



**CHALMERS**  
UNIVERSITY OF TECHNOLOGY

## **Towards an interpretable deep learning model of cancer**

Downloaded from: <https://research.chalmers.se>, 2025-04-03 00:47 UTC

Citation for the original published paper (version of record):

Nilsson, A., Meimetis, N., Lauffenburger, D. (2025). Towards an interpretable deep learning model of cancer. *npj Precision Oncology*, 9(1). <http://dx.doi.org/10.1038/s41698-025-00822-y>

N.B. When citing this work, cite the original published paper.

<https://doi.org/10.1038/s41698-025-00822-y>

# Towards an interpretable deep learning model of cancer

Check for updates

Avlant Nilsson<sup>1,2,3</sup>, Nikolaos Meimetis<sup>1</sup> & Douglas A. Lauffenburger<sup>1</sup> ✉

Cancer is a manifestation of dysfunctional cell states. It emerges from an interplay of intrinsic and extrinsic factors that disrupt cellular dynamics, including genetic and epigenetic alterations, as well as the tumor microenvironment. This complexity can make it challenging to infer molecular causes for treating the disease. This may be addressed by system-wide computer models of cells, as they allow rapid generation and testing of hypotheses that would be too slow or impossible to perform in the laboratory and clinic. However, so far, such models have been impeded by both experimental and computational limitations. In this perspective, we argue that they can now be achieved using deep learning algorithms to integrate omics data and prior knowledge of molecular networks. Such models would have many applications in precision oncology, e.g., for identifying drug targets and biomarkers, predicting resistance mechanisms and toxicity effects of drugs, or simulating cell-cell interactions in the microenvironment.

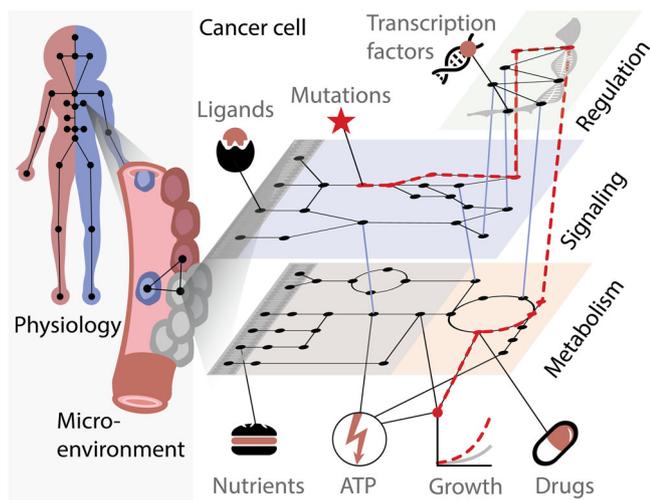
Cancer is a diverse neoplastic disease, where mutations and other alterations drive phenotypes such as sustained cellular proliferation, and evasion of the immune system<sup>1</sup>. Safe and effective cancer treatments could in principle be attained by targeting these specific deviations, i.e., precision medicine<sup>2</sup>. However, in practice, this is challenging because cellular processes are highly interconnected and furthermore depend on the microenvironment (Fig. 1). This complexity hampers our ability to establish causal relations between genetic alterations and disease phenotypes, such as which mutations are driving the disease and which are passengers<sup>3</sup>. It has been found that inhibiting cancer-specific signaling proteins can be highly effective<sup>4</sup>, but such treatments can also be thwarted by negative feedback loops, e.g., an ERK-dependent feedback loop that attenuates the effects of RAF inhibition<sup>5</sup>. Even if a treatment is successful at first, there may emerge cells that are resistant to treatment due to cellular heterogeneity, which may originate from both genomic and epigenetic differences<sup>6</sup>. Furthermore, cancer cells interact with other cells in their microenvironment, which expands their signaling repertoire<sup>7</sup>. Overall, this interplay between mutations, drugs, resistance mechanisms, feedback loops, and environmental factors, gives rise to a combinatorial number of possible causes, which can be challenging to investigate experimentally. A computation approach would therefore be advantageous.

Living cells are dynamical systems that depend on biochemical interactions between a vast number of molecules, which are governed by physico-chemical laws of kinetics and transport. Such systems can be described using mathematical models, and for this, models based on

ordinary differential equations (ODEs) have commonly been used. They can be applied to model interactions between proteins and metabolites, enable predictions of disease-associated molecules, and provide mechanistic explanations of perturbations<sup>8–11</sup>. In principle, if all molecular relations were known and stochastic effects were disregarded, the activity of a cell could be derived from its initial condition, i.e., bottom-up modeling<sup>12</sup>. However, this remains severely daunting for both practical and theoretical reasons<sup>13</sup>, and an appealing alternative is to fit model parameters to systems-level data, i.e., top-down modeling.

Machine learning (ML) and specifically deep learning (DL) models are now becoming broadly employed in biology and medicine. These algorithms are trained, i.e., parameterized, on data to make rapid predictions that generalize to unseen conditions. Together with large-scale high throughput screening datasets<sup>14–16</sup>, these models have been successfully used in a wide variety of tasks<sup>17</sup>, including predicting drug synergy<sup>18</sup>, response to therapy<sup>19</sup>, survival probabilities<sup>20</sup>, disease outcomes<sup>21</sup>, and for cancer histology<sup>22</sup>. Some models include molecular structure in their description of the input, which allows them to extrapolate their predictions to untested molecules<sup>17,23</sup>. Many of these models predict phenotypes of interest without accounting for the underlying mechanisms, and they have therefore been criticized for their lack of biological interpretability<sup>24</sup>. Recently, prior biological knowledge has been integrated with the ML models, e.g., in the form of molecular networks, thereby making them more structured and interpretable, and has shown promising results at both the patient- and cellular level<sup>25–28</sup>.

<sup>1</sup>Department of Biological Engineering, Massachusetts Institute of Technology, Cambridge, MA, USA. <sup>2</sup>Department of Biology and Biological Engineering, Chalmers University of Technology, Göteborg, Sweden. <sup>3</sup>Department of Cell and Molecular Biology, SciLifeLab, Karolinska Institutet, Stockholm, Sweden. ✉e-mail: [lauffen@mit.edu](mailto:lauffen@mit.edu)



**Fig. 1 | Interconnected molecular networks in human cells.** Effects of a mutation in a signaling protein can propagate through gene regulation, and metabolism, to affect cellular growth.

