1

Leveraging Structured Biological Knowledge for Counterfactual Inference: a Case Study of Viral Pathogenesis

Jeremy Zucker[1*], Kaushal Paneri[2*], Sara Mohammad-Taheri[3*], Somya Bhargava[3], Pallavi Kolambkar[3], Craig Bakker[1], Jeremy Teuton[1], Charles Tapley Hoyt[4], Kristie Oxford[1], Robert Ness[5] and Olga Vitek[3†]

Abstract—Counterfactual inference is a useful tool for comparing outcomes of interventions on complex systems. It requires us to represent the system in form of a structural causal model, complete with a causal diagram, probabilistic assumptions on exogenous variables, and functional assignments. Specifying such models can be extremely difficult in practice. The process requires substantial domain expertise, and does not scale easily to large systems, multiple systems, or novel system modifications. At the same time, many application domains, such as molecular biology, are rich in structured causal knowledge that is qualitative in nature. This manuscript proposes a general approach for querying a causal knowledge graph with a causal question and converting the qualitative result into a quantitative structural causal model that can learn from data to answer the question. We demonstrate the feasibility, accuracy and versatility of this approach using two case studies in systems biology. The first demonstrates the appropriateness of the underlying assumptions and the accuracy of the results. The second demonstrates the versatility of the approach by querying a knowledge base for the molecular determinants of a severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2)-induced cytokine storm and performing counterfactual inference to predict the causal effect of medical countermeasures for severely ill COVID-19 patients.

1 Introduction

In systems biology, the basic unit of causality is a sequence of molecular interactions. Each time a cell senses changes in its environment, it marshals a complex choreography of molecular interactions to maintain homeostasis. When a virus infects the cell, this delicate balance is disrupted and can result in a cascade of systemic failures leading to disease.

The emergence of the COVID-19 pandemic has motivated a global community of scientists to understand the molecular mechanisms through which the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) dysregulates the immune response [1]. At the same time, the severity of the pandemic emphasized the need for computational techniques that help accelerate these efforts [2]. Systems and networks biology techniques such as modeling systems dynamics with ordinary differential equations (ODE) or stochastic differential equations (SDE) [3], pathway enrichment analysis [4], [5], [6], boolean/logical modeling [7], [8], [9], signal transduction analysis [10], [11], [12], and various statistical inference techniques applied to signed, directed networks [13], [14], [15], [16] all rely on domain knowledge of causal molecular interactions, and are useful for this task.

Counterfactual inference [17] offers an alternative and complementary avenue for studies of causal molecular interactions. At its heart is a formalism known as a Structural Causal Model (SCM) [17], [18], which represents domain knowledge in terms of causal diagrams, assumes a probability distribution on exogenous variables, and assigns a deterministic function to endogenous variables. Given assumptions in the model that are explicitly encoded in its structure and observational data, a structural causal model enables the estimation of causal effects of interventions.

Unfortunately, in problems with limited structured domain knowledge, such as the mechanism of viral infection with SARS-CoV-2, constructing an SCM is challenging. It is typically generated manually, by searching domain literature and consulting domain experts. At the same time, recent advances have generated natural language processing systems [19], [20], [21], large-scale automated assembly systems [22], and semi-automated curation workflows [23] that automatically capture vast amounts of qualitative domain knowledge. These systems are curated by several organizations [24], [25], [26], [27], [28] and stored in structured knowledge bases [29], [30], [31], [32]. By querying these biological knowledge bases, it is possible to translate a subnetwork into logical models [33], agent-based models [34], probabilisitic graphical models, rule-based models, and ODE [22]. This manuscript demonstrates that such knowledge bases can also serve as foundations for specifying SCM, and performing

*Equal contribution
[1] Pacific Northwest National Laboratory, Richland, WA
[2] Microsoft, Redmond, WA
[3] Northeastern University, Boston, MA
[4] Enveda Therapeutics, Bonn, Germany
[5] Altdeep, Boston, MA
† Corresponding author: ovitek@neu.edu

counterfactual inference in novel problems such as studies of SARS-CoV-2.

2 BACKGROUND

Biological processes Biological processes comprise sequences of causal molecular interactions, called *pathways*, that culminate in high-level functionality such as metabolism, signaling, and gene regulation. An example is mitogen-activated protein kinase (MAPK) signaling, which transmits information sensed at the cell surface to the nucleus, where a genetic program can be initiated to respond to the stimulus [35]. A causal diagram of MAPK consists of an external input S_1 and three proteins Raf, Mek, and Erk:

$$S_1 \rightarrow Raf \rightarrow Mek \rightarrow Erk \tag{1}$$

Each protein has two states, active or inactive. An active protein can activate a downstream protein, causing it to become active. The diagram provides a useful level of abstraction showing that input S_1 activates Raf, Raf activates Mek, and finally Mek activates Erk. Raf must phosphorylate Mek twice before it can become active. Likewise, Mek must phosphorylate Erk twice before it becomes active.

Quantitative modeling of biological processes with ODE/SDE Temporal dynamics of biological processes can be expressed quantitatively using ordinary (or stochastic) differential equations. A small number of high quality, validated models have been published in the literature and stored in a computable form in repositories such as Biomodels [36], [37]. For example, the MAPK signaling in Eq. (1) is well characterized. Denote $R(t)$, $M(t)$, and $E(t)$ the respective amounts of active (i.e., phosphorylated) Raf, Mek, and Erk at time t; T_R, T_M, and T_E their total amounts, which we assume do not change during the considered timeframe; v_R^{act}, v_R^{inh}, v_M^{act}, v_M^{inh}, v_E^{act}, and v_E^{inh} are experimentally derived activation or inhibition kinetic rate constants; and S_1 is the amount of the input signal. Then the system of ODE is specified as follows [38], [39]:

$$\begin{aligned}
\frac{\mathrm{d}R}{\mathrm{d}t} &= v_R^{\text{act}} S_1 (T_R - R(t)) - v_R^{\text{inh}} R(t) \\
\frac{\mathrm{d}M}{\mathrm{d}t} &= \frac{(v_M^{\text{act}})^2}{v_M^{\text{inh}}} R(t)^2 (T_M - M(t)) - v_M^{\text{act}} R(t) M(t) - v_M^{\text{inh}} M(t) \\
\frac{\mathrm{d}E}{\mathrm{d}t} &= \frac{(v_E^{\text{act}})^2}{v_E^{\text{inh}}} M(t)^2 (T_E - E(t)) - v_E^{\text{act}} M(t) E(t) - v_E^{\text{inh}} E(t)
\end{aligned} \tag{2}$$

Given initial conditions, direct simulations from the ordinary differential equation (ODE) can be used to generate the temporal trajectories of the amounts of activated proteins , such as $R(t)$, $M(t)$, and $E(t)$ in the MAPK example. In this manuscript we refer to such simulated data as *observational data*. We define an *ideal intervention* as an event that fixes the amount of an activated protein. For example, if we fix the number of phosphorylated proteins in Mek at $M(t) = m$, the second equality $\frac{\mathrm{d}M}{\mathrm{d}t}$ in Eq. (2) becomes zero. We can simulate data from Eq. (2) with $\frac{\mathrm{d}M}{\mathrm{d}t} = 0$, and refer to these as *interventional data*. Contrasting observational and interventional data helps evaluate the outcome of the intervention [40].

The deterministic ODE ignore the fact that at low concentration, stochasticity becomes a significant factor in determining the reaction [41]. As the collisions between molecules participating in biochemical process become stochastic, a stochastic model is required. In contrast to ODE, a stochastic differential equation

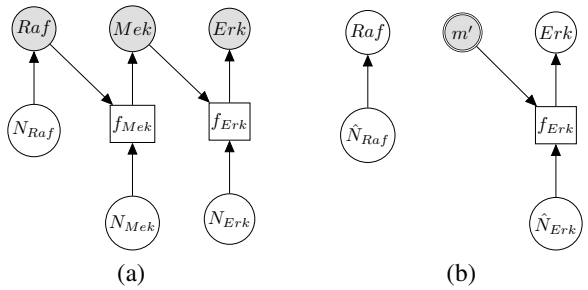


Fig. 1: **Causal modeling of MAPK signaling pathway** Circles are variables, double circles are variables intervened upon, squares are deterministic functional assignments, gray nodes are observed variables, and white nodes are hidden variables. (a) Structural causal model. N_{Raf}, N_{Mek} and N_{Erk} are statistically independent noise variables. Root node Raf is only dependent on noise variable N_{Raf}. Non-root nodes Mek and Erk are dependent on their parent and on the associated noise variable. (b) Counterfactual model. The intervention fixes the count of phosphorylated Mek to m', such that Mek is no longer dependent on Raf and N_{Mek}. Given an observed data point, counterfactual inference infers the noise variables $\hat{N}_{Raf}$, and $\hat{N}_{Erk}$.

model or stochastic differential equation (SDE) specifies biological process as a random process. For example, in the case of MAPK, the random process of the reaction $Mek \rightarrow Erk$ is specified with

$$\frac{dP_E(t)}{dt} = g_E(t, v_E^{act}, v_E^{inh}, M(t)),\ E(0) = e_0 \tag{3}$$

where $P_E(t)$ is marginal probability density of $E(t)$, function g_E determines the probability of a state change between $E(t)$ and $E(s), s > t$, e_0 is initial condition, and $M(t)$ is the value of its parent Mek at t. Once stochastic differential equation are fully specified, one can use, e.g. Gillespie's stochastic simulation algorithm [42] to simulate observational and interventional data, and evaluate the outcomes of interventions.

