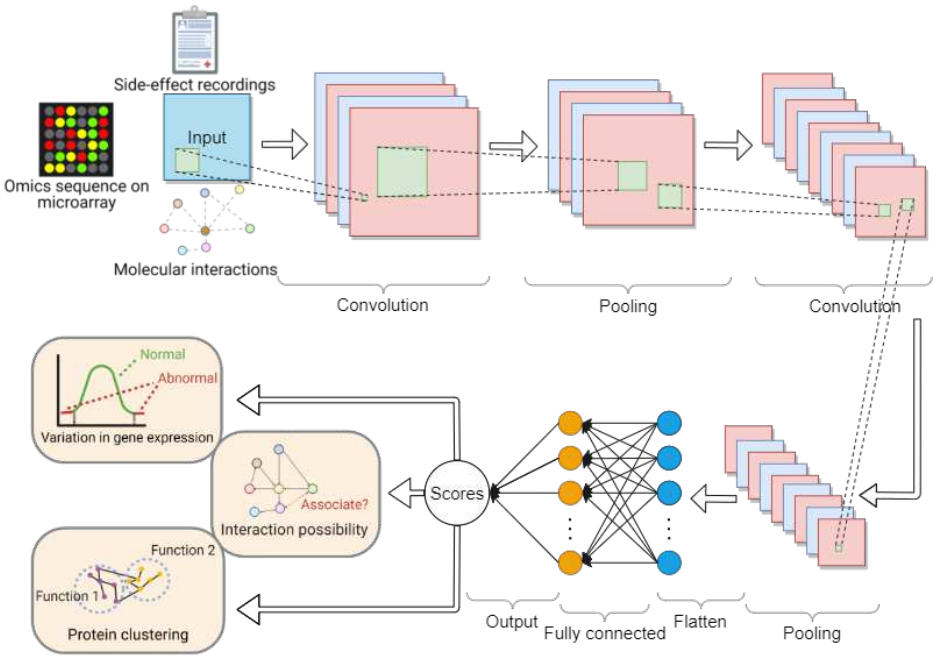



**Figure 4.** A convolutional neural network. This neural network is comprised of an input layer, convolutional layers (that extract features from hyperplanes of input data by projection/convolution), pooling layers (that reduce the spatial size and mitigate the locational sensitivity), flatten layer (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 omics sequence patterns, gene regulation annotations and patterns, omics 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., omics binding profiles) and regression (e.g., quantified molecular binding affinities) when obtaining new samples.

More accurate descriptors will be discovered as more data is collected.

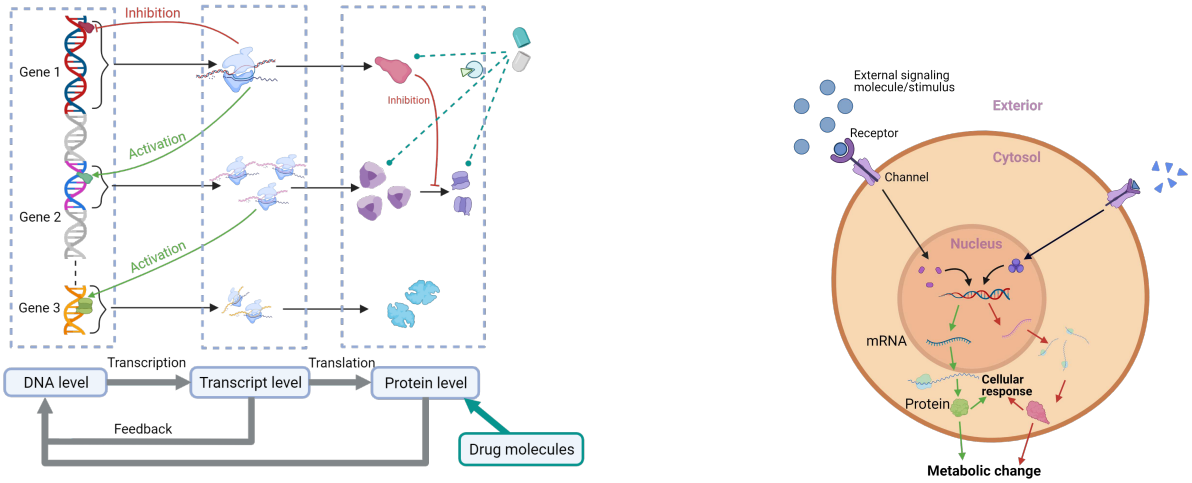
## Dynamic modeling of regulation in organisms

The term dynamic disease can be explained as a disease whose pathogenesis is mainly caused by the appearance of new dynamics of the organism behavior, independently of the underlying pathogenesis[93]. Disease outcome can be predicted by changes in patients' interactome[19]. In dynamic regulation, the systemic view tells us that genes work as part of complex networks instead of acting in isolation to perform cellular processes[94]. Multiple differential co-expression networks outperform single co-expression networks as for inference accuracy of gene pathways[34]. For instance, drug efficacy and reservoir contribution rates cannot be identified from decay-phase data alone[95]. Besides, 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 progression with treatment[95], which help researchers find a functional cure for diseases. Once the dynamic models have been derived, the performance of a kinetic model can be assessed by its ability to generate testable predictions[76]. Notably, the goal of constructing dynamic networks is to identify the biological patterns by modeling the time-course behaviors of gene expressions[27].