The utility of network-based approaches for studying cancer has long been appreciated. These have been used to map the cell's functional organization<sup>29</sup>, for causal integration and mechanistic hypotheses-generation on cancer data<sup>30</sup>, and to characterize the response to drug therapies and their mechanism of action<sup>31</sup>. A plethora of network-centric tools has been created for interrogating the biology of diseases and the mechanism of action of therapeutic solutions<sup>32</sup>. For example, in Cytoscape researchers may integrate large molecular networks and omics data, to perform network statistics and visualization in a fast and simple way<sup>33</sup>. Some approaches, such as CausalR<sup>34</sup>, CARNIVAL<sup>35</sup>, and COSMOS<sup>30</sup>, utilize omics data and prior knowledge to propose causal networks that underlay the observed condition. Additionally, some approaches simultaneously infer gene regulatory networks from single-cell RNA sequencing data and use them to perform *in silico* perturbations, thus interrogating the potential effect of the perturbation of specific molecular species<sup>36,37</sup>. However, most network-based approaches are primarily descriptive; i.e., they are not constructed to make quantitative predictions about unseen conditions, such as combinations of mutations that have not previously been observed. This may therefore require manual interpretation. Furthermore, they do not generally encode the structure of the molecules that induce the observed biological effect, and therefore cannot directly predict the effects of previously untested molecules.

Because cellular processes are highly interconnected in cancer, system-level models are required to predict the effects of mutations and other alterations. In this perspective, we discuss the feasibility of constructing such models in light of recent progress in DL algorithms, and data acquisition methods, along with the accumulated prior knowledge of molecular networks. We also explore potential applications if such a model is successfully implemented.

### Predictive deep learning models

ML, and in particular artificial neural networks (ANNs) facilitate the construction of large-scale predictive models. ANNs approximate unknown, complex functions through a sequence of linear matrix operations and non-linear transformations, with layers of latent variables between them. The models are referred to as 'deep' if multiple layers are used. These models can contain millions of parameters and can rapidly be fitted to paired samples of input- and output data using automatic differentiation. The flexibility and scale of DL make it a promising candidate for fitting models of complex and heterogeneous molecular data. For cancer specifically, DL has been used to predict response to therapy<sup>19</sup>, and tumor phenotypes after perturbations with high performance<sup>38</sup>, and other ML models have been used to predict

patient response using transfer learning approaches from pre-clinical models<sup>39</sup>.

However, while DL models excel at predictions, their relation to the underlying mechanisms that they approximate is often opaque, i.e., the black box problem<sup>25</sup>. It has therefore been proposed that more interpretable DL models should be used for biological systems, and explainable AI has indeed gained a lot of interest over the last years<sup>40,41</sup>. For example, Keyl et al. used an interpretation method termed layer-wise relevance propagation<sup>40</sup>, coupled with an ANN model for predicting protein interaction networks from proteomic data for individual patients<sup>42</sup>. Laurie and Lu developed a DL architecture for survival prediction, whose components capture specific aspects of tumor dynamics<sup>43</sup>.

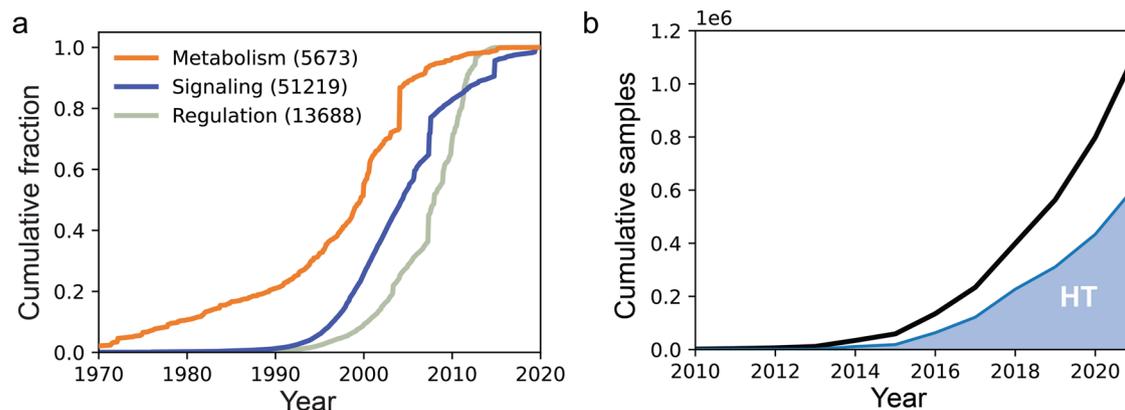
We reason that, while purely predictive models can be useful for many tasks, such as prioritizing perturbations for further experimental validation, the development of safe and effective therapeutics, should be rooted in a purposeful process. Particularly, they should be designed to affect specific targets, with known mechanisms of action, and well-characterized dynamics. This may be achieved by a model of the human cell. According to a survey of the biomodelling community<sup>44</sup>, such a model should, at a minimum, cover signal transduction, metabolism, and gene regulation, including transcription, translation, and degradation processes. While this survey pertained to a framework that mixed different model-types, an integrated model using DL would be advantageous for automatizing parameterization using different types of experimental data.

DL models have now been developed using prior knowledge for each of the major cellular subsystems. For signaling, we developed a model that predicts transcription factor (TF) activities or cell viability from ligand- or drug stimulation<sup>27</sup>. This model simulates signal propagation using a recurrent neural network (RNN) with a prior knowledge signaling network as a scaffold. We recently expanded this approach to simulate cancer cell signaling under perturbations with small molecule drugs, while simultaneously inferring their off-target effects<sup>28</sup>. For metabolism, a model has been developed that predicts metabolic rates from metabolite concentrations for *E. coli* bacteria<sup>45</sup>. Finally, for gene regulation, a model has been developed that reproduces the chemistry of TF-DNA binding and predicts gene expression levels from TF concentrations<sup>46</sup>. Using similar formulations, an integrated model of these processes could be reconstructed.