Unfortunately, even simple ODEs such as the one in the MAPK example are difficult to build *de novo*. This is nearly impossible for novel and poorly studied systems that lack the existence or findability of experimental information describing the structure or boundaries of the process, kinetic equations governing their dynamics [43], rate constants for these equations, or rules governing each agents' states and functions.

Equilibrium enzyme kinetics Simpler and more general quantitative models can be specified when a reaction reaches the state of chemical equilibrium [3]. One commonly used such model is *Hill function* in the form of

$$X = \beta \frac{\mathbf{PA}_X^n}{K^n + \mathbf{PA}_X^n} \tag{4}$$

where X is the abundance of a protein in a causal diagram (such as Erk in Eq. (1)), $\mathbf{PA}_X$ is the set of its parents, n is a parameter interpreted as the number of ligand binding sites of the protein, and β is the total number of molecules of the protein. A special and frequently used case of Hill function, called *Michaelis-Menten* function, occurs when $n = 1$. Although simple to use, these models are deterministic, and do not describe the stochasticity that is a distinctive property of biological systems at low concentrations.

Modeling biological processes with Structural Causal Models The stochastic nature of biological processes at equilibrium can be represented by Structural Causal Models (structural causal model

(SCM) [18]) [40], [44]. SCM are probabilistic generative causal models, typically specified in form of a causal graph such as in Fig. (1) (a) in the case of MAPK signaling. SCM represent the dependencies between a child node X and its parents $\mathbf{PA}_X$ in terms of a deterministic function $X = f_X(\mathbf{PA}_X, N_X)$ called *structural assignment*, and a noise variable N_X. In Fig. (1) (a), f_{Mek} and f_{Erk} are linear or non-linear structural assignments, and N_{Raf}, N_{Mek}, and N_{Erk} are statistically independent noise variables with defined probability distributions

$$Raf = N_{Raf};\ Mek = f_{Mek}(Raf; N_{Mek}) \qquad (5)$$
$$Erk = f_{Erk}(Mek, N_{Erk})$$

An ideal intervention in an SCM is performed on a functional assignment. For example, an ideal intervention on Mek sets $Mek = m'$, defining a new SCM

$$Raf = N_{Raf}; Mek = m'; Erk = f_{Erk}(Mek, N_{Erk}) \quad (6)$$

Ideal intervention can also be thought of as a process of mutilating the causal graph. For example, upon intervention on Mek, it is no longer dependent on its parent Raf, and therefore the edge from Raf to Mek is removed as shown in Fig. (1) (b). Contrasting observational and interventional data simulated directly from such two SCM graphs helps predict the outcome of an intervention.

Counterfactual inference with SCM Beyond direct model-based predictions, SCM enable *counterfactual inference*, i.e., the process of inferring the unseen outcomes of an intervention given an observed outcome without the intervention [17]. For example, for the MAPK signaling pathway, we may be interested in the counterfactual query: *Having observed $Raf = r$, $Mek = m$, $Erk = r$ without intervention, what would have been the number of phosphorylated Erk had the intervention fixed the number of phosphorylated Mek to m''?* We use Pearl's notation $do(X = x')$ [45] to refer to an intervention. In the case of MAPK, the example counterfactual query is

$$P(Erk_{do(Mek=m')}|Raf = r, Mek = m, Erk = e) \qquad (7)$$

The probability distribution in Eq. (7) is estimated with the following steps: (1) Given the observational data, infer the noise parameters and estimate the noise distributions. In the MAPK example, having observed $Raf = r$, $Mek = m$, $Erk = r$ we consider an intervention on Mek, and estimate the noise variables $\hat{N}_{Raf}$ and $\hat{N}_{Erk}$. Several inference algorithms are available for this task, e.g. Markov Chain Monte Carlo [46], Gibbs sampling [47], or no-u-turn Hamiltoninan Monte Carlo (HMC) [48]. In recent years, gradient-based inference algorithms such as stochastic variational inference [49] have grown in popularity because of their scalability, and their ability to convert inference problem into an optimization problem. (2) Apply the intervention to the SCM to generate a new SCM based on mutilated causal graph, as in Fig. (1) (b). In the new SCM, replace the noise variables with the estimated noise variables from (1). In MAPK model, $\hat{N}_{Raf}$ and $\hat{N}_{Erk}$ are replaced with N_{Raf} and N_{Erk} in Fig. (1) (a). (3) (4) Sample from the new SCM to obtain the counterfactual probability distribution.

Causal effect Causal effect quantifies the difference in the expected values between the observational and the counterfactual probability distribution of a variable. Our counterfactual definintion of the term causal effect is equivalent to the definition of "individual causal effect" in causal inference literature. In the MAPK example, a causal effect on Erk upon an intervention that fixes $Raf = r'$, conditional on one observational data point, is

$$\left\{Erk - Erk_{do(Raf=r')}\right\}|Raf = r, Mek = m, Erk = e \quad (8)$$

The causal effect shares stochastic components of the noise variables between observational and interventional data, and is therefore often more precise than a comparison based on a direct simulation [40].

In cases where domain knowledge is available to describe the systems dynamics in the form of an SDE, the system at equilibrium can be translated into an SCM to enable counterfactual reasoning and estimation of causal effects [40], [50]. Unfortunately, this process is challenging in novel and poorly studied systems, due to our limited ability to establish the structure of the causal graph.

Structured knowledge graphs Although there exist a multitude of biological knowledge bases that are manually curated from the literature [24], [25], [26], [27], [28], the systems biology community has coalesced around a small number of structured knowledge representations that differ mainly in their intended purpose. For example, the Systems Biology Markup Language (SBML) [30] was designed for interoperability among quantitative modeling software [30], the Biological Pathway Exchange Language (BioPAX) [29] was designed for pathway database integration [29], and the Systems Biology Graphical Notation (SBGN) [31] was designed for graphical layout [31]. In contrast, the Biological Expression Language (BEL) [32] was specifically designed for manual extraction and automated integration of causal relationships among biological entities, biological processes, and cellular-level observable phenomena [23].

The syntax of a BEL statement is comprised of a triple in the form of {*subject*, *predicate*, *object*}. Each subject and object represents a concept such as a protein abundance or a biological process that is grounded using terms from standard namespaces and encapsulated with information about its type. Predicates express the causal relationship between the subject and object, such as *directlyIncreases* or *regulates*. BEL statements can be chained together from the object of the first statement to the subject of the next statement, as shown in Fig. (2) for the case of the MAPK pathway.

BEL provides a number of valuable features for causal modeling. First, the restriction of BEL edges to causal relations implies the topology of the BEL graph can be reflected in the topology of the causal model. Second, the language is expressive enough for humans to manually curate a wide range of biological concepts, but formal enough to serve as a training corpus for natural language processing of biomedical literature (BioNLP) competitions. [51] Third, the BEL ecosystem is sufficiently mature that causal knowledge represented in other languages can be readily converted to BEL. [52], [53]

To bridge the gap between qualitative structured knowledge graphs and quantitative modeling, Gyori *et al.* developed the Integrated Dynamical Reasoner and Assembler (INDRA) [22] system to automatically assemble a set of causal statements into modeling formalisms including logical causal models [33], agent-based models [34], probabilisitic graphical models, rule-based models [54], [55], and ODEs. However, none of these formalisms can support counterfactual inference. In this manuscript, we demonstrate an approach for querying a knowledge base such as INDRA to export a qualitative causal model represented

```
act(p(fplx:RAF), ma(kin)) directlyIncreases act(p(fplx:MEK), ma(kin))
act(p(fplx:MEK), ma(kin)) directlyIncreases act(p(fplx:ERK, pmod(Ph)))
```

Fig. 2: **Example BEL statement** The statement details the processes in the MAPK signaling pathway in Eq. (1). The first line states that the active RAF kinase phosphorylates MEK, causing it to become an active kinase. The second line states that the active MEK kinase phosphorylates ERK, causing it to become an active kinase.

in the Biological Expression Language and compiling it into to a quantitative structural causal model that can be used to answer counterfactual questions about systems biology and viral pathogenesis.

3 METHODS

3.1 Notation, definitions and assumptions

Let $\mathbf{X} = \{X_i\}$ be a set of variables, such as proteins in a signaling network. Let $\mathbf{P} = \{P_j\}$ be a set of causal predicates that link these variables, such as increases, or regulates. Using this notation, we define a knowledge graph $\mathbb{K}$ as a set of k triples

$$\mathbb{K} = \{X_i, P_j, X_{i'} \,|X_i \in \mathbf{X}, P_j \in \mathbf{P}, X_{i'} \in \{\mathbf{X} \setminus X_i\}\}_{j=1}^{k} \quad (9)$$

We define a causal query $\mathbb{Q}$ as a set $\{\mathbf{X}^{\mathrm{c}},\ \mathbf{X}^{\mathrm{e}},\ \mathbf{X}^{\mathrm{z}}\}$ of variables that are potential causes, effects and covariates of interest for the biological investigation, where

$$\mathbf{X}^{\mathrm{c}} \subset \mathbf{X},\ \ \mathbf{X}^{\mathrm{e}} \subset \mathbf{X} \backslash \mathbf{X}^{\mathrm{c}}, \text{ and } \mathbf{X}^{\mathrm{z}} \subset \mathbf{X} \backslash \mathbf{X}^{\mathrm{c}} \backslash \mathbf{X}^{\mathrm{e}}$$

A pathway $\mathbb{P}(X_1, X_{k'+1})$, $k \leq k'$ is a sequence of a subset of triples from $\mathbb{K}$, where the object of the previous triple is subject of the next triple

$$\{(X_1, P_1, X_2), (X_2, P_2, X_3), \ldots, (X_{k'}, P_{k'}, X_{k'+1})\} \quad (10)$$

Our goal is to query the knowlege graph to generate a qualitative causal model $\mathbb{B}$ that links the causes, the effects and the covariates of interest. Importantly, the query result $\mathbb{B}$ induces a directed acyclic graph G with p variables from $\mathbf{X}$ as nodes, and causal relations from $\mathbf{P}$ as edges.