**Interactions in gene regulatory network** Dynamic analysis in organisms focuses on the omics variations, which are the results of gene expressions. Genes interact with each other through their RNA and protein expression products, thereby governing the rates at which genes in the network are transcribed into messenger RNA[93]. The causal information in gene expression data, e.g., key drivers of complex traits in phenotype-related gene interaction network[11], can be identified by variations in DNA and gene expression-related traits[27,39]. Gene regulatory networks (GRN) are thus used to detect potential mechanisms based on dynamic behaviors of epigenomic activity between normal and disease state[94]. Notably, GRN is regarded as the reverse engineering for investigating gene regulatory relations by going backwards with the observed gene expression[33]. The schematic diagram of GRN is shown in Fig.5a.

GRNs have been used to analyze infectious diseases[14,96] to detect gene regulations related to the infectious and viral mechanism. A combination of gene regulatory components including myeloid and lymphoid has been studied to identify cell fate specification[10]. The genes related to the altered magnitude of the plasma hemagglutination-inhibition response, which are known to be regulated by XBP-1 (a transcription protein related to immune system), have been identified in the study of human innate and adaptive responses to vaccination against influenza[70] through the GRN. Interferon-related genes regulation interactions were used to get the molecular signatures for antibody response prediction based on data from the vaccinated patients and the mice trials. A Differential SparsE Regulatory Network model[94] has been proposed to identify topological changes in gene-regulator dependence networks for cancer study. The gene expression levels were modeled by a sparse linear model that enabled the researchers to capture conditional dependencies efficiently from genome-wide expression data. In inference regulatory interaction networks[29], nodes were genes coded with expression levels, and edges were the genes interactions. To construct GRNs, one can use the time-course microarray profiles with network inference[27], the combination of regulatory proteins and disease-related genes[33].

**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[34]. At the protein level, signaling pathways are made up of protein interactions covering almost all the biological functions in living cells. It captures the inter- and intracellular regulatory mechanisms of gene transcription and protein synthesis[29,33]. For example, the analysis of dynamic signal transduction networks helps explore cancer cell phenotype, tumor cell heterogeneity and drug resistance, since cancer is considered a genetic disease[36]. When analyzing the interferon-stimulatory DNA-sensing pathway, the researchers have measured dsDNA response and built the connections between the predicted genomics with a known signaling network to identify the disease-related proteins and the functional participants[96]. Errors in signal transduction can lead to altered development and incorrect behavioral decisions in organisms, whose dysfunction may result in uncontrolled cell growth or tumorigenesis[33,36,97]. The dynamics of signaling pathways can also be a powerful measurement to model molecular regulatory behaviors of fungal pathogen infections in a fungal signaling network[7]. Similar to DEG, multiple differential modules[34] have been used to predict both unique and shared gene modules across multiple differential co-expression networks. This work explored the systematic profiling of the transcriptome in a heart failure study based on RNA-Seq data. The individual pathways were identified by the specificity inference of multiple differential modules, and the enrichment analyses were also performed for genes in differential modules with known cardiovascular phenotypes. Notably, the pathway-wide association can also be used to identify valuable information from background noise and the context-specific logic of GRN[41]. Transcription factors partly determine the cell fate specification. It regulates the development of innate and adaptive cells of the immune system[10]. The dynamic subnetworks of transcriptional and translational architectures have been explored to construct and model the entire trigger mechanism of the innate response regulated by intercellular and intracellular heterogeneity[14]. Its transcription factors in the downstream processes were generated




**(a)** Gene regulatory network

**(b)** Signal transduction

**Figure 5.** Schematic diagrams in dynamic modeling. (a) Gene regulatory network adapted from the work[27]. 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 cease of cell apoptotic that results in unlimited growth and division.

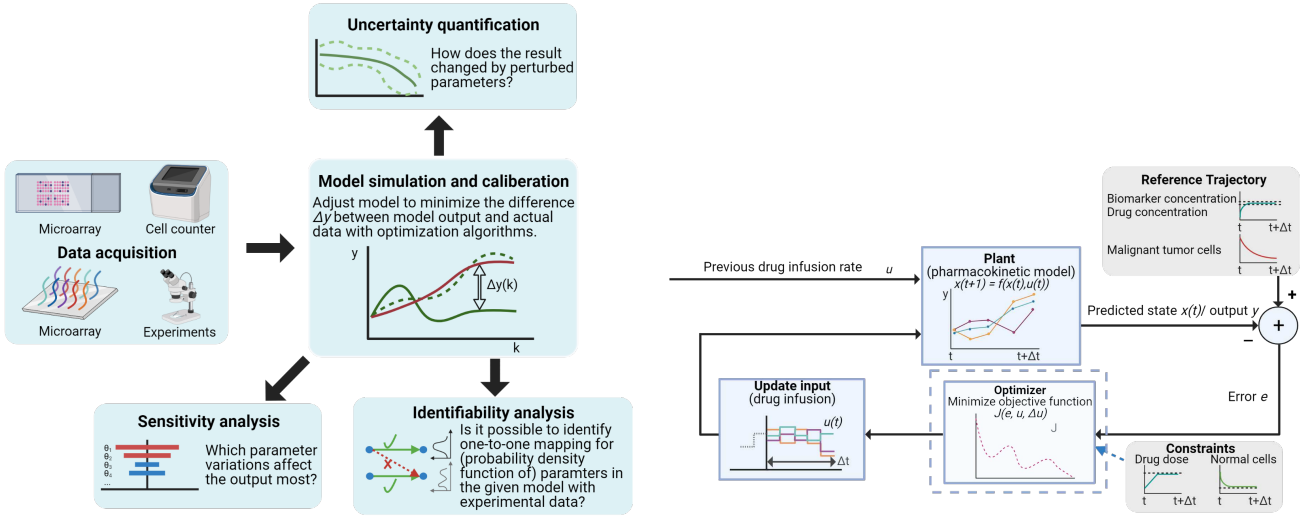
or activated for further target gene-related translations and transcriptions with the detected pathogen ligands. A diagram of signal transduction is shown in Fig.5b.