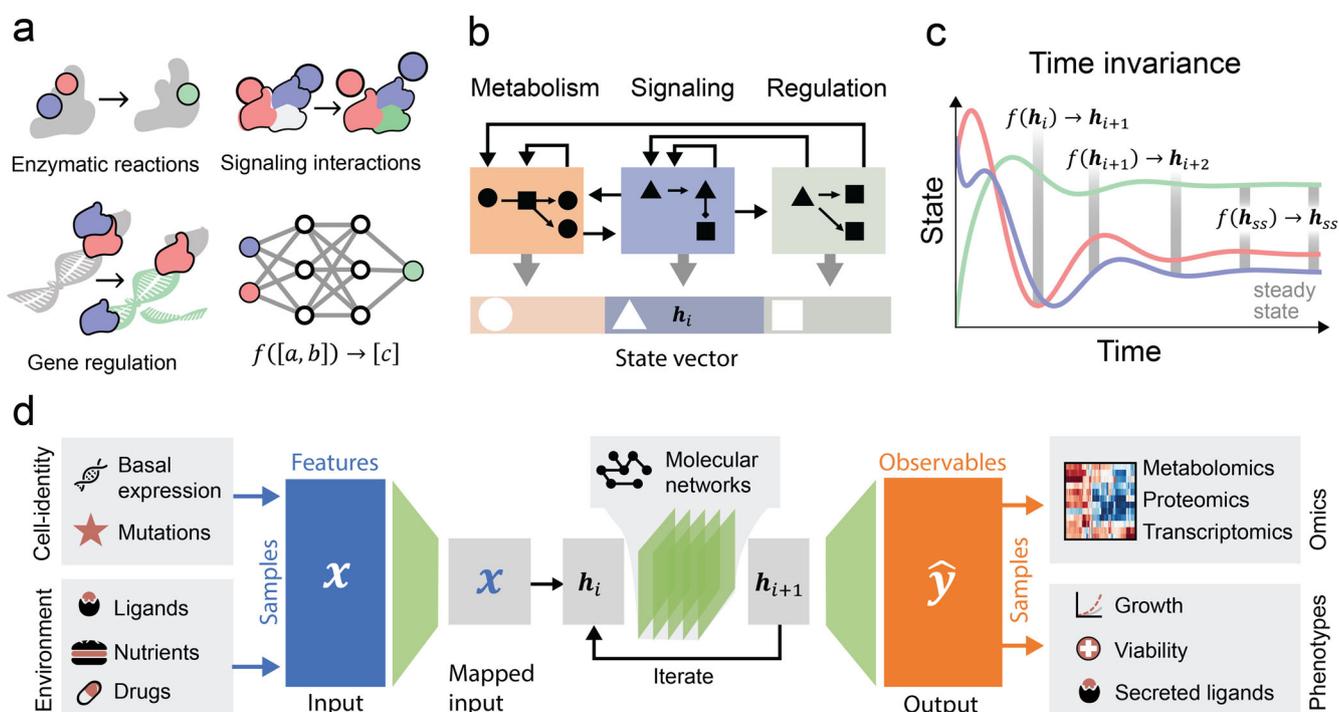
### Saturating knowledge and accumulating data

Our accumulated knowledge of molecular networks has now reached the genome-scale, thus enabling the creation of genome-wide predictive models. Advancements in both experimental<sup>47–50</sup> and computational<sup>151–54</sup> techniques have resulted in the curation of prior knowledge networks with thousand molecular interactions<sup>55,56</sup>. Prior knowledge of molecular relations is in principle finite since it is limited to specifying which interactions can take place if the participating molecules are present at appropriate concentrations. These relations place a structural constraint on which cell states are possible, although they are not always in effect, since not all molecules are present in every cell and cellular condition. Using this type of knowledge, large networks have been reconstructed for metabolism, signal transduction, and gene regulation<sup>55–58</sup>. An analysis of the first publication date for the references behind each interaction suggests that the rate of new discoveries is slowing (Fig. 2a). One interpretation may be that our prior knowledge of molecular interactions is reaching completion, although other possible interpretations may include a shift in research interests from basic biochemistry, or a lag between discovery and addition to databases. In particular, for the gene-regulatory network, the low coverage of interactions with literature support (8%, corresponding to less than one interaction per gene) may reflect a shift in the valuation of evidence, from individual published studies to inclusion in public databases.

Simultaneously, the amount of well-annotated high-throughput data continues to increase rapidly. It can in principle increase infinitely, as it quantifies the cell state or phenotype under particular experimental conditions. To successfully model these data, both the experimental design (or metadata) and the cellular responses should ideally be recorded in databases.



**Fig. 2 | Saturating knowledge and accumulating data.** **a** Prior knowledge by date of discovery. Interactions for metabolism (biochemical reactions and transport in Human1<sup>79</sup>), signaling (protein-protein interactions in OmniPath<sup>56</sup>), and gene regulation (transcription factor-gene interactions in Dorothea<sup>58</sup>) by publication date of the oldest reference in the database. The total number of interactions with at least one PubMed reference is given in parenthesis and covers 50%, 98%, and 8% of the total interactions respectively. **b** Human transcriptomics samples in Gene Expression Omnibus (GEO) using RNAseq technology. A majority are from studies that include more than 250 samples, here labeled as high throughput (HT).



**Fig. 3 | Model properties.** **a** Neural networks can approximate input-output relations between biomolecules such as metabolite concentrations, phosphorylation states of proteins, and transcription factor activities. **b** Integrated processes, the state vector partitioned into classes such as metabolite concentrations, signaling states, and protein concentrations. Each class is updated by independent functions that interact through their dependencies on the shared state space. **c** State trajectories constructed from repeated applications of a function on the state. **d** Inputs ( $x$ ) describing the samples' condition are mapped into their corresponding state elements. Analogously, outputs ( $\hat{y}$ ) are extracted from the state to omics data and phenotypes. Prior knowledge of molecular networks is encoded into the architecture that updates the state vector.

