scientific reports

OPEN

Proof of concept for quantitative adverse outcome pathway modeling of chronic toxicity in repeated exposure

Shigeaki Ito[1,3]✉, Sayak Mukherjee[2,3], Kazuo Erami[1], Shugo Muratani[1], Akina Mori[1], Sakuya Ichikawa[1], William White[2], Kei Yoshino[1] & Dawn Fallacara[2]

Adverse Outcome Pathway (AOP) is a useful tool to glean mode of action (MOE) of a chemical. However, in order to use it for the purpose of risk assessment, an AOP needs to be quantified using in vitro or in vivo data. Majority of quantitative AOPs developed so far, were for single exposure to progressively higher doses. Limited attempts were made to include time in the modeling. Here as a proof-of concept, we developed a hypothetical AOP, and quantified it using a virtual dataset for six repeated exposures using a Bayesian Network Analysis (BN) framework. The virtual data was generated using realistic assumptions. Effects of each exposure were analyzed separately using a static BN model and analyzed in combination using a dynamic BN (DBN) model. Our work shows that the DBN model can be used to calculate the probability of adverse outcome when other upstream KEs were observed earlier. These probabilities can help in identification of early indicators of AO. In addition, we also developed a data driven AOP pruning technique using a lasso-based subset selection, and show that the causal structure of AOP is itself dynamic and changes over time. This proof-of-concept study revealed the possibility for expanding the applicability of the AOP framework to incorporate biological dynamism in toxicity appearance by repeated insults.

Abbreviations

AOP	Adverse outcome pathway
NAMs	New approach methodologies
BN	Bayesian network
DBN	Dynamic Bayesian network
GBN	Gaussian Bayesian network
DAG	Directed acyclic graph
MIE	Molecular initiating event
KE	Key event
AO	Adverse outcome
BM	Biomarker
LOOCV	Leave-one-out cross validation
VUS	Volume under surface

Regulatory agencies such as the U.S. Environmental Protection Agency (EPA), the U.S. Food and Drug Administration (FDA), European Chemical Agency (ECHA), and others are adopting the use of new approach methodologies (NAMs) to reduce animal use as part of their 3Rs (refinement, reduction, and replacement) agenda[1]. Fit-for-purpose in vitro NAMs are now used to interrogate a "system" at different layers of biological organization, from molecular to phenotypical. Integrating all these information together into a cohesive predictive model of an adverse outcome however is not so easy.

One solution to this problem is the adverse outcome pathway (AOP)[2,3] framework. AOP framework proposes a simplification of holistic and systemic approaches where an apical adverse outcome (AO) is causally linked to

[1]Scientific Product Assessment Center, Japan Tobacco Inc., 6-2, Umegaoka, Aoba-ku, Yokohama, Kanagawa 227-8512, Japan. [2]Battelle, 505 King Ave., Columbus, OH 43201, USA. [3]These authors contributed equally: Shigeaki Ito and Sayak Mukherjee ✉email: shigeaki.ito@jt.com

an early molecular initiating event (MIE), like a receptor binding or production of reactive oxygen species, by subsuming hundreds and thousands of intermediate biological processes, from molecular, cellular, to organism level, into a handful of causally linked intermediate key events (KEs).

Although the AOP framework is expected to be a crucial tool for regulatory decision-making, given the qualitative nature of AOPs, direct application in quantitative risk assessment is challenging. This limitation highlights the need for quantitative AOP (qAOP) to bridge the gap between qualitative AOPs and the quantitative requirements of risk assessment. To circumvent this difficulty, several mathematical approaches were proposed and reviewed to render AOPs quantitative (qAOP) starting from simple dose–response modeling to more complicated Bayesian network modeling and differential equation (DE) based Systems Biology approaches[4–6]. Bayesian network (BN) formalism gained popularity as it can harmonize different types of data, provide a robust paradigm for causal modeling, and can be used prospectively for exploring multiple hypothesis. BNs have been used in studies spanning reproductive toxicity, developmental neural toxicity, cardiotoxicity, kidney injury etc.[4,5,7–12] (the others can be found in Spinu's article in 2020). In particular, Zgheib et al. reported several options for qAOP modeling approach for chronic toxicity, in which they applied Dynamic Bayesian Network (DBN) approach for chronic kidney disease by adaptation of time-series in vitro data into the qAOP model[13]. However, appearance of chronic toxicity/disease is often more complex, especially those associated with phenotypical changes in tissues or organ level from chronic exposure to low doses of chemical. For example, very low-level exposure to cadmium is not a serious concern as cadmium is contained in food to some extent. However, repeated (chronic) exposure to even low dose of cadmium can cause severe health impacts[14]. Cigarette smoking is a known risk factor for chronic diseases, but a single puff is hardly a health risk. As such, while an adverse outcome typically arise following the rapid biological responses triggered by certain reactive chemicals, there are the others that require cumulative biological reactions elicited by repeated and chronical exposure.

With advancement of in vitro NAMs, it is now feasible to perform experiments that can probe chronic toxic endpoints from repeated exposures[15–19]. Such technological innovations make it possible to expand the scope of the AOP framework: AOP which consist of acute phase responses and chronic phase response could be now realistic application for risk assessment strategy if appropriate risk quantification method is adopted.

Here, as a proof of concept, we developed a "hypothetical" AOP that has two distinct but interconnected modules, one for acute phase KEs and the other of chronic phase KEs (Fig. 1). We assumed that the chronic phase KEs can be elicited only after repeated/chronic exposures to an agent. Furthermore, in the context of repeated exposures, there is individual variation in the timing of manifestation of chronic toxicity, much like in vivo, where the onset of chronic toxicity may vary among donors of primary cells even under the same repeated exposure. To reflect this, in this proof-of-concept study, the timing of chronic toxicity appearance is varied for each donor. We used virtual data—generated using realistic assumptions—for this proof-of-concept study and quantified the AOP using a combination of static and dynamic Bayesian network formalisms (BN and DBN respectively). We show that the probabilities of adverse outcome can be estimated based on activation of upstream KEs earlier. We also use a data driven AOP pruning approach to quantify the dynamic nature of causal links in an AOP, as well as how these links evolve as a function of repeated insults.
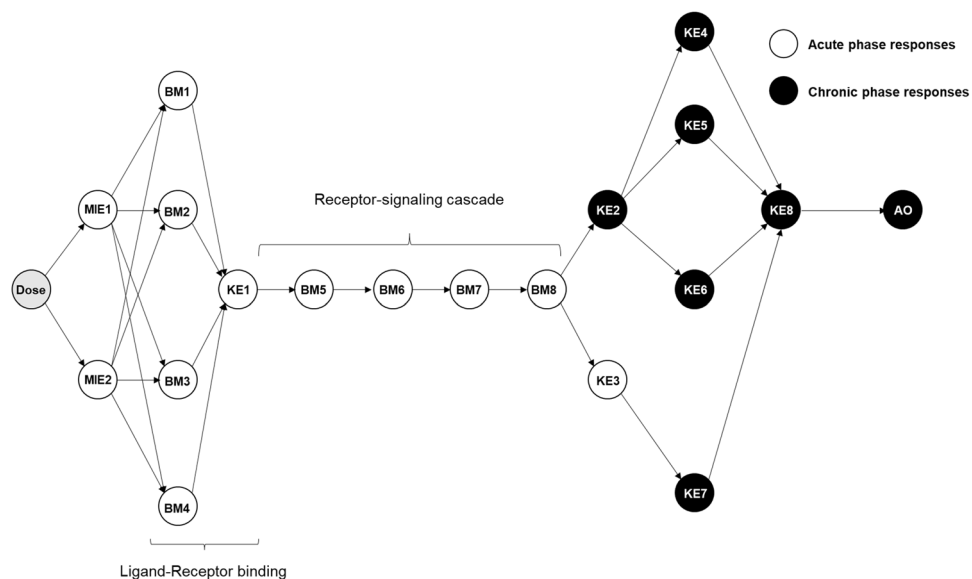