We assume that every variable in $\mathbb{B}$ is continuous. We denote $\mathbb{D} = \{X_{1j}, X_{2j}, ..., X_{pj}\}_{j=1}^{m}$ the observational data of m samples from the joint distribution $\mathcal{P}(\mathbf{X}; \theta)$. The distribution is specified in terms of parameters θ. We denote $\mathbf{R} \subset \mathbf{X}$ a set of nodes in G without parents.

3.2 Querying a knowledge graph to obtain a qualitative causal model

Given a biological knowledge graph $\mathbb{K}$ and a causal query of interest $\mathbb{Q}$, our first objective is to generate a qualitative causal model $\mathbb{B}$ capable of answering the query. To this end, we need to explore all potential paths in $\mathbb{K}$ from the cause to the effect in $\mathbb{Q}$, and then consider all covariates that may act as confounders of the causal question. This is done with the steps in Alg. 1. The algorithm can be implemented on any knowledge graph that represents causal relationships as directed edges, such as BEL or the Systems Biology Graphical Notation Activity Flow (SBGN-AF) language [56]. In the case of MAPK the counterfactual query in Eq. (7) corresponds to $\mathbb{Q} = \{\text{Mek}, \text{Erk}, \text{Raf}\}$.

We execute Alg. 1 step 2 to obtain all pathways from the cause to the effect:

$$\text{Mek} \rightarrow \text{Erk}$$

Algorithm 1 *Query to BEL algorithm*
Inputs: knowledge graph $\mathbb{K}$
causal query $\mathbb{Q} = \{\mathbf{X}^{\mathrm{c}},\ \mathbf{X}^{\mathrm{e}},\ \mathbf{X}^{\mathrm{z}}\}$
Outputs: $\mathbb{B}$

1: **procedure** QUERY2BEL($\mathbf{X}^{\mathrm{c}}, \mathbf{X}^{\mathrm{e}}, \mathbf{X}^{\mathrm{z}}, \mathbb{K}$)
2: ▶ *Get all pathways from cause to effect*
3: **for each** cause $X_i^{\mathrm{c}} \in \mathbf{X}^{\mathrm{c}}$ and for each effect $X_j^{\mathrm{e}} \in \mathbf{X}^{\mathrm{e}}$ **do**
4: find all pathways $\{\mathbb{P}(X_i^{\mathrm{c}}, X_j^{\mathrm{e}})\}$
5: ▶ *Get all pathways from covariates to causes*
6: **for each** covariate $X_i^{\mathrm{z}} \in \mathbf{X}^{\mathrm{z}}$ and for each cause $X_j^{\mathrm{c}} \in \mathbf{X}^{\mathrm{c}}$ **do**
7: find all pathways $\{\mathbb{P}(X_i^{\mathrm{z}}, X_j^{\mathrm{c}})\}$
8: ▶ *Get all pathways from covariates to effects*
9: **for each** covariate $X_i^{\mathrm{z}} \in \mathbf{X}^{\mathrm{z}}$ and for each effect $X_j^{\mathrm{e}} \in \mathbf{X}^{\mathrm{e}}$ **do**
10: find all pathways $\{\mathbb{P}(X_i^{\mathrm{z}}, X_j^{\mathrm{e}})\}$
11: $\mathbb{B} = \{\mathbb{P}(X_i^{\mathrm{c}}, X_j^{\mathrm{e}})\} \cup \{\mathbb{P}(X_i^{\mathrm{z}}, X_j^{\mathrm{c}})\} \cup \{\mathbb{P}(X_i^{\mathrm{z}}, X_j^{\mathrm{e}})\}$
12: **return** $\mathbb{B}$

We execute Alg. 1 step 5 to obtain all pathways from the covariate to the cause:

$$\text{Raf} \rightarrow \text{Mek}$$

We execute Alg. 1 step 8, but since there are no new pathways from the covariate RAF to the effect ERK, we obtain the empty set. The final returned model is:

$$\text{Raf} \rightarrow \text{Mek} \rightarrow \text{Erk}$$

3.3 Compiling a qualitative causal model to a quantitative structural causal model

Our second objective is to express the qualitative causal structure in $\mathbb{B}$ into a quantitative SCM, and estimate the parameters of the SCM from experimental data. These steps are described in Algorithm 2.

Input The algorithm takes as input a BEL causal query result $\mathbb{B}$ and observed measurements on its variables $\mathbb{D}$.

Get network structure G from $\mathbb{B}$ (Alg. 2 line 3) Since a BEL statement identifies parents and children, it induces a causal network structure. We determine this structure by traversing BEL statements with the breadth first search approach, starting with root variables (such as Raf in Figure 2). For all the non-root variables, the algorithm waits until all the parents are traversed.

For each root node R, use $\mathbb{D}$ to estimate parameters θ of $\mathcal{P}(R; \theta)$(Alg. 2 line 5) In order to specify the SCM, we need to define the type and parameters of the marginal probability distributions of the root variables $\mathcal{P}(R; \theta)$. The BEL statement provides a prior knowledge about the distribution in a parametric form. Therefore, this step involves techniques such as maximum likelihood to estimate the parameters of this distribution.

For example, in a stochastic MAPK system at equilibrium the root variable the number of active Raf in a cell follows a Binomial distribution. When the maximum number of active or inactive particles in the system is large, the Binomial distribution can be approximated with a Normal distribution with $\theta_{Raf} =$

Algorithm 2 *BEL2SCM algorithm*
Inputs: BEL statements $\mathbb{B}$
$\mathbb{D} \sim P(X_1, ..., X_p)$
Outputs: $SCM\ \mathbb{M} = \{f_i(\mathbf{PA}_i, N_i)\}_{i=1}^{p}$

procedure BEL2SCM($\mathbb{B}$, $\mathbb{D}$)
  $\mathbb{M} = \{\}$
  Get network structure G from $\mathbb{B}$.
  for each $R \in \mathbf{R}$ in G **do**
    ▶ *Use* $\mathbb{D}$ *to estimate parameters* θ *of* $\mathcal{P}(R;\theta)$
      $\theta = arg\,max_\theta \mathcal{P}(R;\theta \mid \mathbb{D})$
    ▶ *Reparameterize* $\mathcal{P}(R;\theta)$ *in terms of* f_R *and* N_R
      $N_R \sim \mathcal{N}(0,1)$
      $f_R(N_R) = F^{-1}_{\mathcal{P}(R;\theta)}(N_R)$
    $\mathbb{M}$.Add($f_R(N_R)$)
  for each $X \in \{\mathbf{X} \setminus \mathbf{R}\}$ in G **do**
    ▶ *Estimate parameters* $\mathbf{w}$ *and* b *of sigmoid function*
      $\log(\frac{X}{\beta_X - X}) = \mathbf{w}'\mathbf{PA}_X + b$
    ▶ *Define distribution of* N_X *from model residuals.*
      $residual = X - \frac{\beta_X}{1+\exp(-\mathbf{w}'\mathbf{PA}_X - b)}$
      $N_X \sim \mathcal{N}(0, MSE(residual))$
    ▶ *Get* $f_X(\mathbf{PA}_X, N_X)$ *with additive* N_X.
      $f_X(\mathbf{PA}_X, N_X) = \frac{\beta_X}{1+exp(-\mathbf{w}'_X\mathbf{PA}_X - b_X)} + N_X$
    $\mathbb{M}$.Add($f_X(\mathbf{PA}_X, N_X)$).
  return $\mathbb{M}$

$(\mu_{Raf}, \sigma^2_{Raf})$. We then estimate θ_{Raf} using maximum likelihood from the observed Raf in $\mathbb{D}$.

For each root node R, reparameterize $\mathcal{P}(R;\theta)$ in terms of f_R and N_R (Alg. 2 line 7) The specification of an SCM requires us to separate the deterministic and the stochastic components of variation of each variable as shown in Fig. (1). We accomplish this using a reparameterization technique popularized by variational autoencoders [57], which was shown to make counterfactual inference consistent with core biological assumptions [58]e. In the case of root nodes, we reparameterize $\mathcal{P}(R;\theta)$ with Uniform(0,1), and then pass it to the inverse CDF of $\mathcal{P}(R;\theta)$, as follows

$$\begin{aligned} \text{Original :} \quad & R \sim \mathcal{P}(R;\theta) \\ \text{Reparametrized :} \quad & N_R \sim \text{Uniform}(0,1) \\ & f_R(N_R) = F^{-1}_{\mathcal{P}(R;\theta)}(N_R) \end{aligned} \tag{11}$$

where $F^{-1}_{\mathcal{P}(R;\theta)}(N_R)$ is inverse cumulative distribution function of $\mathcal{P}(R;\theta)$. In the case of MAPK, since Raf follows Normal distribution with parameters θ_{Raf}, the reparameterization simplifies even further to

$$\begin{aligned} \text{Original :} \quad & Raf \sim \mathcal{N}(\mu_{Raf}, \sigma^2_{Raf}) \\ \text{Reparametrized :} \quad & N_{Raf} \sim \mathcal{N}(0,1) \\ & f_{Raf}(N_{Raf}) = \sigma_{Raf} N_{Raf} + \mu_{Raf} \end{aligned} \tag{12}$$

Add R to $\mathbb{M}$ (Alg. 2 line 10) For each root node, we add the corresponding function $f_R(N_R)$ and its noise variable N_R to $\mathbb{M}$. For example, since MAPK has only one root node Raf, the Algorithm adds $f_{Raf}(N_{Raf})$ to $\mathbb{M}$.