One such database, the Gene Expression Omnibus (GEO) database<sup>59</sup>, has now surpassed a million human samples (Fig. 2b). There are public datasets<sup>14–16,60–62</sup>, that cover numerous perturbations such as gene knock-outs, and stimulation by different drugs and ligands, in thousands of cell-lines, e.g., the CLUE platform hosts transcriptome profiles for more than 780,000 unique conditions<sup>16</sup>. These offer a vast amount of data to study and model the transcriptomic profile of diseased and perturbed samples in bulk, or even at the single-cell level.

Analyzing samples at the single-cell level, via molecular barcoding (e.g., nucleotide barcoding), has enabled the characterization of groups of cells in

heterogeneous populations, which has allowed for a detailed characterization of cell types in different organs<sup>63</sup>. This can be particularly useful for clinical tumor samples, where the mixture of different cell types and their individual states would be challenging to deconvolute from bulk measurements<sup>64</sup>. After identifying individual cell types in a sample, it is possible to perfume pseudo-bulking, an analysis where gene expression data from cells derived from the same cell type is aggregated into distinct pseudo-bulk samples, reducing the impact of technical variability and dropout effects, often present single-cell RNA sequence data. Single-cell analysis has also allowed the exploration of variability in seemingly homogeneous cell

populations<sup>65</sup>. Finally, the developments in molecular barcoding have also enabled systematic screening of combinations of stimuli, e.g., the transcriptomic response to 420 different drug combinations was measured in a pooled experiment, by using unique barcodes for each drug<sup>66</sup>.

Using different experimental techniques, molecular quantities of different modalities are now routinely characterized at the genome-scale<sup>67</sup>, including metabolite concentrations (metabolomics), concentrations of mRNA transcripts (transcriptomics), protein concentrations (proteomics), and protein phosphorylation levels (phospho-proteomics), which relates to their signaling state. It is also increasingly common to quantify multiple data types for the same group of cells or subjects (multi-omics)<sup>67</sup> and, alternatively, to study the same set of cell lines with different techniques in different studies, e.g., one study-related metabolic profiles to differences in growth rate<sup>61</sup>, among the 60 cancer cell lines in the NCI60 panel, while another quantified their signaling responses to different drugs using phospho-proteomics<sup>62</sup>. However, so far it has proven challenging to integrate data from different studies, data modalities, and conditions<sup>67</sup>. This is in part due to a lack of unified analysis frameworks and because of difficulties in handling samples and subjects with missing data.

### Proposed structure for an integrated model

The activity of a cell consists of a series of molecular interactions that alter its molecular composition. To model these processes, the interactions could be represented by subfunctions that approximate the input-output relations between the molecular quantities involved (Fig. 3a). These interactions can be broadly classified as belonging to a particular cellular subsystem such as metabolism, signaling, or gene regulation that interact through shared molecular quantities. A unified model may be constructed from a large number of such relations (Fig. 3b). For example, an enzyme can synthesize a metabolite that affects a signaling protein in control of a TF that regulates the expression of a gene. Such a modular structure with well-defined processes acting on state variables corresponding to observable molecules would ensure the model's interpretability.

As for many other physical systems, it can be assumed that these molecular functions do not directly depend on which cell type, or cellular compartment they occur in, i.e., that they are space invariant. It can be assumed that the *de facto* observed differences between cells and compartments originate from differences in molecular concentrations and post-translational modifications that can all be considered inputs to the functions. The use of invariances has been found useful in other neural network applications as it allows parameters to be shared between seemingly different conditions<sup>17</sup>, e.g., for image recognition, where the same convolution function is applied across all parts of an image. A potential challenge for this is the existence of cell-type-specific versions of molecules, such as proteins translated from genes with multiple splicing variants, i.e., isoforms. These may either be represented by independent functions or integrated into a unified function by using the isoform composition as an additional input. Another challenge may be ambient environmental factors such as temperature or pH that affect biochemical rates and thus the input-output relations, but in principle, these could also be used as input.

It can also be assumed that only molecular quantities, not the functions themselves, change over time, i.e., that they are time-invariant. This implies that molecular trajectories can be constructed from repeated application of the same functions (Fig. 3d), with different inputs at different times, starting with some initial condition and potential perturbations. Based on this assumption, time dynamics can be simulated using recurrent neural networks (RNNs), which are DL architectures that iteratively model sequential operations<sup>17</sup>, and which have as output intermediate hidden states of a model, that are used as input to calculate the next state. Black box RNNs have been used to recapitulate predictions by ODE models of signaling<sup>68</sup> and we have used an RNN to directly model signal propagation from ligands to transcription factors, under the assumption that they reach a (pseudo-) steady state<sup>27</sup>.

A unified model operating on defined molecular quantities allows for a straightforward mapping of inputs, and extraction of outputs. In this