## Analysis of dynamic modelling

**Mathematical modelling for dynamic regulation** Dynamics of omics data reflects organism's response to the changes of interior milieu or environmental factor. Variations in metabolisms and phenotypes can be huge even if most of the corresponding genes are the same[98]. Modeling of the dynamics exerts mathematical tools that quantify the rate of state change in dynamic regulation[27] in different conditions and time sequences. With knowledge of molecular interactions, the Differential Equations (DEs) can be used to capture how the system reacts to the variations caused by disease or drugs. These equations integrate information from different levels into a unified form, which agrees to the integration of omics information in systems biology[35]. DE models the dynamic variation in transcription[29], gene expression in regulation[27], the metabolite concentrations[99], the reaction rate[35], and the factors in signal transduction pathway[98]. It also quantifies signal flow in pathways and explores the effect of oncogenic mutations on dynamics of ligands[76]. For example, as for disease modeling, DEs capture the dynamics of viral infection, drug response, and interactions between HIV and the specific immune response in an HIV model[100]. As for drug response, DE has been applied to study drug intervention in an irradiation-induced cellular senescence model[101]. A system of DEs, which consist of drug concentration and populations of tumor cell, effector-immune cell and circulating lymphocyte, were summarized to model the growth, death, and interactions of these populations with chemotherapy[102]. Components in these DEs can be the metabolite concentration (in nonlinear biochemical models)[99], the signal transduction molecules (in dynamic cell compartment models)[35], transcription factors and regulatory site[29].

Additionally, the established dynamic model will have the specified inputs and outputs that enable the prediction of crucial components in biological regulation, and a good dynamic model should be robust to perturbations, which allows investigators to track abnormal variations[24]. With more interactive terms identified to be involved in disease progression and drug reaction, DEs can model organisms' actual regulatory mechanisms more precisely. Note that molecules should not be added into model equations unless it is of significance for the given processes.

**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[103,104]. In systems biology, it is usually part of an iterative process




**(a)** Parameter estimation

**(b)** Model Predictive Control

**Figure 6.** (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 model-predicted output, this control strategy updates control input (e.g., drug infusion) 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.)

to develop data-driven models for organisms that should have predictive value[105]. One way to perform parameter estimation is to optimize the fit of a collection of parameters iteratively to a given data set[104] with random disturbance, which explores the parameter space extensively and limits the number of non-convergent solutions[101]. The least-square estimation (LSE) and the maximum likelihood estimation (MLE) are two widely used techniques for this task. LSE estimates parameters for a nonlinear dynamic system by minimizing a quadratic criterion function that comprises of the state estimates and measurements, subject to the constraints of the system[99,106]. MLE is used to search for the value of the parameter vector that maximizes the likelihood function, which is most likely to describe the observed data based on probability distribution[107]. MLE has been used to directly estimate the efficacy of antiviral drugs from the viral load data[95]. As an optimal linear estimator, Kalman Filter (KF) recursively generates the maximum likelihood estimates for a linear dynamic system from a series of noisy measurements[108,109]. It handles the approximate modeling of high-dimensional noisy data and a low number of samples[110]. Based on KF, its expanded versions have been utilized to deal with nonlinear models in biology. For example, the Extended KF, which is the nonlinear version of KF, can be used to estimate states and parameters in a nonlinear system[111,112]. Note that the noise uncertainty of KF is subject to Gaussian distribution. When dealing with non-Gaussian processes, Particle Filters (PFs) can be a good choice. By randomly drawing samples from numerical simulation, the Monte Carlo method can be used to estimate parameter values[95] and quantify their uncertainties[113]. As a sequential Monte Carlo method, PFs obtain weighted samples from the non-Gaussian posterior probability of the state to deal with nonlinear systems[114]. It has been studied that a PF with an orthogonal basis (that was used to approximate the posterior by an orthogonal series expansion) outperformed Extended KF when estimating the Wiener model parameters in an anesthesia delivery model[112]. The process of parameter estimation is shown in Fig.6a.

Sensitivity analyses assess how sensitive the models' outputs are to the fitted models' parameters changes. The integral of absolute sensitivities can be used to analyze the effect of relative changes within a time slot, and the irrelevant sensitivities which have peaks close to zero should be discarded[35]. Dynamic sensitivity analysis has found that treatment effectiveness got gradually reduced and all interventions became largely ineffective later in a cellular model for drug response[101]. Sensitivity can then quantify model uncertainty through finite differencing or variational equations[105]. Additionally, the variability that tells model variance can be assessed by a principal component analysis[101]. Besides, identifiability checks the reliability of the estimates and assesses how well the model explains experimental data[109,115]. The fitted model by parameter estimation may not have the generality for other datasets (e.g., clinical data varies for different patients). This may be solved by identifying the parameters that affect the variances between individuals.