For each $X \in \{\mathbf{X} \setminus \mathbf{R}\}$, estimate parameters $\mathbf{w}$ and b of sigmoid function (Alg. 2 line 12) In order to specify the SCM for non-root nodes, we need to define the form (polynomial, linear, non-linear, sigmoid, etc.) of functional assignments linking the measurements on the parent nodes to the measurements on the child. We chose the functional assignment in the form of a sigmoid function

$$\log\left(\frac{X}{\beta_X - X}\right) = \mathbf{w}'\mathbf{PA}_X + b \tag{13}$$

where β_X is the maximal number of activated protein molecules. For a node X with q parents, $\mathbf{PA}_X$ is a $q \times 1$ vector of measurements on the parent nodes, $\mathbf{w}$ is a $1 \times q$ vector of weights, and b is a scalar bias. Parameters $\mathbf{w}$ and b of the sigmoid function are estimated from the data, e.g. using smooth L_1 loss function.

In the example of the MAPK pathway, f_{Mek} has only one parent. Therefore f_{Mek} has the form

$$f_{Mek}(Raf, N_{Mek}) = \frac{\beta_{Mek}}{1 + exp(-w_{Mek}Raf - b)} + N_{Mek} \tag{14}$$

We use the sigmoid function in Eq. (13) as a special case of the Hill equation. The full parametric description of the Hill equation has a nuanced precise biochemical interpretation. The sigmoid function maintains the Hill equation's functions, but with a reduced set of parameters that are easier to estimate. Fig. (3) shows that the approximation is reasonable for a range of parameter values.

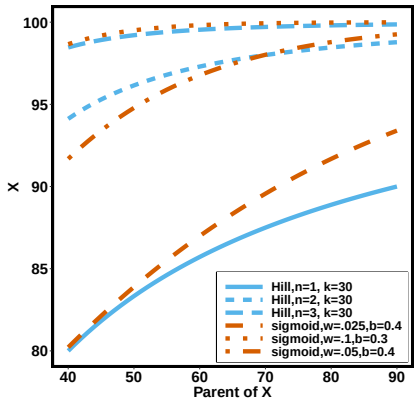


Fig. 3: **Examples of Hill function and sigmoid function for two variables** X is a single node that has a single parent $\mathbf{PA}_X$. We use Hill function ($X = \beta\frac{\mathbf{PA}_X^n}{K^n + \mathbf{PA}_X^n}$) and sigmoid function as in Eq. (13) to predict the value of X given its parent value. In the Hill function, K is the activation rate, n defines the steepness of function and β is fixed at 100. Blue lines correspond to Hill equation with $K = 30$ and $n \in \{1, 2, 3\}$. Brown lines correspond to sigmoid function where $b \in \{0.4, 0.3, 0.4\}$ and $w \in \{0.025, 0.1, 0.5\}$

Define distribution of N_X from model residuals (Alg. 2 line 14) Similarly to the root variables, for non-root variables we assume that the noise variables follow Normal distribution with 0 mean. The variance of this distribution is estimated from the residuals of the model fit in the previous step. For example, in the MAPK pathway, f_{Mek} has only one parent Raf. Therefore, the residuals of the sigmoid curve fit for Mek are defined as

$$residual_{Mek} = Mek - \frac{\beta_{Mek}}{1 + \exp(-w_{Mek}Raf - b)} \tag{15}$$

and the distribution of the noise variable is defined as $N_{Mek} \sim \mathcal{N}(0, MSE(residual_{Mek}))$

Get $f_X(\mathbf{PA}_X, N_X)$ with additive N_X (Alg. 2 line 17) The step combines the sigmoid functional assignment and the independent

noise variable. In the example of Mek in the MAPK pathway, the step outputs

$$f_{Mek}(Raf, N_{Mek}) = \frac{\beta_{Mek}}{1 + exp(-w_{Mek}Raf - b)} + N_{Mek} \tag{16}$$

Add $f_X(\mathbf{PA}_X, N_X)$ to SCM (Alg. 2 line 19) The step iteratively adds (f_X, N_X) for all $X \in \mathbf{X}$.

Output (Alg. 2 line 20) The algorithm returns a generative structural causal model $\mathbb{M} = \{f_i(\mathbf{PA}_i, N_i)\}_{i=1}^{p}$ where $\mathbf{PA}_i \subset \mathbf{X}$. For example, in the case of the MAPK model, it returns $[N_{Raf}, N_{Mek}, N_{Erk}, f_{Raf}(N_{Raf}), f_{Mek}(Raf, N_{Mek}), f_{Erk}(Mek, N_{Erk})]$.

3.4 Counterfactual inference

The generated SCM enables counterfactual inference using the standard procedure [17]. Given a new observation $\mathbb{D}^{new}$,

- **Step 1 (abduction):** Update the probability $P(N_X)$ to obtain $P(N_X|\mathbb{D}^{new})$.
- **Step 2 (action):** Replace the equations determining the variables in set $\mathbf{X}^{\mathrm{c}}$ by $\mathbf{X}^{\mathrm{c}} = \mathbf{x}^{\mathrm{c}\prime}$.
- **Step 3 (prediction):** Sample from the modified model to generate the target distribution $\mathbf{X}^{\mathrm{e}}_{do(\mathbf{X}^{\mathrm{c}}=\mathbf{x}^{\mathrm{c}\prime})}$.

After generating the target distribution of the intervention model, we estimate causal effects. Alg. 3 in Appendix describes the detailed steps of counterfactual inference.

3.5 Implementation

Causal query to Biological Expression Language (QUERY2BEL) was implemented manually using a publicly available instance of BioDati Studio [59], then validated using INDRA's interactive dialogue system Bob with BioAgents [22]. Biological Expression Language to Structural Causal Models (BEL2SCM) was implemented in PyTorch.

SCM-based counterfactual inference was performed with Pyro [60], due to its ability to perform interventions on probabilistic models and scalability to larger models, as described in Alg. 3 in Appendix. Specifically, the implementation relies on the following functionalities in Pyro. The `pyro.do` method is an implementation of Pearl's do-operator used for causal inference. The `pyro.infer.Importance` method performs posterior inference by importance sampling. The `pyro.infer.EmpiricalMarginal` method performs marginal distribution over a single site (or multiple, provided they have the same shape) from the TracePosterior's model.

Experiments in this manuscript took between 13 to 82 seconds depending on the graph size on a system with Intel Core i7 8th Gen CPU, 16 GB RAM and Ubuntu 18.04 Operating System. The code is available at https://github.com/bel2scm.

4 Case Studies

Below we introduce two biological case studies investigated using the approach proposed in this manuscript. The first case study allows us to evaluate the accuracy of the results based on known ground truth. The second uses counterfactual reasoning to pinpoint the mechanism by which SARS-CoV-2 infection can lead to a Cytokine Release Syndrome (cytokine storm) in severely ill coronavirus disease 2019 (COVID-19) patients. The details of the case studies, parameter values of the simulations, and of the results are at https://github.com/bel2scm.

Algorithm 3 *Estimating causal effect on T upon intervening on U*
Inputs: New data point $\mathbb{D}^{new}$
target node T
observational data for target node $\mathbb{D}_T$
intervention value u
node to intervene upon U
number of iteration I
network structure G
SCM $\mathbb{M}$
Outputs: Causal Effect CE

procedure GETCAUSALEFFECT($\mathbb{D}^{new}, T, \mathbb{D}_T, u, U, I, G, \mathbb{M}$)
  $\hat{N} = \{\}$
  ▶ *Interventional data for target node T*
  $\mathbb{ID}_T = \{\}$
  for I **do**
    for each $X \in \{\mathbf{X} \setminus U\}$ in G **do**
      ▶ *Apply stochastic variational inference (SVI)*
      $\hat{N}_X = SVI(\mathbb{D}^{new})$
      $\hat{N}$.Add($\hat{N}_X$)
    ▶ *Apply intervention on U*
    $CM = pyro.do(\mathbb{M}, U = u)$
    ▶ *Get posterior of CM with importance sampling*
    $CMP = pyro.infer.Importance(CM, \hat{N})$
    ▶ *Get EmpiricalMarginal (EM) for T*
    $CMM = pyro.infer.EM(CMP, T)$
    $\mathbb{ID}_T$.Add(CMM)
  $CE = \mathbb{D}_T - \mathbb{ID}_T$
  return CE

4.1 Case study 1: the IGF signaling system

The system The insulin-like growth factor (insulin-like growth factor (IGF)) signaling system (Figure 4) regulates growth and energy metabolism of a cell. Activated by external stimuli, the insulin-like growth factor or the epidermal growth factor (EGF) trigger the signaling events, which include the MAPK pathway in Eq. (1). Similarly to Eq. (1), nodes in the system are proteins, and edges are enzyme reactions. However, the system is larger and more complex. It includes two different paths from Ras to Erk, one direct and the other through $PI3K$ and Akt. This challenges predictions of outcomes of interventions. In the following, we assume that the IGF system is a closed world, and has no unobserved confounders.