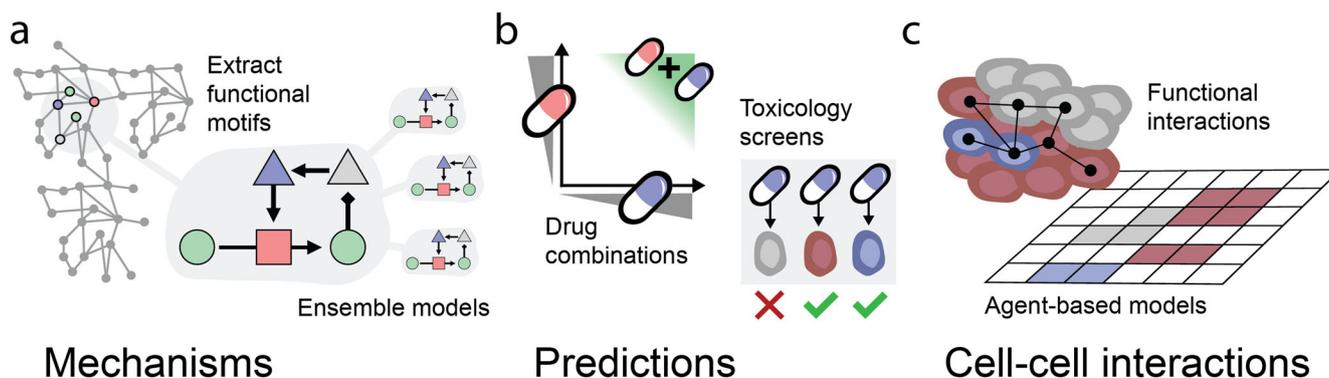
formulation, the cell state is expressed as vectors of molecular features, such as transcription factors, metabolites, proteins, etc. The inputs encode the cell type, environment, and perturbation of each specific sample (Fig. 3e). A mapping submodule can then assign the inputs to the affected molecules, for example drugs can be mapped to their targets, mutations to the affected proteins, and ligands to their receptors. The differences between cell types can be encoded through their basal molecular concentrations. Analogously, predictions pertaining to experimental data can be extracted from the state of the corresponding molecule, e.g., the expression of a gene from the representation of its mRNA concentration (Fig. 3e). The cell state is also a suitable input for submodules that predict cellular phenotypes, such as cell viability. For this, the state as a whole may be used or a knowledge-based subset, e.g., a core set of TFs that regulate cell proliferation has been used to predict viability<sup>9</sup>. This general representation would allow virtually any experimental conditions in any cell type to be integrated using the same model with different inputs. Because cellular processes are known to operate at different timescales, it would sometimes be warranted to separate slower processes into separate conditions, that can be simulated independently.

The purpose of the model is to predict unobserved cell states and the effects of untested perturbations for drug development. To train the model, the difference between expected and predicted values of the molecular quantities and phenotypes can be minimized using Mean Squared Error or the negative log-likelihood. This can be done for the final steady state of the model or alternatively across different iteration steps corresponding to integration across multiple time points. For single-cell data the model could be trained on pseudo-bulked profiles of each cell type, alternatively, the full distributions of gene expression in each specific cell type may be used a formulation inspired by evidential deep learning<sup>69</sup>, where the model estimates the statistical moments of an output distribution. To confirm the model's ability to extrapolate to unseen perturbations, cross-fold-validation should be utilized by holding out perturbations during training that are dissimilar from the ones used in each training fold. Finally, experimental validation is necessary to evaluate interesting results of perturbations and states, where no experimental data exists, and to confirm these findings.

In the long term, one could envision that the neural networks could also model the experimental setting so that much less preprocessing would be required. Currently, preprocessing is a challenge due to differences in sequencing depth (and dropout) as well as differences in experimental protocol, such as RNA extraction, proteomics methods, etc. The cells of individuals represent natural experiments that perturb the input to an idealized cell. Cancer corresponds to a subspace of these perturbations. The challenge that these models could address is that the circumstances of every specific cancer are slightly different, and the optimal treatment (such as a currently existing or putative drug) will be patient-dependent (i.e., the motivation for precision medicine). These models could act as a scaffold to unify data from different circumstances to predict which treatment has the potential to be effective in this particular case. Predictive models (such as neural networks) have the advantage over descriptive (statistical) models that they may extrapolate to make a prediction in a case that has previously not been observed. Although in practice, models may not be directly deployed in clinics, they may help provide biomarkers to stratify patients into different treatment regimes that can be clinically tested.

### Challenges and limitations

For models based on prior knowledge, the accuracy and completeness of the networks are of high importance. It can be expected that the prior knowledge both includes miss-annotated interactions and that it is incomplete. Critical examination and curation of interactions that are found to have high importance for the model predictions, in combination with validation experiments if the evidence is dubious, may be required to alleviate the limitation of miss-annotations in the prior knowledge. The incompleteness is expected to decrease as knowledge accumulates further, but this could also be addressed by including terms that model the influence of unknown factors<sup>70</sup>. If these terms significantly improve the model fit and generalization, they may constitute novel interactions to be experimentally



**Fig. 4 | Applications for an integrative model.** **a** Condition-specific functional motifs may be extracted from larger models. Ensembles of models may address model-model variability. **b** Synergistic effects of drug combinations may be predicted. Models of healthy cells could be used to anticipate toxicity effects. By modeling the effects of drugs in both healthy and diseased cells of different types,

toxicity effects can be simulated. This can help estimate which drug combinations and doses will be tolerated to select promising candidates for further experimental or clinical validation. **c** The functional exchange of metabolites or ligands between cells of different types may be required. Artificial neural network models may act as surrogates for rules in agent-based models of cell-cell interactions.

validated, although inference of protein-protein interactions would not be the main goal of this endeavor. Similarly, processes beyond the model's scope may manifest as missing interaction between molecular species, e.g., effects of splicing variants, microRNAs, or epigenetic modulation. An analysis of which conditions cause the model to fail and which measurements are affected may provide insight into the importance and magnitude of external influences.

Missing values and other irregularities often occur in biological data, and can be problematic for modeling. This may involve data collected from different cells at different time points, batch-batch variation between different groups of samples, or failed phenotype measurements<sup>67</sup>. Although, not apparent directly from their formulation, by applying a few modifications, DL models are well suited to handle these issues. Because ANNs estimate parameters using some version of stochastic gradient descent (SGD), missing data points can be omitted from the gradient calculations and allowed to take on any value. Furthermore, because SGD predicts gradients for the input variables, missing values can be automatically estimated to best agree with the experimental observations. In principle, batch effects could be accounted for by modeling differences in experimental equipment or other confounding variables. This type of “end-to-end” parameterization compared to manual feature selection, and standardization, simplifies the analysis and has proven useful in areas such as image processing<sup>17</sup>. Furthermore, semi-supervised learning approaches have been established to train models in cases where output data is not available for all of the samples<sup>71</sup>.

Discrete RNNs that may be used to simulate cellular dynamics are coarse-grained by design, which may pose a challenge for detailed simulations. While this limitation could in principle be circumvented using continuous time RNN architectures, these have hitherto relied on ODE solvers and thus suffer from issues with speed and scalability<sup>72</sup>. For some problems there exist closed-form approximations of the dynamics that can be used directly in the RNN, thereby achieving speedups by several orders of magnitudes<sup>72</sup>. Notably, a direct correspondence has been established<sup>73</sup> between an RNN with specific architecture and a common numerical ODE solver, thereby further reducing the gap between the two approaches.