Tools have been integrated for modelling and assessment. Systems Biology Markup Language based Parameter Estimation

Tool[103] has been developed for parameter estimation under specific customized requirements (e.g., changing signal strength). The evolutionary optimization with stochastic ranking was implemented in this tool. Recently, the 'Data2Dynamics' modeling environment has been widely used for dynamic modeling since it implements different efficient approaches for parameter estimation as well as uncertainty evaluation[116]. The functional package 'Uncertainpy' quantifies the uncertainty and performs sensitivity analysis with visualized results[113]. Time cost is considered in the modeling, which depends on the speed of getting solutions of those DEs. The solutions should be the optima for the estimation problem. The global optima for parameter estimation are always desired, leading the multiple approaches to converge to the same solution ultimately[99]. It is often used with a computational time limit to prevent an endless search[105]. The global optimization problem is stated as the minimization of a weighted distance measure between experimental and predicted values of the state variables in a nonlinear biochemical dynamic model with DEs[99]. When estimating parameters, calling the single clusters recursively may cause the absence of global parameters[35]. Also, we may get stuck in local optimum[105], which is misleading since it matches the data well in the current data set. However, it will give a bad fit to data from other sources or fails to converge to a satisfying result[117].

Once the numerical values of model parameters have been estimated, the approximation of dynamic omics regulation and pharmacokinetics from a systemic view is obtained for patients. Next, drug scheduling can be designed for patients' recovery based on the models.

**Drug administration with control theory** In control system, feedback signal provides error information to adjust the input and track the reference such that the system can be brought to the desired states. The meanings for feedback loops from biological aspects may vary. The negative feedback refers to inhibition and the positive feedback refers to activation from the downstream to the upstream[36], and the feedback control in signal transduction networks includes positive feedback (signal amplification) and negative feedback (signal attenuation). Notably, feedforward signaling exists in specific transcriptional procedures[14]. Diverse properties of regulation in organisms come from a coherent or incoherent feed-forward loop by combining the inhibition and activation[36,93]. The feedback signal adjusts the drug infusion, which may consider the changing drug concentration, toxicity level, and cell population. Drug concentration in patients is delicate, especially for the fast body function change (e.g., patients do not react to anesthesia delivery during surgeries[118]). The dangerous therapeutic window of drug concentration in plasma defines whether the concentration level is tolerable and the drug is effective for patients[17,67].

The components in a control system include the states (e.g., time-evolved behaviors of the system), the inputs (e.g., variables that are manipulated by the control system such as drug doses), targets (that are the reference trajectories such as drug concentration), and the outputs (that are the elements we want to observe). For example, in an HIV-1 epidemic model[119], state variables can be the density of virus, recombinant virus and the double-infected cells. Its control objectives are to use drugs to prevent new cells from being infected and decrease virus generation. Notably, dynamic models of drug response in real-world will be nonlinear for most of the cases. However, when taking drug infusion rate and plasma concentration as input and output respectively, a pharmacokinetic model can be treated as linear[118].

As a classical control strategy, the Proportional, Integral and Derivative (PID) controller[120] maintains drug concentration at reference level in organisms. For example, in a neuromuscular blockade model[118], a PID controller determines the dose of nondepolarizing types of muscle relaxant. The PIDs acted as local controllers in the proposed supervised multimodel adaptive control scheme. It was used to cope with the uncertainty in the system's dynamics caused by interpatient variability and time variations. The concise scheme of PID[120] makes it flexible for functional expansion. Based on a circulation model, a IPD controller was implemented in a chemotherapy control scheme that achieved and maintained drug concentration at the desired level[18]. When implementing PID, the tuning process can improve controllers' performance and robustness[121]. Recently, an auto-tuning algorithm for PID based on system specifications (e.g., response to a given input) has been proposed that almost met equivalent specifications compared to the design based on knowledge of the full process[122]. To achieve an adequate paralysis level during surgery, a time-varying PID controller paired with a tuning procedure assists the closed-loop drug infusion system recover from sustained oscillations in an autonomous way[112]. The instability of equilibrium caused by dynamic alternations in the PID model was addressed by a stable limit cycle, and bifurcation analysis has been performed on the nonlinear dynamics of the designed control scheme to prevent the oscillation at the unstable equilibrium points caused by plant uncertainty.

The optimal control problem is to find the control input on a time interval to minimize control errors (e.g., drive the systems along trajectories) and control efforts (e.g., less energy consumption)[123] subject to system dynamics. In chemotherapy, its goal is to find the balance between cancer cells elimination and drug toxicity. The drug concentration can be maintained within an effective therapeutic window, and side-effect caused by toxicity can be minimized. The general scheme is to design a proper cost function based on time evolution, and then solve the optimization problems subject to certain constraints. The control signals are the injection rates of drugs[124]. Usually, the objective functions are relevant to drug cost and patients tolerance during the treatment. As for chemotherapy model, the control problem can be to minimize cancerous cell growth, the average tumor burden[125] and the summation of the kinetic energies of all the cancerous cells[124] by searching the best distribution of a cumulative drug administration. The objective function in chemotherapy can be composed of tumor cell decrement, the average level of toxicity (tolerable drug concentration) during the treatment[18]. In a two-drug treatment model, the objective is

the summation of the tumor volume and the doses of anti-angiogenic therapy and cytostatic drug[126]. In HIV, the goal is 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[100].