Intervention We considered two interventions. First, we considered an ideal intervention on the IGF system, which fixes phosphorylated Mek to 40, and unphosphorylated Mek to 60. Second, we considered an ideal intervention which fixes the number of phosphorylated Ras to 30 and unphosphorylated Ras to 70, and considered the effect of the interventions on Erk.

Causal effects of interest We are interested in two counterfactual questions. First, *Having observed a new datapoint $\mathbb{D}^{new}$ without any interventions for all the variables in IGF pathway, what would have been the number of phosphorylated Erk had the intervention fixed the number of phosphorylated Mek to 40?*, the second query is as above, but with the intervention fixing the number of phosphorylated Ras to 30. More formally, we are interested in these causal effects

$$\left\{Erk - Erk_{do(Mek=40)}\right\} | \mathbb{D}^{new} \tag{17}$$

$$\left\{Erk - Erk_{do(Ras=30)}\right\} | \mathbb{D}^{new} \tag{18}$$

$\mathbb{D}^{new}$ is a new datapoint generated from ODE-based simulation.

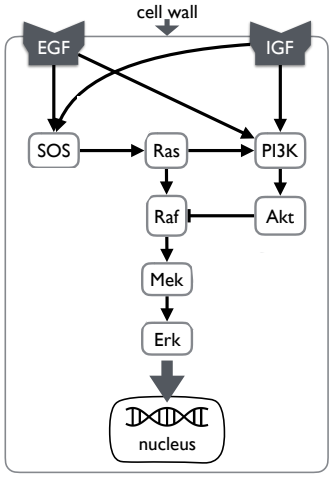


Fig. 4: **Case Study 1: the IGF signaling system** The insulin-like growth factor (IGF) and epidermal growth factor (EGF) are receptors of external stimuli, triggering downstream signaling pathways that include the MAPK pathway. All the relationships between abundances of activated proteins in this network are of the type *increase*, except for the relationship between Akt and Raf which is of the type *decrease*.

Evaluation The IGF system has been extensively investigated, and its dynamics is well characterized in form of ODE and SDE models [38]. We therefore evaluate the accuracy of SCM-based prediction of causal effect with respect to results of direct ODE- and SDE-based simulations. While generating observational data to train the SCM, we simultaneously sampled data points from SDE upon the intervention, and viewed the distribution of differences of simulated counts of phosphorylated Erk as the ground truth. A similar evaluation is also performed with deterministic ODE, where causal effect is a single value.

4.2 Case study 2: host response to viral infection

The system Retrospective studies indicate that high levels of pro-inflammatory cytokine Interleukin 6 (IL6) are strongly associated with severely ill COVID-19 patients [61]. One recently proposed explanation for this observation is the viral induction of a positive feedback loop, known as Interleukin 6 Amplifier (IL6-AMP) [62]. IL6-AMP is stimulated by simultaneous activation of nuclear factor kappa-light-chain-enhancer of activated B cell (NF-κB) and Signal Transducer and Activator of Transcription 3 (STAT3) [63]. This in turn induces various pro-inflammatory cytokines and chemokines, including Interleukin 6, which recruit activated T cells and macrophages. This strengthens the Interleukin 6 Amplifier into a positive feedback loop leading to a cytokine storm [64], which is believed to be responsible for the tissue damage observed in patients with acute respiratory distress syndrome [62].

Intervention Originally developed to treat autoimmune disorders such as rheumatoid arthritis [65], Tocilizumab (Toci) is an antibody that targets the soluble Interleukin 6 receptor and can effectively block the IL6 signal transduction pathway [66]. Tocilizumab has emerged as a promising drug repurposing candidate to reduce mortality in severely ill COVID-19 patients [67], [68].

Causal effect of interest We define a severely ill COVID-19 patient as someone with CytokineStorm > 65. We are interested in the causal effect of the treatment on the treated.

$$\left\{\text{CytokineStorm} - \text{CytokineStorm}_{do(Toci=0)}\right\} | \mathbb{D}^{new} \qquad (19)$$

where $\mathbb{D}^{new}$ is an observation that a patient received Tocilizumab treatment and became severely ill. We wish to know what would have happened had she had not received the treatment.

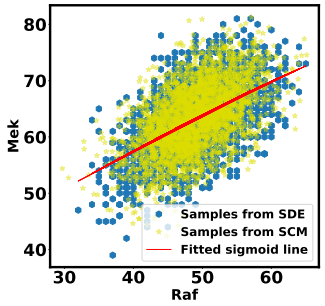
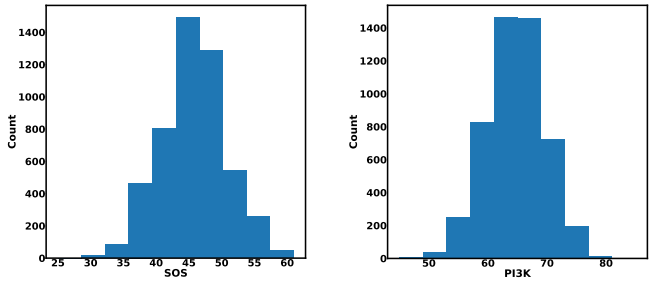


Fig. 5: **Case study 1: IGF model** Scatter plot of Mek versus Raf. Blue points are the data points generated by SDE. Yellow points are the predicted samples from SCM. The red line is the fitted sigmoid curve in Alg. 2 line 12.



(a) (b)

Fig. 6: **Case study 1: probability distributions of the root nodes of IGF model** (a) Histogram of SOS generated from SDE simulation (b) As in (a), for $PI3K$.

Evaluation Although this case study has no ground truth, we know that Tocilizumab has a strong inhibitory effect on soluble Interleukin 6 receptor, and expect that the severity of the cytokine storm would have been worse had she not received treatment. We therefore expect the causal effect to be negative.

We further evaluated the estimation of causal effects in two ways. First, we compared the estimates of causal effects in Alg. 2 to estimates obtained with the SCM of the same form, but with known weights of functional assignments. This comparison allows us to characterize the impact of weight estimation on the accuracy of causal effects. Second, we compared the estimates of causal effects in Alg. 2 with the average causal effect obtained by direct simulations from two SCMs of the same form, with and without the intervention, and with known weights (as opposed to conditioning on $\mathbb{D}^{new}$ and performing counterfactual inference). This comparison allows us to characterize the usefulness of counterfactual inference when estimating causal effects.

5 Results

5.1 Case study 1: the IGF signaling system

Generating BEL causal model The BEL representation of the IGF system was manually curated using BioDati studio [59], to match

the existing ODE and SDE. The BEL representation of the IGF system specified all the node types as in category *abundance*. All the relationships between parents and children nodes were of type *increase*, except for the parent node Akt, where the relationship was of type *decrease*.

Observational data We mimicked the process of collecting observational data by simulating protein counts from the corresponding ODE and SDE. In both models, each reaction was parametrized with a rate of activation and deactivation. The initial number of phosphorylated proteins was set to 100, and the number of unphosphorylated proteins was set to 0 for all the proteins. Therefore, the total number of protein molecules $\beta_X = 100$. The initial number of particles for the receptor was 37 for EGF and 5 for IGF. The deterministic simulation numerically solved the ODE using the *deSolve* [69] R package. The stochastic simulation used the Gillespie algorithm [42] from the *smfsb* [70] R package.

Appropriateness of model assumptions SCM-based predictions with sigmoid approximation of functional assignments were well within the range of the SDE-based data (as shown for Raf and Mek in Fig. (5)). Similar results were obtained for predictions of Ras, $PI3K$, AKT, Raf, and Erk. The fitted functional assignment had little curvature. This indicates that a more complicated function with more parameters, such as Hill equation, was unnecessary in this case.

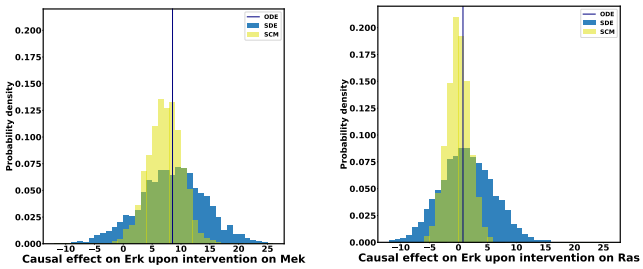
To further evaluate the plausibility of the assumptions, Fig. (6) shows the histograms of the SDE-generated abundances of root nodes, which were not affected by functional assignments in SCM. The shape of the histograms indicate that the assumption of Normal distribution was plausible.

Accuracy of causal effects Fig. (7) shows the estimated causal effects on Erk upon intervention on Mek and on Ras, compared to the estimates from ODE and SDE-based simulations. For both interventions, the expected values of causal effects based on ODE, SDE and SCM were consistent. However, the SCM-based had a narrower variance, reflecting a smaller uncertainty in the causal effect. This is expected, and illustrates the benefits of counterfactual inference. Since the counterfactual inference shares stochastic components in experiments with and without intervention, it reduces nuisance variation. In contrast, direct SDE-based simulations are negatively impacted by different stochasticity in simulation with and without the intervention.

The causal effect of intervening on Mek was higher than the causal effect of intervening on Ras, for the following reason. While Mek directly influences Erk (i.e., there is a single path from Mek to Erk), Ras has two pathways to Erk. The path through AKT has an inhibiting (deactivation) effect on Raf, and estimated negative weights in the sigmoid function in Eq. (13). The alternative path, a cascade from Ras to Erk, has the opposite (activating) effect on Erk. The two paths mitigate the overall causal effect from Ras on Erk.