### Applications and outlook

If successfully implemented and trained, these models would provide a succinct representation of cellular states and dynamics. This is expected to have many applications. The structured relation between state variables enables mechanistic interpretation of the modeling results (Fig. 4a). This may provide insight into emergent behavior and non-trivial dynamics, such as complex feedback structures that are not apparent in smaller-scale models<sup>13</sup>.

Models may also be used to predict responses to non-tested conditions and counterfactuals (Fig. 4b). This may have applications for drug discovery,

for example, drug resistance can be mitigated by drug combinations<sup>4</sup>, however, this is challenging to screen for experimentally as it grows exponentially for high-order drugs combinations. Another application may be to predict the effects of drugs on healthy cells. Drugs normally have off-target effects, and while this is sometimes important for their effectiveness<sup>74</sup>, it can affect healthy cells adversely and prevent the drugs' translation to clinical use.

Yet another application may be to study cell-cell interactions (Fig. 4c). It is common to interrogate this through the expression pattern of ligands and their receptors in different cell types<sup>75</sup>, and this principle could be advanced to predict the effects on downstream signaling targets and predict which ligands and metabolites are secreted. One way to simulate the effects of cell-cell interactions at scale is through agent-based models (ABM) that represent different cell types by rules that describe their input-output relations<sup>76</sup>. However, the rules used in ABM models (such as Boolean models of signaling) are typically manually encoded, which is laborious and biases the result. It has been shown that these rules can be substituted by DL models<sup>77</sup>, e.g., allowing an ANN to model the complete signaling structure which incorporates multiple potential logical rules, which could relieve the need for manually defining the interactions, thus allowing for rapid data-driven model construction.

After constructing models that can simulate cellular responses and cell-cell communication, -potentially with single-cell resolution, the next natural step is to make use of clinical data for patient-level predictions. For example, the proposed framework may be used to map genomic mutations and gene expression for individual cell types, via signaling, creating this way interactome-based DL models that can include the transcriptomic and genomic identity of cells as input. Then the model could leverage approaches where given the single-cell gene expression profile of individual patients and the identification of each cell type present, the patient response can be predicted<sup>78</sup>.

The aim of constructing integrative models is to enable end-to-end simulation of genotype-phenotype relations. Such a project may help bridge the areas of metabolism, signaling, and gene regulation that are currently mostly studied in isolation. Through simulations, models will allow rapid generation and testing of hypotheses that would be impractical or impossible in the lab and clinical setting. This is anticipated to have many applications, for identifying drug targets, biomarkers, and treatment options.

### Data availability

No datasets were generated or analysed during the current study.

Received: 26 September 2024; Accepted: 27 January 2025;