Figure 1. Directed acyclic graph of hypothetical AOP. Directed acyclic graph was used for Bayesian statistics-based approaches. BM1 through BM4 are ligands of the receptor (KE1). KE1 induces receptor signaling cascades of BM5 through BM8, which results in the downstream causal responses including chronic phase KEs. MIEs, acute-phase KEs, and acute-phase BMs are represented in white. Chronic phase KEs and AO are shown in solid black circles. MIE; Molecular initiating event, KE; Key event, BM; Biomarker, AO; adverse outcome.

Methods

Notation convention

Exposure repetition is indexed by $e = 1, 2, ., ., ., E$. Donors are denoted by $n = 1, 2, ., ., ., N$. The dose (i.e., exposure concentration) is denoted by $d \in D$, where D is the set of all doses including non-treatment control. MIEs, KEs, BMs, and AO are collectively referred to as nodes and denoted by $v \in V$, where V is the set of all nodes. The total number of elements in a set is denoted by $|..|$; $|V|$ for example, stands for total number of nodes. Replicates for each donor will be denoted by $r = 1, 2, ., ., ., \mathcal{R}$. Matrices are written in double struck upper case letters: $\mathbb{X}_{M \times |V|}$ is a matrix with M rows and $|V|$ columns. Furthermore, a symbol such as $\mathbb{X}^{(e)}$ means a matrix $\mathbb{X}$ for a specific exposure repetition e. Similarly, $\mathbb{X}^{(e,n)}$ is understood as a matrix for a specific exposure repetition e and a donor n. Elements of a matrix and scalars in general are written in lower-case fonts. A vector is always a column vector and denoted in bold: $\boldsymbol{\beta} \in R^P$ is a vector with P components. A one vector, $1_{P \times 1} \in R^P$, is a column vector of P ones. Transpose of a matrix or a vector is denoted by the capital letter T in the superscript. A node v is assumed to take continuous values and in Bayesian formalism, the probability that it takes a value x depends only on the values of its parents, i.e. $P(x^{(v)} | \boldsymbol{x} \in R^V) = P(x^{(v)} | \boldsymbol{x}^{(\Pi_v)} \in R^{|\Pi_v|})$, where Π_v denotes the set of parents of node v.

Software and algorithms

We used Microsoft Excel for generation of the primary dataset. All the other analyses and visualization were performed using R statistics software version 4.0.4. The algorithms and used R packages are summarized in the Supplementary material 1.

Structure of hypothetical AOP

The network structure is shown in Fig. 1. To model chronic toxicity, we constructed a hypothetical AOP with two MIEs, two acute-phase KEs, eight acute-phase biomarkers (BMs), six chronic-phase KEs, and an AO. Activation of some biomarkers, such as ligands for receptor-ligands binding, and activation of downstream signaling molecules for signal transduction are crucial for elicitation of AO. We included them in our AOP and denoted them as BMs. While BMs are usually not included in AOPs, we included them in this proof-of-concept study to highlight the fact that additional assays that capture relevant endpoints upstream of or related to a specific KE are usually obtained in a lab. For example, one may collect activation of biomarkers in MAP kinase pathway upstream of any apoptotic KE. Moreover, from a modeling point of view, any number of BMs can be included in the modeling if their wiring does not violate the fundamental assumption of Bayesian Network analysis, namely generation of feedback loops and cycles. To keep the discussion as general as possible, biomarkers that lead to KEs are also included in the model.

Specifically, BM1 through BM4 mimic ligands for KE1 (e.g., receptor activation), and BM5 through BM8 represents the downstream signal transduction cascade. KE2 and KE3 represent the products of signal transduction, but KE2 only appears in chronic phase, while KE3 appears upon every exposure. Once chronic-phase response appears, a series of chronic-phase responses simultaneously appear as a chain reaction. We have collectively referred to them as nodes (19 nodes in total). In this AOP, we assumed that twelve biological events including MIEs, acute-phase KEs and biomarkers were elicited in a dose dependent manner for every exposure repetition. Chronic-phase KEs were also elicited in dose dependent manner but only after being subjected to a given number of exposure repetition. Specific timing of a chronic phase KE varied across donors (Table 1).

Primary virtual date generation

For this proof-of-concept study, we used a fully virtual dataset rather than actual data, because there are hardly any available data appropriate for chronic toxicity related repeated exposure based qAOP modeling.

The assumptions and parameters used for the primary virtual dataset are:

i. Four doses including non-treated control ($|D| = 4$), total number of donors $N = 8$, and the number of exposures $E = 6$
ii. Acute-phase biological responses (i.e., MIEs, KE1, KE3, and BM1 through BM8) show robust dose-dependence for all exposures.

Donor	Appearance of chronic toxicity	Severity
1	Exposure rep 4	Severe
2	Exposure rep 3	Severe
3	Exposure rep 4	Moderate
4	Exposure rep 3	Moderate
5	Exposure rep 4	Weak
6	Exposure rep 3	Weak
7	No	No
8	No	No

Table 1. Assumption of donor dependent appearance of chronic toxicity.

iii. The timing of appearance of chronic-phase responses (i.e., KE2 and KE4 through KE8, and AO) are donor-dependent (Table 1).

iv. Once elicited, chronic phasae responses also increase in a dose-dependent manner. They also exhibit a robust exposure-repetition dependence upon elicitation.

v. Denoting the replicate average and standard deviation of fold change at a dose d by $\overline{f}^{(d)}$ and $s^{(d)}$, we assumed that $s^{(d+1)}$ is affected by the difference in $\overline{f}^{(d)}$ and $\overline{f}^{(d+1)}$: larger the difference, larger is $s^{(d+1)}$.

vi. The fold-changes follow a log-normal distribution.