Similarly, to eliminate the spread of HIV-1 by increasing the number of the normal CD4+ T cells, reducing virions, and minimizing the cost of treatment, the objective function consists of beneficiary T Cells and the weighted systemic costs of the therapy[119]. For multi-drug treatment, the goal can be to identify the optimal ratio of co-administration between chemotherapy agents and anti-angiogenic agent[124]. Constraints can be the boundary for toxicity level and the certain therapeutic window for drug concentration[18]. The constraints can also be certain bounds for drug doses and the volume of blood vessels in the neighborhood of the tumor[126]. Algorithms have been proposed for the constrained optimization problems. For instance, in a multi-objective genetic algorithm, the Pareto optimal sets have been used to search for lower values for 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[126]. The steepest descent method can be used to search numerical solutions iteratively to minimize cost function[124]. The numerical solutions can be derived with Miser3/Matlab[102]. To realize precision medicine, the objective functions should be evaluated for the individuals, and the solutions should be checked by the rank of the degree of dominance[18]. Model Predictive Control (MPC) can handle the mismatch between nominal and actual processes[127]. Its essence is the ability to handle constrained optimization problems[128–130]. MPC solves a finite horizon open-loop optimal control problem to obtain control actions with predicted states from models during each sampling period[129]. Similar to optimal control, objective functions in MPC mainly consist of disease-related molecules and drug infusion[131]. Control signals are formulated to minimize the difference between model values and the predefined reference trajectories (that can be learned based on a comparison of measurement to the desired values[130]). The local asymptotic stability of the control law is guaranteed when giving sufficiently long time horizon[128]. The weight of each term in the objectives reflects the focus of the optimization problem. The control law calculated by minimizing the cost can be piecewise constants or linear[128] in the finite time horizon. For multiple control inputs (e.g., chemotherapy and immunotherapy), multiple MPC can be applied to design the optimal combination ratio with lower drug toxicity[132]. The loop of model predictive control is shown in Fig.6b. Notably, discrete drug therapy should be considered to allow the normal cells to rebuild in clinical treatment[119]. Besides, controllers can also be "learned" from complex situations by applying learning-based strategies, such that optimal parameterized control signals can be derived[133]. If the equilibrium exists in the chemotherapy model, it should be the case that the tumor is eliminated or cancer cell concentration is at a safe level, so that no more treatment is needed for recovery[124]. With the estimated parameters in dynamic models, various control laws can be designed for drug administration to suppress disease progression. Optimization has been widely used for therapeutic design since it handles patients' physical constraints and the efficacy of the treatment at the same time. Similar to the optimization problem in parameter estimation, the time cost of getting an optimal solution depends on the solvers' convergence. The solvers with higher efficiency are always desired. Also, to design personalized treatments, elements relevant to the variability among patients may be added to the objective functions.

The static network identifies potential elements to update the dynamic models for accurate disease mechanisms and drug responses. With experimental and clinic data, numerical parameters are fitted in dynamic models. Next, the drug infusion rate for better therapeutic performance is designed by control laws.

## Outlook

Researchers can identify potential molecules or interactions in the existing biological networks with interactive patterns and make predictions for the incoming molecules. Since these molecules do not work individually, the functional clusters may be identified to account for organisms' particular functions. With a larger map of the biological process in organisms, the potential molecules or connections in omics and drug interactions can be identified for disease mechanisms and drug intervention. These existing models could have novel components to describe the same biological mechanisms in dynamic modeling. This makes the modeling more comprehensive and close to the real phenomena. The resulting models may contain the elements (from omics levels) to differentiate drug responses between individuals. In that case, it may enable personalized medicine that generalizes the models to serve patients with different health conditions. Besides, with the identified disease regulators through the molecular interactions and dynamic pathways, we may only use the drug molecules that bind these particular targets, which indicates that the key regulators should be the only targets for novel therapy design when desiring fewer drugs. On the one hand, with the detected off-targets, the side-effect can be minimized by selecting drug molecules properly according to the potential DTIs. For individuals, the identified off-targets' binding properties can be used to assess the side-effect of the drug molecules for all the available drugs. Then drug molecules can be combined by avoiding these off-targets (only if no DDI will be detected for side-effects within these drugs). Besides, the severity and frequency of the side effects should also be taken into consideration[79]. In this way, the safe drug combinations would be feasible for reducing the side-effect events. On the other hand, side effects should be avoided for safety issues, while this does not mean that all drugs that cause off-target effects will be abandoned. The treatments such as chemotherapy harm normal cells, so we may select cancer drugs that cause less damage

according to the identified DTI and combine with the advanced control strategies to get higher treatment performance but lower impairment. Besides, data sets in hand are always low-throughput[134] with high dimensionality. So, a good capability of noise resistance is needed for the feature extraction algorithms. The changing nature of omics data[8] may involve the irrelevant molecules in the models, which reduces the robustness of the prediction. An advanced algorithm, e.g., the learning-based algorithm, is desired to solve these constraints. Moreover, multi-constraints in systems biology make optimization an NP-hard problem[117], which exists in parameter estimation and the constrained problem. A solver that avoids getting stuck in local optima and handles the constraints simultaneously is necessitated.