5.2 Case study 2: host response to viral infection

Generating BEL causal model In accordance with the inputs to Alg. 1, we defined the knowledge base $\mathbb{K}$ as the Covid-19 knowledge network automatically assembled from the COVID-19 Open Research Dataset (CORD-19) [2] document corpus using the INDRA workflow. We defined the cause $\mathbf{X}^{\mathrm{c}}$ as sIL6Rα, the effect $\mathbf{X}^{\mathrm{e}}$ as cytokine storm, and the covariates $\mathbf{X}^{\mathrm{z}}$ as SARS-CoV-2



(a) (b)

Fig. 7: **Case study 1: estimated causal effect on Erk.** Since SCM- and SDE-based simulations are stochastic, the causal effects have a distribution. In contrast, ODE-based simulation is deterministic, and its estimated causal effect is a constant. (a) Upon intervention on Mek (b) Upon intervention on Ras.

and Toci. Therefore the causal query of interest was defined as $\mathbb{Q} = \{\text{sIL6R}\alpha, \text{CytokineStorm}, \{\text{SARS-CoV-2}, \text{Toci}\}\}$.

Alg. 2 line 2 generated all pathways from soluble Interleukin 6 receptor to Cytokine Release Syndrome, resulting in

$$\text{sIL6R}\alpha \rightarrow \text{IL6-STAT3} \rightarrow \text{IL6-AMP} \rightarrow \text{CytokineStorm}$$

Next, line 5 generated all pathways from Tocilizumab to soluble Interleukin 6 receptor:

$$\text{Toci} \rightarrow \text{sIL6R}\alpha$$

We then generated all pathways from severe acute respiratory syndrome coronavirus 2 to soluble Interleukin 6 receptor:

$$\text{SARS-CoV-2} \dashv \text{ACE2} \dashv \text{Angiotensin II} \rightarrow \text{AGTR1} \rightarrow \\ \rightarrow \text{ADAM17} \rightarrow \text{sIL6R}\alpha$$

Line 8 found no new branches from Tocilizumab to Cytokine Release Syndrome. Finally, we generated all pathways from severe acute respiratory syndrome coronavirus 2 to Cytokine Release Syndrome, which resulted in three new branches

$$\text{SARS-CoV-2} \rightarrow \text{PRR} \rightarrow \text{NF-}\kappa\text{B} \rightarrow \text{IL6-AMP}$$
$$\text{ADAM17} \rightarrow \text{EGF} \rightarrow \text{EGFR} \rightarrow \text{NF-}\kappa\text{B}$$
$$\text{ADAM17} \rightarrow \text{TNF}\alpha \rightarrow \text{NF-}\kappa\text{B}$$

The steps above produced the qualitative causal model illustrated in Fig. (8), and the corresponding BEL causal model $\mathbb{B}$.

Observational data At the time of writing, there were no publicly available COVID-19 datasets quantifying protein abundances of cytokine storm at the single-cell level, which is required to train the parameters of the sigmoidal structural causal model. Therefore, we simulated data from a fully specified structural causal model, where parameters were carefully chosen to ensure that the variables were in the desired range of 0–100. The root nodes SARS-CoV-2 and Tocilizumab were sampled from a Normal distribution with mean of 50 and standard deviation of 10. The non-root nodes were sampled from a sigmoid function as in Eq. (13). We randomly generated two new observations $\mathbb{D}^{new}$ with Cytokine Release Syndrome > 65 that mimic severely ill patients. The first observation had a higher value of cytokine storm and a lower value of Toci. The second observation had a lower value of cytokine storm and a higher value of Toci.

Estimation of causal effect Fig. (9) compares the estimates of causal effect of Tocilizumab on cytokine storm based on the SCM

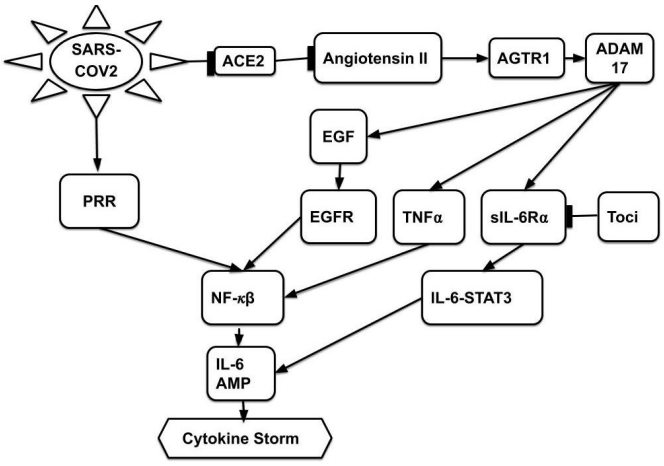


Fig. 8: **Case study 2: host response to viral infection** Pointed edges represent relationships of type *increase*; flat-headed edges represent relationships of type *decrease*. Nodes SARS-COV2 and Toci are external stimuli.

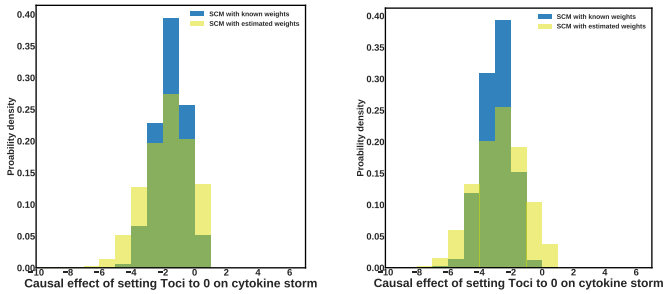


Fig. 9: **Case study 2: causal effect on cytokine storm upon intervention with $Toci = 0$, estimated as in Alg. 2, and based on an SCM of the same form but with known weights of functional assignments.** The noise associated with each exogenous variable in the model induces a distribution on the estimation of the causal effect. (a) $\mathbb{D}^{new}$ has a higher value of SARS-CoV-2 and a lower value of Toci. (b) $\mathbb{D}^{new}$ has a lower value of SARS-CoV-2 and a higher value of Toci.

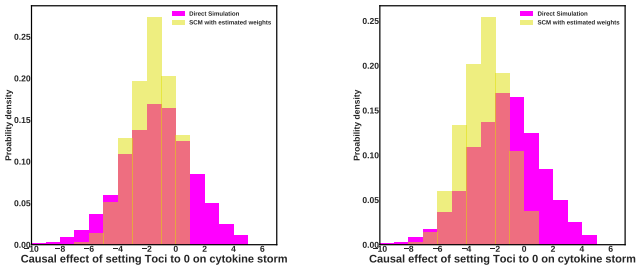


Fig. 10: **Case study 2: comparison of causal effects estimated as in Alg. 2 and causal effects obtained from direct model-based simulations.** Yellow histogram is the same as in Fig. (9)(a) $\mathbb{D}^{new}$ has a higher value of SARS-CoV-2 and a lower value of Toci. (b) $\mathbb{D}^{new}$ has a lower value of SARS-CoV-2 and a higher value of Toci.

in Alg. 2, and based on the SCM with weights of structural assignments used to generate the data (which we view as ground truth). The causal effects were conditional on two different new observations $\mathbb{D}^{new}$. In both cases, the histograms were well aligned, indicating that the weights of functional assignments were accurately estimated. The causal effects with weights estimated based on Alg. 2 had a larger variance. This is expected, and reflects the additional uncertainty induced by weight estimation.

Also as expected, the histograms were centered away from zero towards negative causal effects, indicating that a severely ill patient receiving the Tocilizumab treatment was indeed less likely to experience cytokine storm. The locations of the distributions of causal effects in Fig. (9) (a) and (b) shift, depending on the values of the new observation $\mathbb{D}^{new}$. Specifically, the causal effect is stronger with patients with lower level of SARS-CoV-2 and higher level of Toci. This emphasizes the utility of counterfactual inference, and its ability to make individualized predictions.

To further evaluate the practical utility of counterfactual inference, Fig. (10) contrasts estimates of causal effects in Alg. 2 with estimates obtained by direct simulations from two SCMs of the same form, with and without the intervention. As can be seen, the distribution of causal effects obtained by direct simulation does not depend on the individual patient, its variance is much broader, and the distribution contains 0. This variance will increase even further in situations where weights of structural assignments are unknown and need to be estimated. This leads to the simulation-based causal effects being less precise, and emphasizes the ability of counterfactual inference to provide more precise and individualized recommendations.

6 Discussion

We proposed a general approach that leverages structured qualitative prior knowledge, automatically generates a quantitative SCM, and enables answers to counterfactual research questions. Its application to the IGF signaling system demonstrated the appropriateness of the underlying assumptions, and the accuracy of the results. The application to a study of host response to SARS-CoV-2 infection demonstrate the feasibility, versatility and usefulness of our approach as applied to a novel biological problem.

In particular, the use of causal BEL grammar allowed us to leverage a rapidly developing knowledge bases, and specify an SCM and perform individualized counterfactual inference in a quick and automated manner, which would otherwise require a substantial manual effort.

The approach opens multiple directions for future research. In particular, future work will extend the configurability of the BEL2SCM algorithm by incorporating the rich type information of predicates in BEL, and mapping parent-child type signatures to functional forms such as post-nonlinear models, neural networks, mass action kinetics and Hill equations, and incorporating additional data types such as binary variables, categorical variables, and continuous variables with constraints on their domains.