First, the *mean fold-changes* for all nodes for $d = 0$ (air control) were set to 1, i.e., $\overline{f}^{(v,d=0,n)} = 1 \forall v, n$. Mean fold-changes for subsequent doses and exposures for all *acute phase nodes* (MIE, KE1, KE3, and all BMs) were then generated using the formulae $\overline{f}^{(d+1,v,n,e)} = \overline{f}^{(d,v,n,e)} + \zeta_d$ and $\overline{f}^{(e+1,v,n,d)} = \overline{f}^{(e,v,n,d)} + \zeta_e$ respectively, where $\overline{f}^{(d,e,n,v)}$ stands for mean fold-change of node v of donor n at dose d, exposure e, and $\zeta_d \in U(0,1)$ and $\zeta_e \in U(0,0.2)$ are uniform random numbers. Denoting onset time of a chronic response for a donor n by e_{ch}^n (Table 1), mean fold change for chronic phase KEs were generated using the formulae above for exposure repetitions $e > e_{ch}^n$. For $e < e_{ch}^n$, the mean fold changes for the chronic phase KEs were randomly populated using the formula $0.8 + 0.4\zeta \in U(0,1)$, thereby making sure that they were not elicited. Similarly, standard deviation of acute phase biomarkers for all exposures and chronic phase biomarkers for $e > e_{ch}^n$ were generated using the formula $s^{(d+1,v,n,e)} = \left(\overline{f}^{(d,v,n,e)} + \zeta \times \left(\overline{f}^{(d+1,v,n,e)} - \overline{f}^{(d,v,n,e)}\right)\right)/(4 + 6\zeta)$, where $s^{(d+1)}$ is the standard deviation of fold-change for that dose.

The mean fold changes of each node as functions of dose and exposure repetition for all $N = 8$ donors are summarized in Supplementary material 2.

Consideration of correlation co-efficient for resampling the primary dataset

For each donor, dose, and exposure repetition, we log transformed the mean fold changes and constructed the *mean log-fold change* vector $\boldsymbol{\mu}^{(n,e,d)} \in R^{|V|}$ such that its components $\mu^{(n,e,d,v)}$, $v = 1, 2, ., |V|$, were given by the formula

$$\mu^{(n,e,d,v)} = ln\left(\frac{\left(\overline{f}^{(n,e,d,v)}\right)^2}{\sqrt{\left(\overline{f}^{(n,e,d,v)}\right)^2 + (s^{(n,e,d,v)})^2}}\right). \tag{1}$$

We know that biological responses elicited by a common stimulus are usually correlated with each other, especially when they show clear dose responses. Therefore, we accounted for response-response correlation at the donor level for each exposure repetition. Correlations were calculated from the primary dataset. The elements of response-response covariance matrix $\mathbb{O}_{|V|\times|V|}^{(n,e,d)}$, were given by

$$o_{vv'}^{(n,e,d)} = \left(\sigma^{(n,e,d,v)}\right)^2 = ln\left(1 + \frac{\left(s^{(n,e,d,v)}\right)^2}{\left(\overline{f}^{(n,e,d,v)}\right)^2}\right) for v = v'$$

$$o_{vv'}^{(n,e,d)} = \rho_{vv'}^{(e)} \sqrt{o_{vv}^{(n,e,d)} * o_{v'v'}^{(n,e,d)}} for v \neq v' \tag{2}$$

where the correlation co-efficient $\rho_{vv'}^{(e)} = \sigma_{vv'}^{(e)}/\sqrt{\sigma_{vv}^{(e)} \times \sigma_{v'v'}^{(e)}}$.

Using Eq. (1) and Eq. (2), we drew samples from a multivariate normal distribution $\mathcal{MN}\left(\boldsymbol{\mu}^{(n,e,d)}, \mathbb{O}_{|V|\times|V|}^{(n,e,d)}\right)$ for each donor, exposure repetition, and dose.

Bayesian updating approach for resampling the primary dataset

In addition, we considered the possibility that the data at a previous exposure repetition could be informative for the next exposure repetition. To account for that, we explicitly introduced causation between each node with a Gaussian Bayesian network (GBN) update scheme.

Commonly, parameter learning of a GBNs entails maximizing a Gaussian likelihood of the form $\mathcal{MN}\left(\boldsymbol{x}^{(v)}|\mathbb{X}_{M\times|\Pi_v|}\boldsymbol{\beta} + \beta_0, \sigma^2\mathbb{I}\right)$, where Π_v is the parent set of a node v, M is the number of samples, σ^2 is the residual, and $(\boldsymbol{\beta}, \beta_0)$ are the parameters to be estimated. Maximum Likelihood Estimate (MLE) for a Gaussian likelihood is simply $\widehat{\boldsymbol{\beta}}_{OLS} = \left(\mathbb{X}_{M\times\Pi_v}{}^T\mathbb{X}_{M\times\Pi_v}\right)^{-1}\mathbb{X}_{M\times\Pi_v}{}^T\boldsymbol{x}^{(v)}$ (ordinary linear regression estimates). Alternatively, we can estimate $\boldsymbol{\beta}$ by using a Gaussian prior of the form $\mathcal{MN}\left(\boldsymbol{\beta}|\boldsymbol{\beta}_{pr}, \mathbb{V}_{pr}\right)$. When σ^2 is known, the posterior of $\boldsymbol{\beta}$ is a multivariate Gaussian $\mathcal{MN}\left(\boldsymbol{\beta}|\widehat{\boldsymbol{\beta}}_{pos}, \mathbb{V}_{pos}\right)$, where

$$\widehat{\boldsymbol{\beta}}_{pos} = \mathbb{V}_{pos}\left(\mathbb{V}_{pr}{}^{-1}\boldsymbol{\beta}_{pr} + \frac{1}{\sigma^2}\mathbb{X}_{M\times\Pi_v}{}^T\boldsymbol{x}_v\right)$$

$$\mathbb{V}_{pos}^{-1} = \mathbb{V}_{pr}{}^{-1} + \frac{1}{\sigma^2}\mathbb{X}_{M\times\Pi_v}{}^T\mathbb{X}_{M\times\Pi_v} \tag{3}$$

First, for each exposure e, dose d and donor n, we drew κ samples $\mathbb{X}_{\kappa\times|V|}^{(n,e,d)} \sim \mathcal{MN}\left(\boldsymbol{\mu}^{(n,e,d)}, \mathbb{O}_{|V|\times|V|}^{(n,e,d)}\right)$. Then, we learned the parameters $\widehat{\boldsymbol{\beta}}_{pos}^{(n,e,d)}$ and $\mathbb{V}^{(n,e,d)}{}_{pos}$ iteratively using $\mathbb{X}_{\kappa\times\Pi_v}$ (relevant sub-matrix of $\mathbb{X}_{\kappa\times|V|}^{(n,e,d)}$) and

Eq. (3) by replacing $\boldsymbol{\beta_{pr}}$ and $\mathbb{V}_{pr}$ in Eq. (3) with $\widehat{\boldsymbol{\beta}}_{pos}^{(n,e-1,d)}$ and $\mathbb{V}^{(n,e-1,d)}{}_{pos}$ respectively. For the $e = 0$, OLS estimate for $\boldsymbol{\beta}$ was used. Usually, σ^2 is unknown. Prior over σ^2 (Inverse Wishart distribution) can be used and standard result can be applied. Instead, we pretended to know σ^2 by setting it to the ordinary least square estimate

$$\left(\widehat{\sigma}^{(n,e,d)}\right)^2 = \frac{\boldsymbol{x}^{(v)T}\left(\boldsymbol{x}^{(v)} - \mathbb{X}_{\kappa \times |\Pi_v|}^{(n,e,d)}\widehat{\boldsymbol{\beta}}_{OLS}^{(e)}\right)}{\kappa - |\Pi_v|}. \tag{4}$$

Please note that we explicitly used the causal structure of the AOP (Fig. 1) for estimating these parameters.