## References


1. Yang, K., Bai, H., Ouyang, Q., Lai, L. & Tang, C. Finding multiple target optimal intervention in disease-related molecular network. *Mol. systems biology* **4**, 228 (2008).
2. Dickson, M. & Gagnon, J. P. Key factors in the rising cost of new drug discovery and development. *Nat. reviews Drug discovery* **3**, 417–429 (2004).
3. Lo, Y.-C., Rensi, S. E., Torng, W. & Altman, R. B. Machine learning in chemoinformatics and drug discovery. *Drug discovery today* **23**, 1538–1546 (2018).
4. Tanne, J. H. Pfizer stops clinical trials of heart drug (2006).
5. Forrest, M. J. *et al.* Torcetrapib-induced blood pressure elevation is independent of cetp inhibition and is accompanied by increased circulating levels of aldosterone. *Br. journal pharmacology* **154**, 1465–1473 (2008).
6. Liebler, D. C. & Guengerich, F. P. Elucidating mechanisms of drug-induced toxicity. *Nat. reviews Drug discovery* **4**, 410–420 (2005).
7. Horn, F. *et al.* Systems biology of fungal infection. *Front. microbiology* **3**, 108 (2012).
8. Hasin, Y., Seldin, M. & Lusis, A. Multi-omics approaches to disease. *Genome biology* **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, The J. Am. Soc. Hematol.* **120**, 5231–5236 (2012).
10. Laslo, P., Pongubala, J. M., Lancki, D. W. & Singh, H. Gene regulatory networks directing myeloid and lymphoid cell fates within the immune system. In *Seminars in immunology*, vol. 20, 228–235 (Elsevier, 2008).
11. Schadt, E. E. *et al.* An integrative genomics approach to infer causal associations between gene expression and disease. *Nat. genetics* **37**, 710–717 (2005).
12. Wang, B. *et al.* Integrative omics approach to identifying genes associated with atrial fibrillation. *Circ. research* **126**, 350–360 (2020).
13. Li, J., Lu, Q. & Wen, Y. Multi-kernel linear mixed model with adaptive lasso for prediction analysis on high-dimensional multi-omics data. *Bioinformatics* **36**, 1785–1794 (2020).
14. Subramanian, N., Torabi-Parizi, P., Gottschalk, R. A., Germain, R. N. & Dutta, B. Network representations of immune system complexity. *Wiley Interdiscip. Rev. Syst. Biol. Medicine* **7**, 13–38 (2015).
15. Kitano, H. Computational systems biology. *Nature* **420**, 206–210 (2002).
16. AY, M. *et al.* Drug–target network. *Nat. biotechnology* **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. & biological engineering & computing* **49**, 51–65 (2011).
19. Taylor, I. W. *et al.* Dynamic modularity in protein interaction networks predicts breast cancer outcome. *Nat. biotechnology* **27**, 199–204 (2009).
20. Zhao, S. & Iyengar, R. Systems pharmacology: network analysis to identify multiscale mechanisms of drug action. *Annu. review pharmacology toxicology* **52**, 505–521 (2012).
21. Sharan, R. *et al.* Conserved patterns of protein interaction in multiple species. *Proc. Natl. Acad. Sci.* **102**, 1974–1979 (2005).
22. Durmuş, S., Çakır, T., Özgür, A. & Guthke, R. A review on computational systems biology of pathogen–host interactions. *Front. microbiology* **6**, 235 (2015).

**23.** 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).

**24.** Albert, R. Network inference, analysis, and modeling in systems biology. *The Plant Cell* **19**, 3327–3338 (2007).

**25.** Langhauser, F. *et al.* A diseasome cluster-based drug repurposing of soluble guanylate cyclase activators from smooth muscle relaxation to direct neuroprotection. *NPJ systems biology applications* **4**, 1–13 (2018).

**26.** Menche, J. *et al.* Uncovering disease-disease relationships through the incomplete interactome. *Science* **347** (2015).

**27.** 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).

**28.** Goh, K.-I. *et al.* The human disease network. *Proc. Natl. Acad. Sci.* **104**, 8685–8690 (2007).

**29.** Husmeier, D. Sensitivity and specificity of inferring genetic regulatory interactions from microarray experiments with dynamic bayesian networks. *Bioinformatics* **19**, 2271–2282 (2003).

**30.** 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 & Rheum.* **60**, 2585–2594 (2009).

**31.** 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).

**32.** Keiser, M. J. *et al.* Predicting new molecular targets for known drugs. *Nature* **462**, 175–181 (2009).

**33.** Wang, E., Lenferink, A. & O'Connor-McCourt, M. Cancer systems biology: exploring cancer-associated genes on cellular networks. *arXiv preprint arXiv:0712.3753* (2007).

**34.** Ma, X. *et al.* Revealing pathway dynamics in heart diseases by analyzing multiple differential networks. *PLoS Comput. Biol* **11**, e1004332 (2015).

**35.** Bentele, M. *et al.* Mathematical modeling reveals threshold mechanism in cd95-induced apoptosis. *The J. cell biology* **166**, 839–851 (2004).

**36.** 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).

**37.** Lamb, J. *et al.* The connectivity map: using gene-expression signatures to connect small molecules, genes, and disease. *science* **313**, 1929–1935 (2006).