Published online: 14 February 2025

## References

- Hanahan, D. Hallmarks of cancer: new dimensions. *Cancer Discov.* **12**, 31–46 (2022).
- Collins, F. S. & Varmus, H. A new initiative on precision medicine. *N. Engl. J. Med.* **372**, 793–795 (2015).
- Brown, A. L., Li, M., Goncarenko, A. & Panchenko, A. R. Finding driver mutations in cancer: elucidating the role of background mutational processes. *PLoS Comput. Biol.* **15**, e1006981 (2019).
- Zhong, L. et al. Small molecules in targeted cancer therapy: advances, challenges, and future perspectives. *Signal Transduct. Target. Ther.* **6**, 201 (2021).
- Kholodenko, B. N. et al. A systematic analysis of signaling reactivation and drug resistance. *Cell Rep.* **35**, 109157 (2021).
- Salgia, R. & Kulkarni, P. The genetic/non-genetic duality of drug ‘resistance’ in cancer. *Trends Cancer* **4**, 110–118 (2018).
- Tape, C. J. et al. Oncogenic KRAS regulates tumor cell signaling via stromal reciprocity. *Cell* **165**, 910–920 (2016).
- Schmucker, R. et al. Combination treatment optimization using a pan-cancer pathway model. *PLoS Comput. Biol.* **17**, e1009689 (2021).
- Fröhlich, F. et al. Efficient parameter estimation enables the prediction of drug response using a mechanistic pan-cancer pathway model. *Cell Syst.* **7**, 567–579.e6 (2018).
- Kok, F. et al. Disentangling molecular mechanisms regulating sensitization of interferon alpha signal transduction. *Mol. Syst. Biol.* **16**, e8955 (2020).
- Steiert, B., Raue, A., Timmer, J. & Kreutz, C. Experimental design for parameter estimation of gene regulatory networks. *PLoS ONE* **7**, e40052 (2012).
- Bruggeman, F. J. & Westerhoff, H. V. The nature of systems biology. *Trends Microbiol.* **15**, 45–50 (2007).
- Babtie, A. C. & Stumpf, M. P. H. How to deal with parameters for whole-cell modelling. *J. R. Soc.* **14**, 20170237 (2017).
- Zheng, M. et al. ChemPert: mapping between chemical perturbation and transcriptional response for non-cancer cells. *Nucleic Acids Res.* **51**, D877–D889 (2023).
- Lachmann, A. et al. Massive mining of publicly available RNA-seq data from human and mouse. *Nat. Commun.* **9**, 1366 (2018).
- Subramanian, A. et al. A next generation connectivity map: L1000 platform and the first 1,000,000. *Profiles Cell* **171**, 1437–1452.e17 (2017).
- Eraslan, G., Avsec, Ž., Gagneur, J. & Theis, F. J. Deep learning: new computational modelling techniques for genomics. *Nat. Rev. Genet.* **20**, 389–403 (2019).
- Kumar, V. & Dogra, N. A comprehensive review on deep synergistic drug prediction techniques for cancer. *Arch. Comput. Methods Eng.* **29**, 1443–1461 (2022).
- Sakellaropoulos, T. et al. A deep learning framework for predicting response to therapy in cancer. *Cell Rep.* **29**, 3367–3373.e4 (2019).
- Vale-Silva, L. A. & Rohr, K. Long-term cancer survival prediction using multimodal deep learning. *Sci. Rep.* **11**, 13505 (2021).
- Buergel, T. et al. Metabolomic profiles predict individual multidisease outcomes. *Nat. Med.* **28**, 2309–2320 (2022).
- Golatkar, A., Anand, D. & Sethi, A. Classification of breast cancer histology using deep learning. In *Image Analysis and Recognition* (eds Campilho, A., Karray, F. & ter Haar Romeny, B.) 837–844 (Springer International Publishing, Cham, 2018).
- Heid, E. et al. Chemprop: a machine learning package for chemical property prediction. *J. Chem. Inf. Model.* **64**, 9–17 (2024).
- Ching, T. et al. Opportunities and obstacles for deep learning in biology and medicine. *J. R. Soc. Interface* **15**, 20170387 (2018).
- Fortelny, N. & Bock, C. Knowledge-primed neural networks enable biologically interpretable deep learning on single-cell sequencing data. *Genome Biol.* **21**, 190 (2020).
- Elmarakeby, H. A. et al. Biologically informed deep neural network for prostate cancer discovery. *Nature* **598**, 348–352 (2021).
- Nilsson, A., Peters, J. M., Meimetis, N., Bryson, B. & Lauffenburger, D. A. Artificial neural networks enable genome-scale simulations of intracellular signaling. *Nat. Commun.* **13**, 3069 (2022).
- Meimetis, N., Lauffenburger, D. A. & Nilsson, A. Inference of drug off-target effects on cellular signaling using interactome-based deep learning. *iScience* **27**, 109509 (2024).
- Barabási, A.-L. & Oltvai, Z. N. Network biology: understanding the cell’s functional organization. *Nat. Rev. Genet.* **5**, 101–113 (2004).
- Dugourd, A. et al. Causal integration of multi-omics data with prior knowledge to generate mechanistic hypotheses. *Mol. Syst. Biol.* <https://doi.org/10.15252/msb.20209730> (2021).
- Iorio, F. et al. Network based elucidation of drug response: from modulators to targets. *BMC Syst. Biol.* **7**, 139 (2013).
- Garrido-Rodriguez, M., Zirngibl, K., Ivanova, O., Lobentanzer, S. & Saez-Rodriguez, J. Integrating knowledge and omics to decipher mechanisms via large-scale models of signaling networks. *Mol. Syst. Biol.* **18**, e11036 (2022).
- Shannon, P. et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* **13**, 2498–2504 (2003).
- Bradley, G. & Barrett, S. J. CausalR: extracting mechanistic sense from genome scale data. *Bioinformatics* **33**, 3670–3672 (2017).
- Liu, A. et al. From expression footprints to causal pathways: contextualizing large signaling networks with CARNIVAL. *npj Syst. Biol. Appl.* **5**, 1–10 (2019).
- Aibar, S. et al. SCENIC: single-cell regulatory network inference and clustering. *Nat. Methods* **14**, 1083–1086 (2017).
- Kamimoto, K. et al. Dissecting cell identity via network inference and in silico gene perturbation. *Nature* **614**, 742–751 (2023).
- Ma, J. et al. Few-shot learning creates predictive models of drug response that translate from high-throughput screens to individual patients. *Nat. Cancer* **2**, 233–244 (2021).
- Mourragui, S. M. C. et al. Predicting patient response with models trained on cell lines and patient-derived xenografts by nonlinear transfer learning. *Proc. Natl. Acad. Sci. USA* **118**, e2106682118 (2021).
- Samek, W., Montavon, G., Lapuschkin, S., Anders, C. J. & Müller, K.-R. Explaining deep neural networks and beyond: a review of methods and applications. *Proc. IEEE* **109**, 247–278 (2021).
- Klauschen, F. et al. Toward explainable artificial intelligence for precision pathology. *Annu. Rev. Pathol.* **19**, 541–570 (2024).
- Keyl, P. et al. Patient-level proteomic network prediction by explainable artificial intelligence. *npj Precis. Onc.* **6**, 1–10 (2022).
- Laurie, M. & Lu, J. Explainable deep learning for tumor dynamic modeling and overall survival prediction using Neural-ODE. *npj Syst. Biol. Appl.* **9**, 1–11 (2023).
- Szigeti, B. et al. A blueprint for human whole-cell modeling. *Curr. Opin. Syst. Biol.* **7**, 8–15 (2018).
- Faure, L. et al. A neural-mechanistic hybrid approach improving the predictive power of genome-scale metabolic models. *Nature Communications* **14**, 4669 (2023).
- Liu, Y., Barr, K. & Reinitz, J. Fully interpretable deep learning model of transcriptional control. *Bioinformatics* **36**, i499–i507 (2020).
- Fields, S. & Song, O. A novel genetic system to detect protein–protein interactions. *Nature* **340**, 245–246 (1989).
- Rual, J.-F. et al. Towards a proteome-scale map of the human protein–protein interaction network. *Nature* **437**, 1173–1178 (2005).
- Furey, T. S. ChIP-seq and beyond: new and improved methodologies to detect and characterize protein–DNA interactions. *Nat. Rev. Genet.* **13**, 840–852 (2012).
- Richards, A. L., Eckhardt, M. & Krogan, N. J. Mass spectrometry-based protein–protein interaction networks for the study of human diseases. *Mol. Syst. Biol.* **17**, e8792 (2021).
- Jansen, R. et al. A Bayesian networks approach for predicting protein–protein interactions from genomic data. *Science* **302**, 449–453 (2003).