Subsequently, we constructed a virtual dataset $\mathbb{X}_{\mathcal{R} \times |V|}^{(n,e,d)}$ by resampling from this trained network using the standard formula $P\left(\boldsymbol{x}^{(v)}{}_{pred}|\mathbb{X}^{(test)}, \sigma^2\mathbb{I}\right) = \mathcal{MN}\left(\boldsymbol{x}^{(v)}{}_{pred}|\mathbb{X}^{(test)}\widehat{\boldsymbol{\beta}}_{pos}, \sigma^2\mathbb{I} + \mathbb{X}^{(test)}\mathbb{V}_{pos}\left(\mathbb{X}^{(test)}\right)^T\right)$. Replicate averaged virtual data $\boldsymbol{x}^{(n,e,d)} = \left(\left(X_{\mathcal{R} \times |V|}^{(n,e,d)}\right)^T 1\right)/\mathcal{R} \in R^V$ was used for all downstream analysis.

Dose–response and response-response analysis
We performed dose–response analysis using a linear mixed model (LMM) with a donor specific random effect of the form

$$\boldsymbol{x}^{(v,n,e)} = \beta_0 1_{|D| \times 1} + \beta_1 \boldsymbol{d}_{|D| \times 1} + \gamma^{(n)} 1_{|D| \times 1} + \varepsilon, \tag{5}$$

where $\gamma^{(n)} \sim \mathcal{N}\left(0, \sigma_n^2\right)$ is the random effect on the intercept. $\boldsymbol{d}$ is the vector of doses, and $\boldsymbol{x}^{(v,n,e)}$ is a replicate averaged fold-change vector for a given node v, donor n, and exposure repetition e.

For response-response relationships, we used two models, a) LMM model with donor specific random effect on the intercept, and b) a cubic polynomial model of forms

$$\left.\begin{array}{l} \boldsymbol{x}^{(v,n)} = \left(\beta_0 + \gamma^{(n)}\right)1_{|D|*E \times 1} + \beta_1 \boldsymbol{x}^{(v',n)} + \varepsilon \\ \boldsymbol{x}^{(v)} = \beta_0 1_{|D|*E*N \times 1} + \beta_1 \boldsymbol{x}^{(v')} + \beta_2\left[\boldsymbol{x}^{(v')}\right]^2 + \beta_2\left[\boldsymbol{x}^{(v')}\right]^3 + \varepsilon \end{array}\right\}, \tag{6}$$

where $\left[\boldsymbol{x}^{(v')}\right]^2, \left[\boldsymbol{x}^{(v')}\right]^3$ imply element wise exponentiation. Owing to large number of possible pairs $\left(v, v'\right)$, we did not analyze the data separately for each exposure repetition, instead analyzed the combined dataset for all exposure repetition.

Bayesian modeling and probability calculation of AO
Here, we looked at each exposure repetition separately. The underlying assumption here is that the time gap between successive exposures is much longer compared to the time scale associated with the emergence of the AO. Clearly, for chronic conditions, this assumption is not true. Yet, we modelled this hypothetical scenario as a proof-of-concept for acute conditions where such an assumption may find more merit. We used a GBN—a directed acyclic graph (DAG)$\mathcal{G} = (V, E)$, where V is the set of nodes and $\mathcal{E}$ is the set of edges between nodes in V. A DAG can be topologically ordered such that the edges point from topological ancestors to their successors. In Bayesian formalism, the joint probability of a realization $\boldsymbol{x} \in R^{|V|}$ can be factorized as follows

$$P(\boldsymbol{x}) = \prod_{v=1}^{|V|} P\left(x^{(v)}|\boldsymbol{x}^{(\Pi_v)}\right), \tag{7}$$

where Π_v is the set of parents of node v, and $\boldsymbol{x}^{(\Pi_v)}$ is the value that they take. The likelihood $P\left(x^{(v)}|\boldsymbol{x}^{(\Pi_v)}\right)$ is given by $P\left(x^{(v)}|\boldsymbol{x}^{(\Pi_v)}\right) = \mathcal{N}\left(x^{(v)}|\boldsymbol{x}^{(\Pi_v)T}\boldsymbol{\beta}, \sigma^2\right)$ where unknown parameter $\boldsymbol{\beta}$ is learned by fitting a linear regression model[20,21].

The DAG of the hypothetical AOP is shown in Fig. 1. Dose was included as root node without parent nodes. Dose was assumed to be a continuous variable. The parameters of the model were learned by fitting the AOP to $\mathbb{X}^{(e)}$ (data for each exposure repetition) using a function from the bnlearn package[22] called bn.fit. The fitted parameter values are summarized in Supplementary tables 1. The trained GBNs were resampled to estimate the probability of activation of KEs ($P(KE > \Delta | d \in [\widehat{d} \pm \varepsilon])$) for each exposure repetition using logic sampling. A typical $\Delta = log_{10}2$ was chosen. We also varied Δ and the dose d to investigate how this probability changed as a function of d and Δ. We also calculated the volume under the probability surface (VUS) using standard trapezoid rule.

Dynamic Bayesian modeling and probability of adverse outcome conditioned on activation of upstream chronic events at earlier exposure repetitions.
A dynamic Bayesian model (DBN) was developed to estimate transition probabilities of AO by conditioning it on the upstream events at a previous exposure repetition ($e - \tau$) using a multivariate Markov process of the form[23].

$$\boldsymbol{x}^{(e,v)} = \mathbb{X}^{(e-1,\Pi_v)}\boldsymbol{\beta}^{(e,v)} + \boldsymbol{\varepsilon}^{(e,v)} \sim \mathcal{N}\left(0, \sigma^2\mathbb{I}\right) \tag{8}$$

Diagrammatic representation of our logic and structure of DBNs are shown in Fig. 2. In brief, we assumed that events at a previous exposure causally influenced events at the current exposure. For example, amplitude of
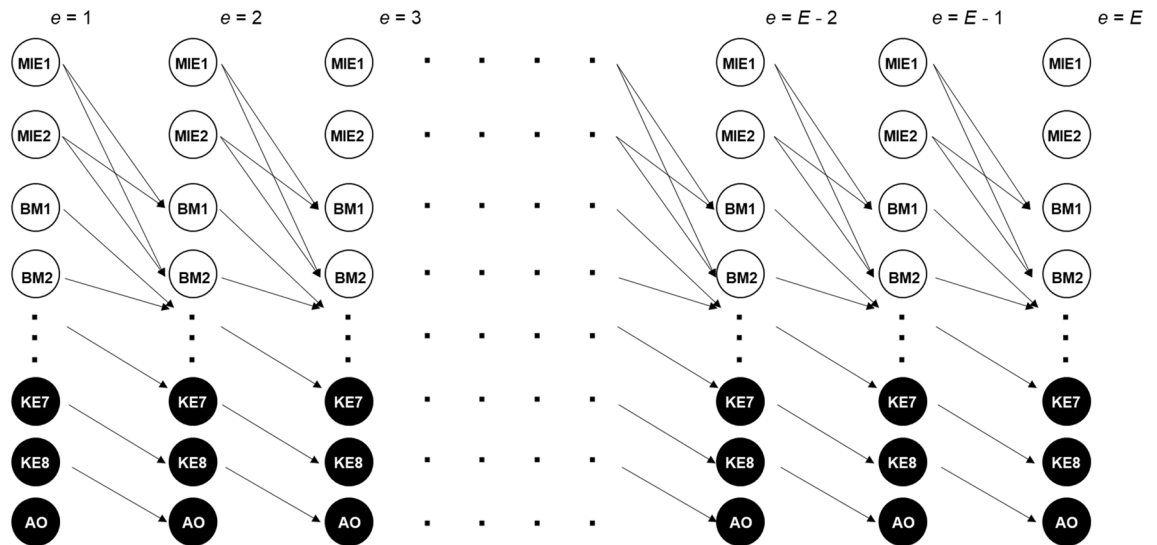
Figure 2. Diagrammatic representation of Dynamic Bayesian Network of the AOP. The nodes are causally connected across exposure repetitions, where a parent node at a previous exposure can influence the child node at a later exposure. We did not include any causal link between nodes at a given exposure. We assumed a first-order Markov process $x^{(e)} \leftarrow x^{(e-1)}$ and did not assume stationarity.