**38.** Chicco, D., Sadowski, P. & Baldi, P. Deep autoencoder neural networks for gene ontology annotation predictions. In *Proceedings of the 5th ACM conference on bioinformatics, computational biology, and health informatics*, 533–540 (2014).

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

**40.** 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).

**41.** 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. genetics* **44**, 841–847 (2012).

**42.** 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).

**43.** Roth, W., Hecker, D. & Fava, E. Systems biology approaches to the study of biological networks underlying alzheimer's disease: role of mirnas. In *Systems Biology of Alzheimer's Disease*, 349–377 (Springer, 2016).

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

**45.** 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).

**46.** Dutta, A. Stabilizing covid-19 infections in us by feedback control based test and quarantine. In *2020 IEEE Global Humanitarian Technology Conference (GHTC)*, 1–6 (IEEE, 2020).

**47.** 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. Medicine* **18**, 1–10 (2020).

**48.** Zhavoronkov, A., Buzdin, A. A., Garazha, A. V., Borissoff, N. & Moskalev, A. A. Signaling pathway cloud regulation for in silico screening and ranking of the potential geroprotective drugs. *Front. genetics* **5**, 49 (2014).

**49.** Smyth, G. K. Limma: linear models for microarray data. In *Bioinformatics and computational biology solutions using R and Bioconductor*, 397–420 (Springer, 2005).

**50.** Quan, Z. *et al.* Protein-protein interaction network and mechanism analysis in ischemic stroke. *Mol. medicine reports* **11**, 29–36 (2015).

**51.** Zhang, B. & Horvath, S. A general framework for weighted gene co-expression network analysis. *Stat. applications genetics molecular biology* **4** (2005).

**52.** Langfelder, P. & Horvath, S. Wgcna: an r package for weighted correlation network analysis. *BMC bioinformatics* **9**, 1–13 (2008).

**53.** Karp, P. D. *et al.* Pathway tools version 19.0 update: software for pathway/genome informatics and systems biology. *Briefings bioinformatics* **17**, 877–890 (2016).

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

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

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

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

**58.** Yamanishi, Y., Araki, M., Gutteridge, A., Honda, W. & Kanehisa, M. Prediction of drug–target interaction networks from the integration of chemical and genomic spaces. *Bioinformatics* **24**, i232–i240 (2008).

**59.** 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. Informatics Assoc.* **19**, 79–85 (2012).

**60.** Liu, S., Tang, B., Chen, Q. & Wang, X. Drug-drug interaction extraction via convolutional neural networks. *Comput. mathematical methods medicine* **2016** (2016).

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

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

**63.** 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).

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

**65.** Fakhraei, S., Huang, B., Raschid, L. & Getoor, L. Network-based drug-target interaction prediction with probabilistic soft logic. *IEEE/ACM Transactions on Comput. Biol. Bioinforma.* **11**, 775–787 (2014).

**66.** 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. pharmaceutics* **14**, 3098–3104 (2017).

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

**68.** 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).

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

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

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

**73.** Wang, R.-S., Zhang, S., Wang, Y., Zhang, X.-S. & Chen, L. Clustering complex networks and biological networks by nonnegative matrix factorization with various similarity measures. *Neurocomputing* **72**, 134–141 (2008).

**74.** Navlakha, S. & Kingsford, C. The power of protein interaction networks for associating genes with diseases. *Bioinformatics* **26**, 1057–1063 (2010).

**75.** Kyrtsos, C. R. & Baras, J. S. A systems biology model studying the role of cholesterol in alzheimer's disease progression. In *2011 10th International Workshop on Biomedical Engineering*, 1–7 (IEEE, 2011).

**76.** Chen, W. W. *et al.* Input–output behavior of erbb signaling pathways as revealed by a mass action model trained against dynamic data. *Mol. systems biology* **5**, 239 (2009).

**77.** Cheng, F. & Zhao, Z. Machine learning-based prediction of drug–drug interactions by integrating drug phenotypic, therapeutic, chemical, and genomic properties. *J. Am. Med. Informatics Assoc.* **21**, e278–e286 (2014).

**78.** Cheng, F. *et al.* Prediction of drug-target interactions and drug repositioning via network-based inference. *PLoS Comput. Biol* **8**, e1002503 (2012).

**79.** Ye, H., Liu, Q. & Wei, J. Construction of drug network based on side effects and its application for drug repositioning. *PloS one* **9**, e87864 (2014).

**80.** Ding, Y., Tang, J. & Guo, F. Identification of drug-side effect association via multiple information integration with centered kernel alignment. *Neurocomputing* **325**, 211–224 (2019).

**81.** Huang, S., Chaudhary, K. & Garmire, L. X. More is better: recent progress in multi-omics data integration methods. *Front. genetics* **8**, 84 (2017).

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

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

**84.** Bleakley, K. & Yamanishi, Y. Supervised prediction of drug–target interactions using bipartite local models. *Bioinformatics* **25**, 2397–2403 (2009).

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

**86.** Stepniewska-Dziubinska, M. M., Zielenkiewicz, P. & Siedlecki, P. Development and evaluation of a deep learning model for protein–ligand binding affinity prediction. *Bioinformatics* **34**, 3666–3674 (2018).

**87.** Gawehn, E., Hiss, J. A., Brown, J. B. & Schneider, G. Advancing drug discovery via gpu-based deep learning (2018).