We also note that experimentalists typically formulate biological processes as linear pathways (e.g., from S_1 to Erk in the MAPK example) that can be effectively perturbed and measured in a laboratory setting. Yet such boundaries of biological processes are quite arbitrary, are therefore highly susceptible to confounders. One way to address this issue is to search the knowledge graph for all common causes of variables in the causal model, use an identification algorithm [71] to find the minimal valid adjustment

set of the augmented model, and then prune all common causes that do not contribute to that set by allowing an exogenous variable to be shared among the two effects of the common cause. This approach will require us to tackle the issues of parameter and causal identifiability in the presence of confounders. We will also investigate the ability of counterfactual inference to improve predictions in case of model misspecification [58].

In some cases, the variables in the model may not be directly observable, but may nonetheless be characterized by means of detectable molecular signatures. For example, even if interferon signaling may not be directly observable using transcriptomics measurements, it may still be possible to infer the activity of interferon signaling by an upregulation of interferon stimulated genes (ISG). Future work will focus on leveraging molecular signature databases to infer the activity of variables in the model, and on learning and/or evaluating the models using experimental data.

ACKNOWLEDGMENTS

This work was supported by funds from the PNNL Mathematics and Artificial Reasoning Systems Laboratory Directed Research and Development Initiative. Knowledge curation environments were provided gratis by BioDati.com and Causaly.com. We would also like to acknowledge Jessica Stothers and Rose Glavin at CoronaWhy.org and Marek Ostaszewski at the COVID-19 Disease Map Initiative for providing valuable feedback about the IL6-AMP model.

REFERENCES

[1] M. Ostaszewski, A. Mazein, M. E. Gillespie, I. Kuperstein, A. Niarakis, H. Hermjakob, A. R. Pico, E. L. Willighagen, C. T. Evelo, J. Hasenauer, F. Schreiber, A. Drager, E. Demir, O. Wolkenhauer, L. I. Furlong, E. Barillot, J. Dopazo, A. Orta-Resendiz, F. Messina, A. Valencia, A. Funahashi, H. Kitano, C. Auffray, R. Balling, and R. Schneider, "COVID-19 disease map, building a computational repository of SARS-CoV-2 virus-host interaction mechanisms." *Scientific Data*, vol. 7, p. 136, 2020.

[2] L. Lu Wang, K. Lo, Y. Chandrasekhar, R. Reas, J. Yang, D. Eide, K. Funk, R. Kinney, Z. Liu, W. Merrill, P. Mooney, D. Murdick, D. Rishi, J. Sheehan, Z. Shen, B. Stilson, A. D. Wade, K. Wang, C. Wilhelm, B. Xie, D. Raymond, D. S. Weld, O. Etzioni, and S. Kohlmeier, "CORD-19: The covid-19 open research dataset." *arXiv*, 2020.

[3] U. Alon, *An Introduction to Systems Biology: Design Principles of Biological Circuits*. CRC press, 2019.

[4] P. Khatri, M. Sirota, and A. J. Butte, "Ten years of pathway analysis: Current approaches and outstanding challenges," *PLoS Computational Biology*, vol. 8, 2012.

[5] A. Liberzon, C. Birger, H. Thorvaldsdóttir, M. Ghandi, J. P. Mesirov, and P. Tamayo, "The molecular signatures database (MSigDB) hallmark gene set collection," *Cell Systems*, vol. 1, p. 417, 2015.

[6] S. Mubeen, C. T. Hoyt, A. Gemünd, M. Hofmann-Apitius, H. Fröhlich, and D. Domingo-Fernández, "The Impact of Pathway Database Choice on Statistical Enrichment Analysis and Predictive Modeling," *Frontiers in Genetics*, vol. 10, p. 654442, 2019.

[7] R.-S. Wang, A. Saadatpour, and R. Albert, "Boolean modeling in systems biology: an overview of methodology and applications," *Physical Biology*, vol. 9, p. 055001, 2012.

[8] A. Naldi, D. Berenguier, A. Fauré, F. Lopez, D. Thieffry, and C. Chaouiya, "Logical modelling of regulatory networks with GINsim 2.3," *Biosystems*, vol. 97, p. 134, 2009.

[9] R. Thomas and M. Kaufman, "Multistationarity, the basis of cell differentiation and memory. II. Logical analysis of regulatory networks in terms of feedback circuits," *Chaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 11, p. 180, 2001.

[10] A. Liu, P. Trairatphisan, E. Gjerga, A. Didangelos, J. Barratt, and J. Saez-Rodriguez, "From expression footprints to causal pathways: contextualizing large signaling networks with CARNIVAL," *Systems Biology and Applications*, vol. 5, p. 1, 2019.

[11] M. R. Hidalgo, C. Cubuk, A. Amadoz, F. Salavert, J. Carbonell-Caballero, and J. Dopazo, "High throughput estimation of functional cell activities reveals disease mechanisms and predicts relevant clinical outcomes," *Oncotarget*, vol. 8, p. 5160, 2017.

[12] A. L. Tarca, S. Draghici, P. Khatri, S. S. Hassan, P. Mittal, J. S. Kim, C. J. Kim, J. P. Kusanovic, and R. Romero, "A novel signaling pathway impact analysis," *Bioinformatics*, vol. 25, p. 75, 2009.

[13] F. Martin, T. M. Thomson, A. Sewer, D. Drubin, C. Mathis, D. Weisensee, D. Pratt, J. Hoeng, and M. C. Peitsch, "Assessment of network perturbation amplitudes by applying high-throughput data to causal biological networks," *BMC Systems Biology*, vol. 6, p. 54, 2012.

[14] F. Martin, A. Sewer, M. Talikka, Y. Xiang, J. Hoeng, and M. C. Peitsch, "Quantification of biological network perturbations for mechanistic insight and diagnostics using two-layer causal models," *BMC Bioinformatics*, vol. 15, p. 238, 2014.

[15] N. L. Catlett, A. J. Bargnesi, S. Ungerer, T. Seagaran, W. Ladd, K. O. Elliston, and D. Pratt, "Reverse causal reasoning: applying qualitative causal knowledge to the interpretation of high-throughput data," *BMC Bioinformatics*, vol. 14, p. 340, 2013.

[16] G. Bradley and S. J. Barrett, "CausalR - extracting mechanistic sense from genome scale data," *Bioinformatics*, vol. 33, p. 3670, 2017.

[17] J. Pearl, *Causality: Models, Reasoning and Inference*. Cambridge, MA, USA,, 2013.

[18] J. Peters, D. Janzing, and B. Schölkopf, *Elements of Causal Inference: Foundations and Learning Algorithms*. MIT press, 2017.

[19] J. F. Allen, M. Swift, and W. De Beaumont, "Deep semantic analysis of text," *Proceedings of the 2008 Conference on Semantics in Text Processing*, vol. 1, p. 343, 2008.

[20] D. D. McDonald, "Issues in the Representation of Real Texts: The Design of Krisp," *Natural Language Processing and Knowledge Representation*, p. 77, 2000.

[21] M. A. Valenzuela-Escárcega, O. Babur, G. Hahn-Powell, D. Bell, T. Hicks, E. Noriega-Atala, X. Wang, M. Surdeanu, E. Demir, and C. T. Morrison, "Large-scale automated machine reading discovers new cancer-driving mechanisms," *Database*, vol. 2018, p. 1, 2018.

[22] B. M. Gyori, J. A. Bachman, K. Subramanian, J. L. Muhlich, L. Galescu, and P. K. Sorger, "From word models to executable models of signaling networks using automated assembly," *Molecular Systems Biology*, vol. 13, 2017.

[23] C. T. Hoyt, D. Domingo-Fernández, R. Aldisi, L. Xu, K. Kolpeja, S. Spalek, E. Wollert, J. Bachman, B. M. Gyori, P. Greene, and M. Hofmann-Apitius, "Re-curation and rational enrichment of knowledge graphs in Biological Expression Language," *Database*, vol. 2019, 2019.

[24] E. G. Cerami, B. E. Gross, E. Demir, I. Rodchenkov, O. Babur, N. Anwar, N. Schultz, G. D. Bader, and C. Sander, "Pathway Commons, a web resource for biological pathway data," *Nucleic Acids Research*, vol. 39, p. 685, 2011.

[25] A. Fabregat, S. Jupe, L. Matthews, K. Sidiropoulos, M. Gillespie, P. Garapati *et al.*, "The Reactome pathway knowledgebase," *Nucleic Acids Research*, vol. 46, p. D649, 2018.

[26] M. Kanehisa, M. Furumichi, M. Tanabe, Y. Sato, and K. Morishima, "KEGG: New perspectives on genomes, pathways, diseases and drugs," *Nucleic Acids Research*, vol. 45, p. D353, 2017.

[27] L. Perfetto, L. Briganti, A. Calderone, A. C. Perpetuini, M. Iannuccelli, F. Langone, L. Licata, M. Marinkovic, A. Mattioni, T. Pavlidou, D. Peluso, L. L. Petrilli, S. Pirró, D. Posca, E. Santonico, A. Silvestri, F. Spada, L. Castagnoli, and G. Cesareni, "SIGNOR: A database of causal relationships between biological entities," *Nucleic Acids Research*, vol. 44, p. D548, 2016.

[28] D. N. Slenter, M. Kutmon, K. Hanspers, A. Riutta, J. Windsor, N. Nunes *et al.*, "WikiPathways: a multifaceted pathway database bridging metabolomics to other omics research." *Nucleic Acids Research*, vol. 46, p. D661, 2018.

[29] E. Demir, M. P. Cary, S. Paley, K. Fukuda, C. Lemer, I. Vastrik *et al.*, "The BioPAX community standard for pathway data sharing," *Nature Biotechnology*, vol. 28, p. 1308, 2010.