KE6 at $e = 4$ was causally influenced by its parent node, i.e., amplitude of KE2 at $e = 3$, who was in turn influenced by its parent node, BM8, at $e = 2$, so and so forth. We further assumed that no causal links exist among nodes for the same exposure repetition. Formally, nodes at a given exposure e were conditionally independent of each other when their values at the previous exposure were known. Furthermore, we have assumed that values at present were only conditionally determined by values in the immediate past, i.e., $x^{(e)} \leftarrow x^{(e-1)}$ (Markov process), but not by remote pasts $e - 2, e - 3$, etc. The last two assumptions were not essential, but they were necessary to avoid number of parameters from exploding and thereby rendering parameter learning unfeasible. In addition, dose was not included as a root node in this framework. We assumed that dose-MIE interaction, which for a vast majority of AOPs are ligand-receptor types of interaction, occurs at a timescale much faster compared to the time-gap between successive exposures. As a result, the dose was treated as a discrete variable, and subsequently, we either pooled all the doses together or examined each of the three doses separately.

Although time-series analysis typically assumes that the process in weakly stationary, we relaxed weak stationarity criteria as evidenced by the fact that our $\boldsymbol{\beta}$ explicitly depended on the exposure repetition index e (see Eq. 8) and needed to be learned for each pair of exposure slices (for example , $\boldsymbol{\beta}^{(e=3)} \neq \boldsymbol{\beta}^{(e=5)}$: $\boldsymbol{\beta}^{(e=3)}$ was estimated using data from exposure slice $(e = 2, e = 3)$, whereas $\boldsymbol{\beta}^{(e=5)}$ was estimated using data from exposure slice $(e = 4, e = 5)$). Moreover, we also observed very strong correlations between chronic phase responses (see Result section), which made straightforward interpretation of ordinary multivariate linear regression difficult. As subset selection algorithms would prune the causal links between nodes of AOP, we did not use them here. Instead, we used ridge regularization[24]. The parameters were learned using glmnet package[25] and appropriate penalty λ^* was selected using leave-one-out-cross-validation (LOOCV). For this exercise, we used an activation cut-off $\Delta = log_{10} 2$. Combining data for all doses and donors, we numerically calculated **a)** $P(AO^{(e)} > \Delta | KE8^{(e-1)} > \Delta)$ for $e = 3, ., 6$ (4 probabilities in total), **b)** $P(AO^{(e)} > \Delta | KE6^{(e-2)} > \Delta)$, $P(AO^{(e)} > \Delta | KE5^{(e-2)} > \Delta)$, $P(AO^{(e)} > \Delta | KE4^{(e-2)} > \Delta)$, $P(AO^{(e)} > \Delta | KE7^{(e-2)} > \Delta)$ for $e = 4, 5, 6$ (3 probabilities in total for each KEs), and finally **c)** $P(AO^{(e)} > \Delta | KE2^{(e-3)} > \Delta)$, $P(AO^{(e)} > \Delta | KE3^{(e-3)} > \Delta)$ for $e = 5, 6$ (2 probabilities in total for each KEs) by resampling data from DBNs and using likelihood weighing. We also calculated probabilities of AO for specific dose, $P(AO^{(e,d)} > \Delta | KE^{(e-\tau,d)} > \Delta)$, where the interaction strengths $\boldsymbol{\beta}$ s were not only learned between successive exposure repetition slices but also for specific doses as shown below

$$x^{(e,v,d)} = \mathbb{X}^{(e-1,\Pi_v,d)} \boldsymbol{\beta}^{(e,v,d)} + \boldsymbol{\varepsilon}^{(e,v,d)} \sim \mathcal{N}(0, \sigma^2 \mathbb{I}). \tag{9}$$

Data driven pruning of AOP

As an additional exercise, we further implemented a data driven approach to restructure AOP, where we pruned AOP in Fig. 1 using a subset selection strategy: for nodes regulated by more than one predictor nodes, we implemented a LOOCV based lasso regularization that shrinks coefficients for uninformative/irrelevant predictors towards zero; for nodes that were regulated by one predictor, we only included it if the p value of the slope parameter was < 0.05. The customized BN was further scrutinized to only include those nodes for which at least one path existed to AO (for example, if the path from KE6 $\longrightarrow$ KE8 had non-zero coefficient but KE8 $\longrightarrow$ AO was zero, and since AO could not be reached from KE6 without visiting KE8, KE6 was pruned as well).

Results

Virtual data generation

The primary dataset was generated using random number generation function to suit the assumptions described earlier (see Methods section "Structure of hypothetical AOP"). Because there are only a few real-life datasets for qAOP modeling of chronic toxicity, we considered the primary dataset as the grand truth. The generated primary dataset is summarized in Supplementary material 2. For large number of replicates $r = 1, 2, ., ., ., \mathcal{R}$, the mean and the standard deviation of fold-change, $\sum_r f^{(n,e,d,v,r)}/\mathcal{R}$, $\sum_r \left(f^{(n,e,d,v,r)} - \overline{f}^{(n,e,d,v)} \right)^2/\mathcal{R}$, would approach the mean fold-change ($\overline{f}^{(n,e,d,v)}$) and standard deviation ($s^{(n,e,d,v)}$) in the primary dataset. However, for small number of replicates (typically ~ 3), and small sample size effect can be significant. One way to account for small sample size is by means of confidence intervals. Alternately, one can re-sample a small number of samples $R \ll \mathcal{R}$ as discussed in the Methods section, and then take the replicate average of fold-change $f^{(n,e,d,v,r)}$. In this study, we resampled the primary dataset using Bayesian updating approach as described in the methods section "Consideration of correlation co-efficient for resampling the primary dataset"-"Bayesian updating approach for resampling the primary dataset". As expected, the virtually generated dataset showed that the chronic phase biomarkers were not readily activated, but after three successive exposure repetitions, moderate to high expression of chronic phase responses was observed (Supplementary Fig. 1).