52. Margolin, A. A. et al. ARACNE: an algorithm for the reconstruction of gene regulatory networks in a mammalian cellular context. *BMC Bioinformatics* **7**, S7 (2006).
53. Zhang, Q. C. et al. Structure-based prediction of protein–protein interactions on a genome-wide scale. *Nature* **490**, 556–560 (2012).
54. Garzón, J. I. et al. A computational interactome and functional annotation for the human proteome. *eLife* **5**, e18715 (2016).
55. Mering, C. et al. STRING: a database of predicted functional associations between proteins. *Nucleic Acids Res.* **31**, 258–261 (2003).
56. Türei, D., Korcsmáros, T. & Saez-Rodríguez, J. OmniPath: guidelines and gateway for literature-curated signaling pathway resources. *Nat. Methods* **13**, 966–967 (2016).
57. Brunk, E. et al. Recon3D enables a three-dimensional view of gene variation in human metabolism. *Nat. Biotechnol.* **36**, 272 (2018).
58. Garcia-Alonso, L., Holland, C. H., Ibrahim, M. M., Turei, D. & Saez-Rodríguez, J. Benchmark and integration of resources for the estimation of human transcription factor activities. *Genome Res.* **29**, 1363–1375 (2019).
59. Edgar, R., Domrachev, M. & Lash, A. E. Gene expression omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Res.* **30**, 207–210 (2002).
60. Wilks, C. et al. recount3: summaries and queries for large-scale RNA-seq expression and splicing. *Genome Biol.* **22**, 323 (2021).
61. Ortmayr, K., Dubuis, S. & Zampieri, M. Metabolic profiling of cancer cells reveals genome-wide crosstalk between transcriptional regulators and metabolism. *Nat. Commun.* **10**, 1841 (2019).
62. Frejno, M. et al. Proteome activity landscapes of tumor cell lines determine drug responses. *Nat. Commun.* **11**, 3639 (2020).
63. Jones, R. C. et al. The Tabula Sapiens: A multiple-organ, single-cell transcriptomic atlas of humans. *Science* **376**, eabl4896 (2022).
64. Fan, J., Slowikowski, K. & Zhang, F. Single-cell transcriptomics in cancer: computational challenges and opportunities. *Exp. Mol. Med.* **52**, 1452–1465 (2020).
65. Lähnemann, D. et al. Eleven grand challenges in single-cell data science. *Genome Biol.* **21**, 1–35 (2020).
66. Mathur, L. et al. Combi-seq for multiplexed transcriptome-based profiling of drug combinations using deterministic barcoding in single-cell droplets. *Nat. Commun.* **13**, 4450 (2022).
67. Wörheide, M. A., Krumsiek, J., Kastenmüller, G. & Arnold, M. Multi-omics integration in biomedical research – A metabolomics-centric review. *Anal. Chim. Acta* **1141**, 144–162 (2021).
68. Samarasinghe, S. & Ling, H. A system of recurrent neural networks for modularising, parameterising and dynamic analysis of cell signalling networks. *Biosystems* **153–154**, 6–25 (2017).
69. Amini, A., Schwarting, W., Soleimany, A. & Rus, D. Deep evidential regression. In *Advances in Neural Information Processing Systems*, (eds Larochelle H, Ranzato M, Hadsell R, Balcan M. F., Lin H.) **33**, 14927–14937 (Curran Associates, Inc., 2020).
70. Engelhardt, B., Kschischo, M. & Fröhlich, H. A Bayesian approach to estimating hidden variables as well as missing and wrong molecular interactions in ordinary differential equation-based mathematical models. *J. R. Soc. Interface* **14**, 20170332 (2017).
71. van Engelen, J. E. & Hoos, H. H. A survey on semi-supervised learning. *Mach. Learn.* **109**, 373–440 (2020).
72. Hasani, R. et al. Closed-form continuous-time neural networks. *Nat. Mach. Intell.* **4**, 992–1003 (2022).
73. Niu, M. Y., Horesh, L. & Chuang, I. Recurrent neural networks in the eye of differential equations. Preprint at *arXiv preprint arXiv:1904.12933* (2019).
74. Lin, A. et al. Off-target toxicity is a common mechanism of action of cancer drugs undergoing clinical trials. *Sci. Transl. Med.* **11**, eaaw8412 (2019).
75. Armingol, E., Officer, A., Harismendy, O. & Lewis, N. E. Deciphering cell–cell interactions and communication from gene expression. *Nat. Rev. Genet.* <https://doi.org/10.1038/s41576-020-00292-x> (2020).
76. Letort, G. et al. PhysiBoSS: a multi-scale agent-based modelling framework integrating physical dimension and cell signalling. *Bioinformatics* **35**, 1188–1196 (2019).
77. Angione, C., Silverman, E. & Yaneske, E. Using machine learning as a surrogate model for agent-based simulations. *PLoS ONE* **17**, e0263150 (2022).
78. Sinha, S. et al. PERCEPTION predicts patient response and resistance to treatment using single-cell transcriptomics of their tumors. *Nat. Cancer* **5**, 938–952 (2024).
79. Jonathan, L. et al. An atlas of human metabolism. *Sci Signal.* **13**, eaaz1482 (2020).

### Acknowledgements

The authors want to thank Carolin Loos, Brian Joughin, Daniel M. Schafer, Nikolaos Meimetus, Caroline Uhler, Joshua M. Peters, Bryan Bryson, Julio Saez-Rodríguez, Pablo Rodríguez Mier, Sebastian Lobentanzer, Attila Gabor, and Fabian Froehlich for valuable input in developing the ideas presented in this perspective. The authors acknowledge funding from the Swedish Research Council, grant no. 2019-06349 (AN), Cancerfonden, grant no. 23 0693 JIA (AN), SciLifeLab & Wallenberg Data-Driven Life Science Program grant no. KAW 2020.0239 (A.N.), and National Institutes of Health, grant AI-201700104 (DAL). The funders played no role in the writing of this manuscript. The authors acknowledge funding from the Swedish Research Council, grant no. 2019-06349 (A.N.), Cancerfonden, grant no. 23 0693 JIA (A.N.), SciLifeLab & Wallenberg Data-Driven Life Science Program grant no. KAW 2020.0239 (A.N.), and National Institutes of Health, grant AI-201700104 (D.A.L.). The funders played no role in the writing of this manuscript.

### Author contributions

A.N. and D.A.L. conceived the paper, A.N. implemented the code, A.N. and N.M. wrote the manuscript, D.A.L. edited the manuscript.

### Competing interests

The authors declare no competing interests.

### Additional information

**Correspondence** and requests for materials should be addressed to Douglas A. Lauffenburger.

**Reprints and permissions information** is available at <http://www.nature.com/reprints>

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

**Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025