[30] M. Hucka, F. T. Bergmann, A. Dräger, S. Hoops, S. M. Keating, N. Le Novère, C. J. Myers, B. G. Olivier, S. Sahle *et al.*, "The Systems Biology Markup Language (SBML): language specification for level 3 version 2 core," *Journal of Integrative Bioinformatics*, vol. 15, 2018.

[31] N. Le Novere, M. Hucka, H. Mi, S. Moodie, F. Schreiber, A. Sorokin, E. Demir, K. Wegner, M. I. Aladjem, S. M. Wimalaratne *et al.*, "The systems biology graphical notation," *Nature Biotechnology*, vol. 27, p. 735, 2009.

[32] T. Slater, "Recent advances in modeling languages for pathway maps and computable biological networks," *Drug Discovery Today*, vol. 19, p. 193, 2014.

[33] T. Helikar, B. Kowal, S. McClenathan, M. Bruckner, T. Rowley, A. Madrahimov, B. Wicks, M. Shrestha, K. Limbu, and J. A. Rogers, "The cell collective: toward an open and collaborative approach to systems biology." *BMC Systems Biology*, vol. 6, p. 96, 2012.

[34] M. Gündel, C. T. Hoyt, and M. Hofmann-Apitius, "BEL2ABM: Agent-based simulation of static models in Biological Expression Language," *Bioinformatics*, vol. 34, p. 2316, 2018.

[35] Y. Li, J. Roberts, Z. AkhavanAghdam, and N. Hao, "Mitogen-activated protein kinase (MAPK) dynamics determine cell fate in the yeast mating response," *The Journal of Biological Chemistry*, vol. 292, p. 20354, 2017.

[36] L. Chen, R. Wang, C. Li, and K. Aihara, *Modeling Biomolecular Networks in Cells: Structures and Dynamics*. Springer Science & Business Media, 2010.

[37] D. Gratie, B. Iancu, and I. Petre, "ODE analysis of biological systems," in *International School on Formal Methods for the Design of Computer, Communication and Software Systems*, 2013, p. 29.

[38] F. Bianconi, E. Baldelli, V. Ludovini, L. Crino, A. Flacco, and P. Valigi, "Computational model of EGFR and IGF1R pathways in lung cancer: a systems biology approach for translational oncology," *Biotechnology Advances*, vol. 30, p. 142, 2012.

[39] E. K. Kim and E.-J. Choi, "Pathological roles of MAPK signaling pathways in human diseases," *Biochimica et Biophysica Acta - Molecular Basis of Disease*, vol. 1802, p. 396, 2010.

[40] R. Ness, K. Paneri, and O. Vitek, "Integrating Markov processes with structural causal modeling enables counterfactual inference in complex systems," in *Advances in Neural Information Processing Systems*, 2019, p. 14211.

[41] K. Paneri, "Integrating markov process and structural causal models enables counterfactual inference in complex systems," 2019.

[42] D. T. Gillespie, "Exact stochastic simulation of coupled chemical reactions," *The Journal of Physical Chemistry*, vol. 81, p. 2340, 1977.

[43] S. K. Jha and C. J. Langmead, "Exploring behaviors of stochastic differential equation models of biological systems using change of measures," *BMC bioinformatics*, vol. 13, p. S8, 2012.

[44] S. Bongers and J. M. Mooij, "From random differential equations to structural causal models: the stochastic case," in *Proceedings of Uncertainty in Artificial Intelligence*, 2019.

[45] J. Pearl, "Causal diagrams for empirical research," *Biometrika*, vol. 82, p. 669, 1995.

[46] M. Jerrum, A. Sinclair, and D. S. Hochbaum, "The Markov chain Monte Carlo method," *Approximation Algorithms for NP-hard Problems*, 1997.

[47] A. E. Gelfand, "Gibbs sampling," *Journal of the American statistical Association*, vol. 95, p. 1300, 2000.

[48] M. D. Hoffman and A. Gelman, "The No-U-Turn sampler: adaptively setting path lengths in Hamiltonian Monte Carlo," *Journal of Machine Learning Research*, vol. 15, p. 1593, 2014.

[49] M. D. Hoffman, D. M. Blei, C. Wang, and J. Paisley, "Stochastic variational inference," *The Journal of Machine Learning Research*, vol. 14, p. 1303, 2013.

[50] T. Blom, S. Bongers, and J. M. Mooij, "Beyond structural causal models: Causal constraints models," in *Proceedings of the 35th Conference on Uncertainty in Artificial Intelligence*, 2019.

[51] S. Madan, J. Szostak, R. Komandur Elayavilli, R. T.-H. Tsai, M. Ali, L. Qian, M. Rastegar-Mojarad, J. Hoeng, and J. Fluck, "The extraction of complex relationships and their conversion to biological expression language (BEL) overview of the BioCreative VI (2017) BEL track." *Database: the Journal of Biological Databases and Curation*, vol. 2019, 2019.

[52] C. T. Hoyt, D. Domingo-Fernández, S. Mubeen, J. M. Llaó, A. Konotopez, C. Ebeling, C. Birkenbihl, O. Muslu, B. English, S. Müller, M. P. de Lacerda, M. Ali, S. Colby, D. Türei, N. Palacio-Escat, and M. Hofmann-Apitius, "Integration of structured biological data sources using biological expression language," *BioRxiv*, 2019.

[53] C. T. Hoyt, A. Konotopez, C. Ebeling, and J. Wren, "PyBEL: a computational framework for biological expression language." *Bioinformatics*, vol. 34, p. 703, 2018.

[54] P. Boutillier, M. Maasha, X. Li, H. F. Medina-Abarca, J. Krivine, J. Feret, I. Cristescu, A. G. Forbes, and W. Fontana, "The Kappa platform for rule-based modeling," *Bioinformatics*, vol. 34, p. i583, 2018.

[55] L. A. Harris, J. S. Hogg, J.-J. Tapia, J. A. P. Sekar, S. Gupta, I. Korsunsky, A. Arora, D. Barua, R. P. Sheehan, and J. R. Faeder, "BioNetGen 2.2: advances in rule-based modeling." *Bioinformatics*, vol. 32, p. 3366, 2016.

[56] H. Mi, F. Schreiber, S. Moodie, T. Czauderna, E. Demir, R. Haw, A. Luna, N. Le Novère, A. Sorokin, and A. Villéger, "Systems Biology Graphical Notation: Activity Flow language Level 1 Version 1.2." *Journal of Integrative Bioinformatics*, vol. 12, p. 265, 2015.

[57] D. J. Rezende, S. Mohamed, and D. Wierstra, "Stochastic backpropagation and approximate inference in deep generative models," *arXiv preprint arXiv:1401.4082*, 2014.

[58] R. Ness, K. Paneri, and O. Vitek, "Integrating Markov processes with structural causal modeling enables counterfactual inference in complex systems," in *Advances in Neural Information Processing Systems*, 2019, p. 14234.

[59] "BioDati, inc." [Online]. Available: https://studio.covid19.biodati.com/

[60] E. Bingham, J. P. Chen, M. Jankowiak, F. Obermeyer, N. Pradhan, T. Karaletsos, R. Singh, P. Szerlip, P. Horsfall, and N. D. Goodman, "Pyro: Deep Universal Probabilistic Programming," *Journal of Machine Learning Research*, 2018.

[61] Z. S. Ulhaq and G. V. Soraya, "Interleukin-6 as a potential biomarker of COVID-19 progression." *Medecine et Maladies Infectieuses*, vol. 50, p. 382, 2020.

[62] T. Hirano and M. Murakami, "COVID-19: A new virus, but a familiar receptor and cytokine release syndrome." *Immunity*, vol. 52, p. 731, 2020.

[63] M. Murakami and T. Hirano, "The pathological and physiological roles of IL-6 amplifier activation." *International Journal of Biological Sciences*, vol. 8, p. 1267, 2012.

[64] H. Ogura, M. Murakami, Y. Okuyama, M. Tsuruoka, C. Kitabayashi, M. Kanamoto, M. Nishihara, Y. Iwakura, and T. Hirano, "Interleukin-17 promotes autoimmunity by triggering a positive-feedback loop via interleukin-6 induction," *Immunity*, vol. 29, p. 628, 2008.

[65] V. Oldfield, S. Dhillon, and G. L. Plosker, "Tocilizumab: a review of its use in the management of rheumatoid arthritis." *Drugs*, vol. 69, p. 609, 2009.

[66] C. Zhang, Z. Wu, J.-W. Li, H. Zhao, and G.-Q. Wang, "Cytokine release syndrome in severe COVID-19: Interleukin-6 receptor antagonist Tocilizumab may be the key to reduce mortality," *International Journal of Antimicrobial Agents*, vol. 55, p. 105954, 2020.

[67] E. A. Coomes and H. Haghbayan, "Interleukin-6 in COVID-19: A systematic review and meta-analysis," *medRxiv*, 2020.

[68] X. Xu, M. Han, T. Li, W. Sun, D. Wang, B. Fu, Y. Zhou, X. Zheng, X. L. Y. Yang, X. Zhang, A. Pan, and H. Wei, "Effective Treatment of Severe COVID - 19 Patients with Tocilizumab," *PNAS*, vol. 117, p. 10970, 2020.

[69] K. E. R. Soetaert, T. Petzoldt, and R. W. Setzer, "Solving differential equations in R: package deSolve," *Journal of Statistical Software*, vol. 33, 2010.

[70] D. Wilkinson, "Package smfsb," 2018.

[71] S. Tikka and J. Karvanen, "Identifying causal effects with theR packagecausaleffect," *Journal of Statistical Software*, vol. 76, p. 1, 2017.