Dose–response and response-response relationship

Figure 3 shows the conditional R^2 values for dose–response and response-response modeling when LMM was used and standard training R^2 value for the polynomial regression model. As expected, the acute phase nodes generally showed strong dose–response for all exposure repetitions, whereas the chronic phase nodes started to show dose response only for $e \geq 3$. Response-response analysis was performed on the combined dataset for all exposure using a LMM and a polynomial regression model. We found that LMM outperformed polynomial regression model overall (Fig. 3), except for chronic-phase nodes, when they were comparable. Using analysis from LMM, we observed that all acute-phase nodes were highly correlated with each other. Similarly, all chronic-phase nodes were also correlated with each other; however, there was a low correlation between acute-phase nodes and chronic-phase nodes (Fig. 1C).

Probability analysis with Bayesian modeling approach

GBNs were fitted to the data for each exposure repetition separately. Figure 4 shows the R^2 values of those fits. Except for KE2 and KE7, R^2 values of other chronic KEs were consistently > 0.6 for exposure repetition $e = 3$ onwards. The fits for KE7 and KE2 did not do well for any exposure repetition because they were both regulated by upstream acute phase nodes KE3 and BM8 with whom they were very poorly correlated (Fig. 4). R^2 values for other acute phase KEs were also > 0.6 by-n-large.

GBNs were resampled and the probabilities of KEs at a given dose, $P(v > \Delta | d \in [d \pm \varepsilon])$, were estimated using logic sampling. Figure 5 shows these probabilities when the activation threshold was set to $\Delta = log_{10}2$. We see that $P(KE1 > \Delta | d \in [d \pm \varepsilon])$ increased monotonically with d and showed a strong dose dependence for all exposure repetitions. Similarly, $P(KE3 > \Delta | d \in [d \pm \varepsilon])$ also increased but the dose dependence was not as strong as KE1. As expected, probabilities $P(v > \Delta | d \in [d \pm \varepsilon])$ for the chronic phase nodes on the other hand were practically zero for $e = 1, 2$, demonstrating that chronic-phase nodes were activated only after a certain exposure repetition. From $e = 3$ onwards, $P(v > \Delta | d \in [d \pm \varepsilon])$ for all chronic phase nodes increased consistently with successive exposures ($P_{e=5}(v > \Delta | d \in [d \pm \varepsilon]) > P_{e=4}(v > \Delta | d \in [d \pm \varepsilon])$, so and so forth), but did not show a strong dose dependence. Additionally, these probabilities are consistently higher than 30% only from $e = 5$ onwards. The choice of $\Delta = log_{10}2$ was adhoc and applied uniformly across all KEs. We also varied Δ and estimated how probabilities $P(v > \Delta | d \in [d \pm \varepsilon])$ varied as a function of both the activation threshold and dose. The plots are shown in Supplementary Fig. 2. A more succinct way of visualizing these surfaces was by calculating the volume under the surface (VUS), $\int P(v > \Delta | d \in [d \pm \varepsilon])$. VUSs plot (Fig. 6) showed clearly that the VUSs for two acute phase key events, KE1 & KE3, were comparable across exposure repetitions, highlighting that their activations were exposure repetition independent, but VUSs for chronic phase KEs progressively grew with exposure repetitions post $e = 3$, showing once again that chronic phase responses required repeated exposures for their activation (Fig. 6).

While this work was done using virtual data, it showed the kind of information researchers can extract from this type of modeling once in vitro data become available. This work also demonstrates the qAOP formalism is highly flexible and can be used as a tool for prospective study.

DBN modeling for AO probability calculation

General trend pooled dose

Here we explored how an earlier activation of an upstream KE influenced AO subsequently. Formally, we estimated transition probability $P\left(AO^{(e)} > \Delta | KE^{(e-\tau)} > \Delta\right)$ by first fitting the DBN (Eq. (8), and Fig. 2) to the data, resampling from the fitted DBNs, and finally estimating the probability using likelihood weighing. For the first part of this work, we pooled all doses together. As shown in the Fig. 2, the arrows causally connected nodes between two successive exposure repetitions. It is worth repeating that unlike time-invariant time-series analysis, the strengths of interactions (the thickness of arrows) between two nodes were not constants, but they changed dynamically with e. For example, the strength of interaction $KE8^{(e-1)} \rightarrow AO^{(e)}$ was not the same for $e = 3$ and $e = 4$: interaction strength was weak for $e = 3$, but strong for $e = 4$ and onwards (Fig. 7). We observed that $P\left(AO^{(e)} > \Delta | KE8^{(e-\tau)} > \Delta\right)$ (denoted as $KE8^{(e-\tau)} \rightarrow AO^{(e)}$ in Fig. 7),
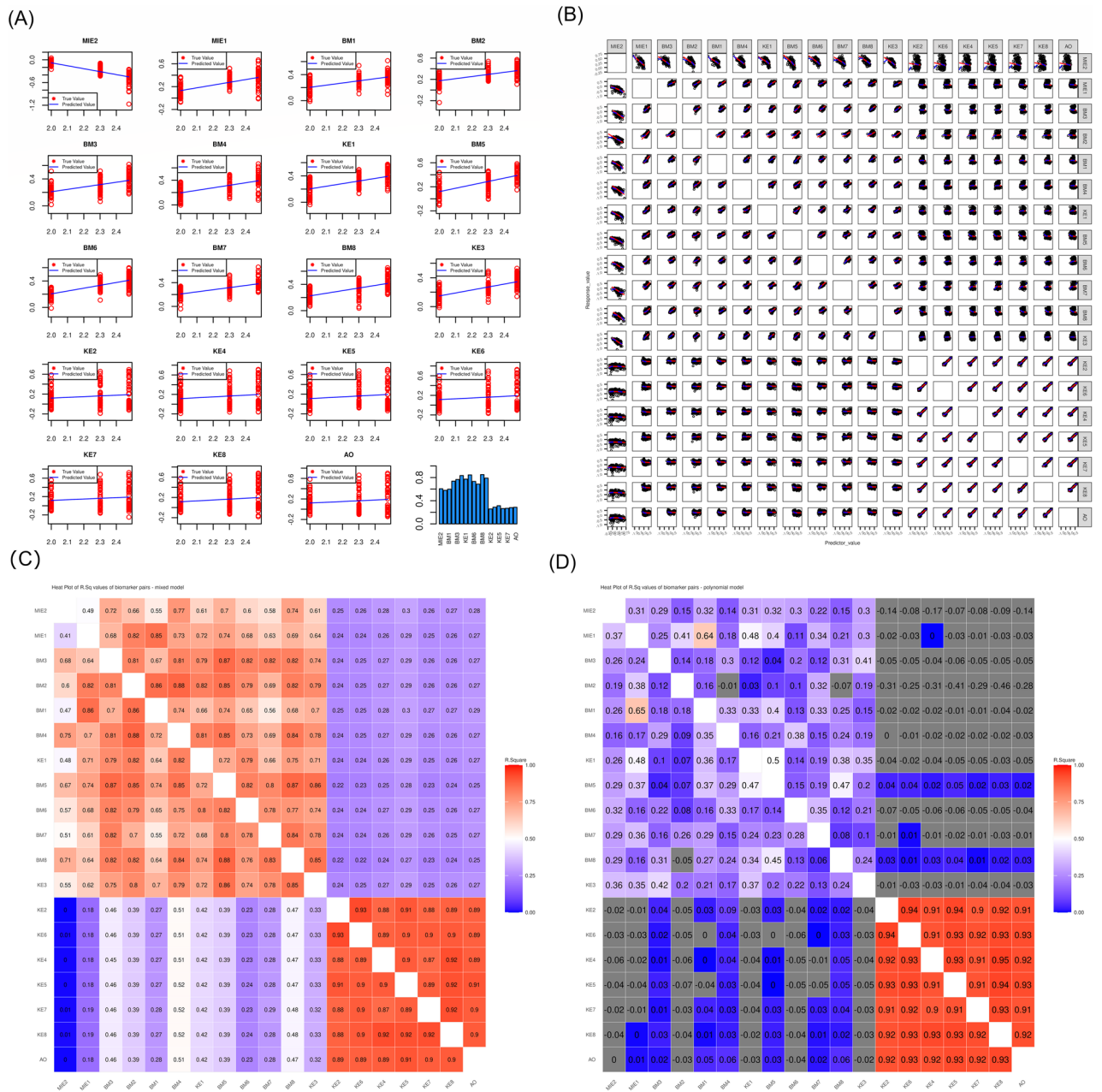
Figure 3. R² values for dose–response and response-response models: Linear mixed model (LMM) was used for both dose–response (Eq. 5) and response-response modeling (Eq. 6a). The conditional R² values[26] were reported for both dose–response (**A**) and response-response modeling (**B**). For dose–response, each exposure repetition was analyzed separately. For response-response modeling, all exposure repetitions were pooled together. (**C**) Heatmap of R² values of biomarker pairs. (**D**) Same as (**B**) and (**C**), except polynomial regression (Eq. 6b) was used and R² coefficient was calculated using standard formula. These analyses and visualizations were performed using R statistics software version 4.0.4.