**88.** Duvenaud, D. K. *et al.* Convolutional networks on graphs for learning molecular fingerprints. In *Advances in neural information processing systems*, 2224–2232 (2015).

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

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

**91.** Chen, X. *et al.* Bnpmda: bipartite network projection for mirna–disease association prediction. *Bioinformatics* **34**, 3178–3186 (2018).

**92.** Miura, M. *et al.* Verification of a cocktail approach for quantitative drug–drug interaction assessment: a comparative analysis between the results of a single drug and a cocktail drug. *Xenobiotica* 1–29 (2020).

**93.** Villoslada, P., Steinman, L. & Baranzini, S. E. Systems biology and its application to the understanding of neurological diseases. *Annals Neurol. Off. J. Am. Neurol. Assoc. Child Neurol. Soc.* **65**, 124–139 (2009).

**94.** Grechkin, M., Logsdon, B. A., Gentles, A. J. & Lee, S.-I. Identifying network perturbation in cancer. *PLoS computational biology* **12**, e1004888 (2016).

**95.** 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).

**96.** Lee, M. N. *et al.* Identification of regulators of the innate immune response to cytosolic dna and retroviral infection by an integrative approach. *Nat. immunology* **14**, 179–185 (2013).

**97.** Sever, R. & Brugge, J. S. Signal transduction in cancer. *Cold Spring Harb. perspectives medicine* **5**, a006098 (2015).

**98.** Cho, K.-H. & Wolkenhauer, O. Analysis and modelling of signal transduction pathways in systems biology (2003).

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

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

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

**102.** de 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).

**103.** Zi, Z. & Klipp, E. Sbml-pet: a systems biology markup language-based parameter estimation tool. *Bioinformatics* **22**, 2704–2705 (2006).

**104.** Aster, R. C., Borchers, B. & Thurber, C. H. *Parameter estimation and inverse problems* (Elsevier, 2018).

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

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

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

**108.** 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).

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

**110.** Pirgazi, J. & Khanteymoori, A. R. A robust gene regulatory network inference method base on kalman filter and linear regression. *PloS one* **13**, e0200094 (2018).

**111.** Zeng, N., Wang, Z., Li, Y., Du, M. & Liu, X. A hybrid ekf and switching pso algorithm for joint state and parameter estimation of lateral flow immunoassay models. *IEEE/ACM Transactions on Comput. Biol. Bioinforma.* **9**, 321–329 (2011).

**112.** Medvedev, A., Zhusubaliyev, Z. T., Rosén, O. & Silva, M. M. Oscillations-free pid control of anesthetic drug delivery in neuromuscular blockade. *Comput. methods programs biomedicine* **171**, 119–131 (2019).

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

**114.** 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. Heal. Monit. The Off. J. Int. Assoc. for Struct. Control. Monit. Eur. Assoc. for Control. Struct.* **16**, 99–123 (2009).

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

**116.** Raue, A. *et al.* Data2dynamics: a modeling environment tailored to parameter estimation in dynamical systems. *Bioinformatics* **31**, 3558–3560 (2015).

**117.** Banga, J. R. Optimization in computational systems biology. *BMC systems biology* **2**, 1–7 (2008).

**118.** Mendonça, T., Lemos, J. M., Magalhaes, H., Rocha, P. & Esteves, S. Drug delivery for neuromuscular blockade with supervised multimodel adaptive control. *IEEE Transactions on Control. systems technology* **17**, 1237–1244 (2009).

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

**120.** Hägglund, T. *PID controllers: theory, design, and tuning*, vol. 2.

**121.** Tan, W., Liu, J., Chen, T. & Marquez, H. J. Comparison of some well-known pid tuning formulas. *Comput. & chemical engineering* **30**, 1416–1423 (2006).

**122.** De Keyser, R., Dutta, A., Hernandez, A. & Ionescu, C. M. A specifications based pid autotuner. In *2012 IEEE International Conference on Control Applications*, 1621–1626 (IEEE, 2012).

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

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

**125.** 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 biology* **459**, 67–78 (2018).

**126.** 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).

**127.** Dutta, A. *et al.* Switched nonlinear predictive control with adaptive references for engagement of wet clutches. *IFAC Proc. Vol.* **45**, 460–465 (2012).

**128.** 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 biomedicine* **108**, 973–983 (2012).

**129.** Dutta, A., Ionescu, C., Loccufier, M. & De Keyser, R. Robust penalty adaptive model predictive control (pampc) of constrained, underdamped, noncollocated systems. *J. Vib. Control.* **22**, 549–558 (2016).

**130.** Dutta, A. *et al.* Comparison of two-level nmpc and ilc strategies for wet-clutch control. *Control. Eng. Pract.* **22**, 114–124 (2014).

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

**132.** Sharifi, N., Ozgoli, S. & Ramezani, A. Multiple model predictive control for optimal drug administration of mixed immunotherapy and chemotherapy of tumours. *Comput. methods programs biomedicine* **144**, 13–19 (2017).

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

**134.** Tyanova, S. *et al.* The perseus computational platform for comprehensive analysis of (prote) omics data. *Nat. methods* **13**, 731 (2016).

**Acknowledgements**
We thank Jianghua Wu, Hezi Zhao and the UConn Writing center for grammatical corrections. We acknowledge the support from the UConn School of Engineering.

**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.

# Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- LaTeXsupplementaryfile.bbl