increased monotonically as $e \to E = 6$ for all $\tau = 1, 2, \& 3$. In addition, activation of KE8 at or after $e = 4$ was highly predictive of adverse outcome (~ 0.96). For KEs that were separated by $\tau = 2$, namely KE4, KE5, KE6, & KE7, we found a similar increasing trend: the probability of AO reached values close to 0.8 for $e - \tau = 4$. For $e - \tau = 3$, the probability of AO was roughly 0.6. KE3 is an acute-phase response in our model and though it regulated KE7, they were poorly correlated (Fig. 4). Consequently, we did not expect the probabilities $P\left(AO^{(e)} > \Delta | KE3^{(e-3)} > \Delta\right), e = 5, 6$ to depend on the specific exposure repetition e. Our analysis confirmed that indeed $P\left(AO^{(e=5)} > \Delta | KE3^{(e-3=2)} > \Delta\right) \cong \left(AO^{(e=6)} > \Delta | KE3^{(e-3=3)} > \Delta\right)$ (Fig. 7).

$P\left(AO^{(E)} > \Delta | KE2^{(e=3)} > \Delta\right)$, oddly, was very close to 1. The reason remains unclear, but we think that it can be because the outgoing degree (number of outgoing edges) for KE2 is highest among all the chronic KEs.
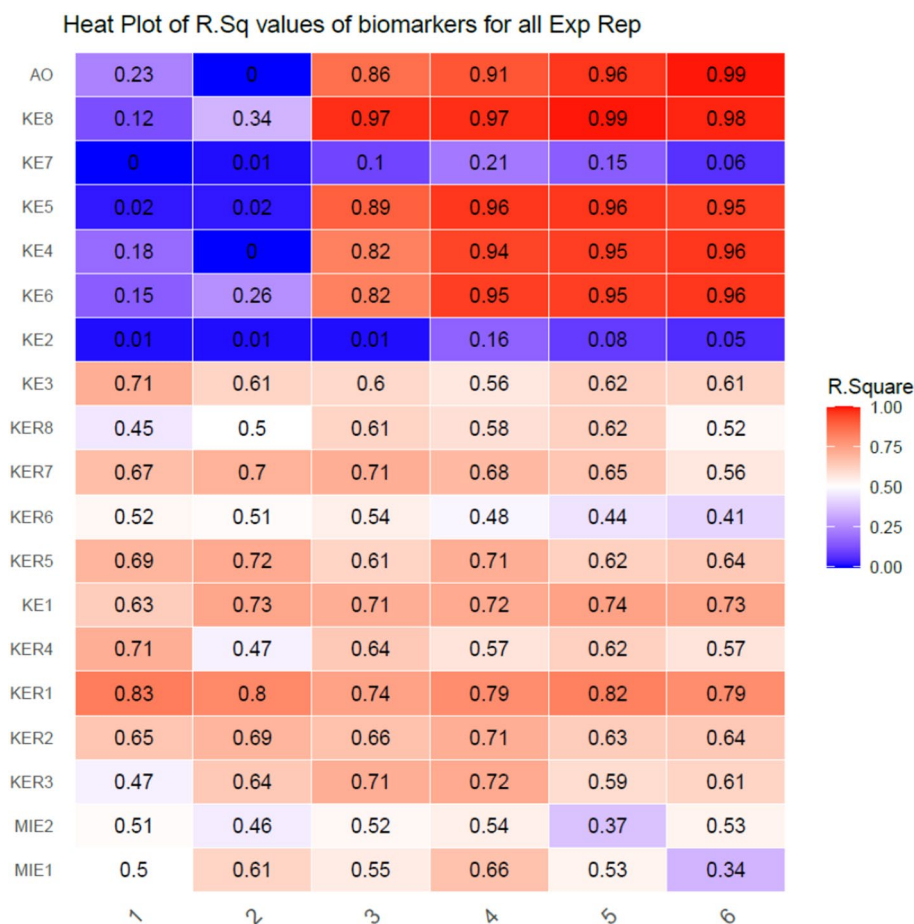
Figure 4. R^2 of biomarkers for all exposure repetitions. R^2 for each node was calculated using bnlearn that implements standard linear regression model. The predictors for each node were their parent set Π_v. The analysis and visualization were performed using R statistics software version 4.0.4.

Dose specific trends

The general trends for dose specific transition probabilities, $P\big(AO^{(e,d)} > \Delta | KE^{(e-\tau,d)} > \Delta\big)$, by-n-large met our expectations: probabilities increased as $e \to E = 6$ for all $\tau = 1, 2, 3$, and as the dose $d \uparrow$. Notable exceptions were $P\big(AO^{(e=4,d=200)} > \Delta | KE6^{(e-2=2,d=200)} > \Delta\big)$, & $P\big(AO^{(e=5,d=200)} > \Delta | KE2^{(e-3=2,d=200)} > \Delta\big)$ (Fig. 8A). One plausible explanation is that the DBNs were fitted with only eight donor level data points for each dose, rendering it error prone. We took a deep dive and visualized log fold changes of $AO^{(e=3,d)}$ vs $KE8^{(e-1=2,d)}$ for all three doses (Fig. 8B). For such an early exposure ($e = 3, e - 1 = 2$), chronic responses were not elicited and no correlation between two chronic events was expected. Yet, we found positive correlation for low dose but negative at mid and high dose between AO and KE8. We considered these inconsistent correlations as purely coincidental and artifacts of small sample size.

Data-driven AOP restructuring

We also developed a data-driven AOP restructuring method based on lasso regression: an interaction was considered important provided pruning that interaction resulted in significant decline in the quality of the fit. In order to accomplish the task of pruning, the GBNs were parameterized using Lasso regression (methods section "Dose–response and response-response analysis"). For nodes with only one parent regulator, we preserved the interaction provided the p-value of the slope parameter was < 0.05.

In the repeated dosing study, it could be expected that the strength of causal links varied based on exposure repetition. In addition, changes in the causal relationship could result in some of the predictors to be uninformative. The data-driven pruning of DBNs are shown in Fig. 9A. The uninformative predictors and their links are shown in grey. We found that KE8 at $e = 2$ did not exert any influence on AO at $e = 3$: the probability of AO at $e = 3$ was conditionally independent of KE8, i.e. $P\big(AO^{(e=3)} > \Delta | KE8^{(e=2)} > \Delta\big) = P\big(AO^{(e=3)} > \Delta\big)$. For all subsequent exposure slices, KE8 exerted influence on AO outcomes. Similarly, we found KE4, KE5, KE6, and KE7 at $e = 2$ had no bearing on AO at $e = 4$; KE4 and KE5 at $e = 3$ conditionally influenced AO at $e = 5$. All these KEs were conditional determinants of AO only for $e = 6$. KE3 was never a conditional determinant, and KE2 at $e = 3$ influenced AO at $e = 6$ but not for any earlier time slices. This data-driven remodeling
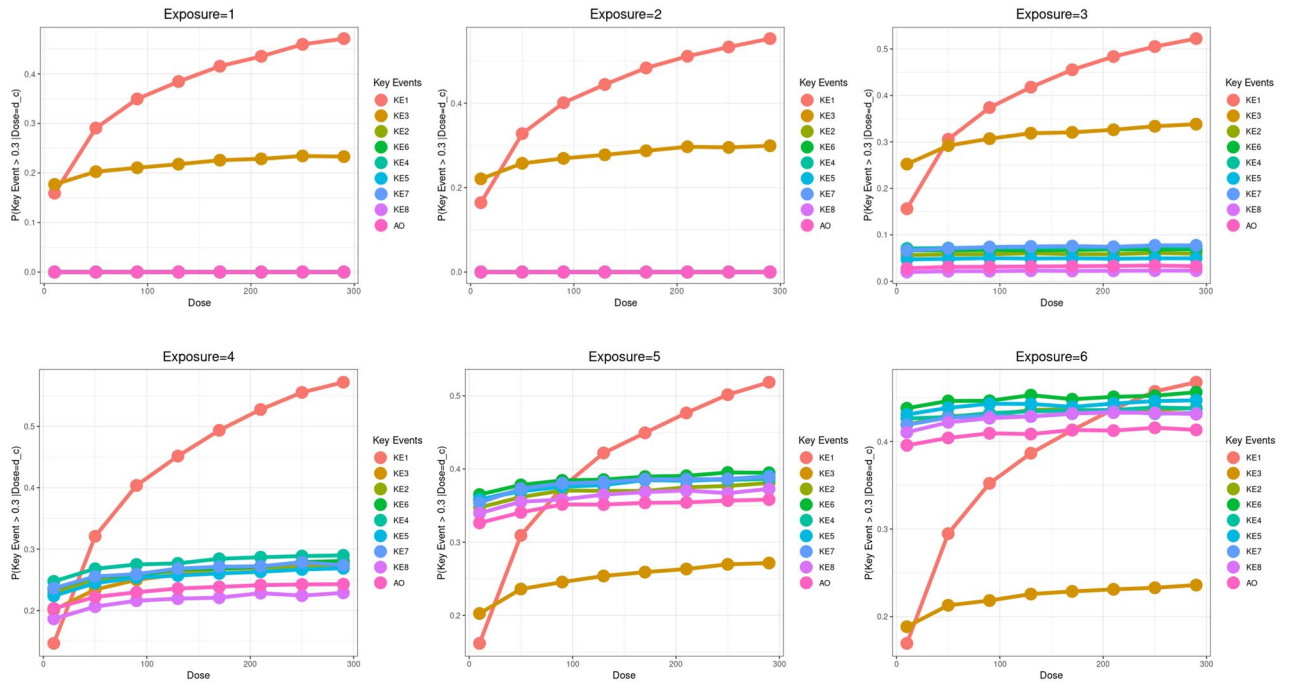
Figure 5. Dose-probability curves of Bayesian Network analysis: The probabilities were calculated by fitting the GBN to the virtual data, followed by logic sampling. The threshold of activation Δ was set to $log_{10}2$. The analysis and visualization were performed using R statistics software version 4.0.4. The lines corresponding to KE2, 4, 5, 6, 7, and 8 at the exposure 1 and 2 are hidden by the line of AO.
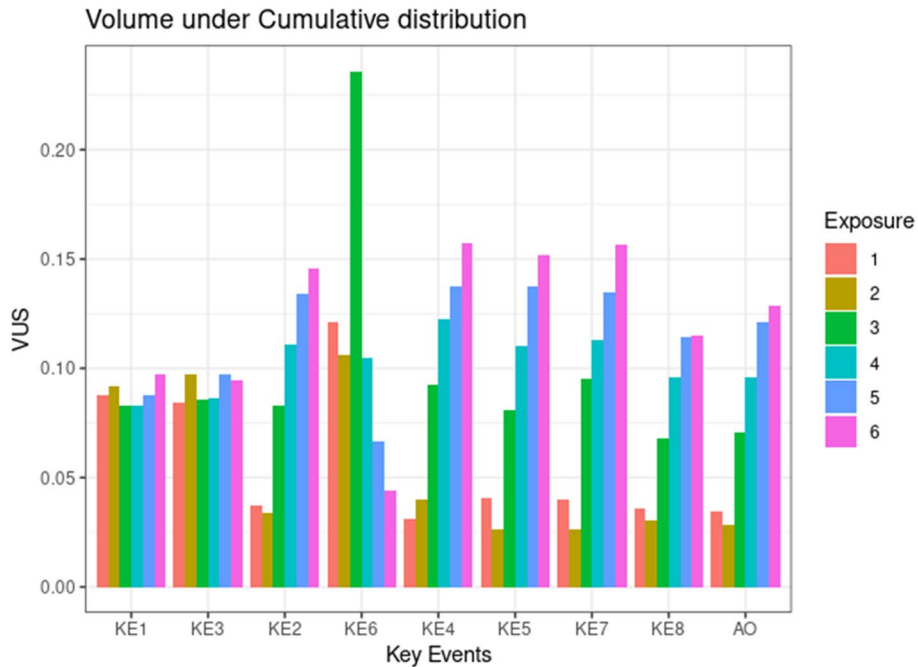

Figure 6. Volume under surfaces of KEs at each exposure repetitions. Probabilities surfaces were calculated by varying the activation threshold Δ and the dose d. The volume under the surface (VUS) of those variables was calculated using trapezoid rule. The analysis and visualization were performed using R statistics software version 4.0.4
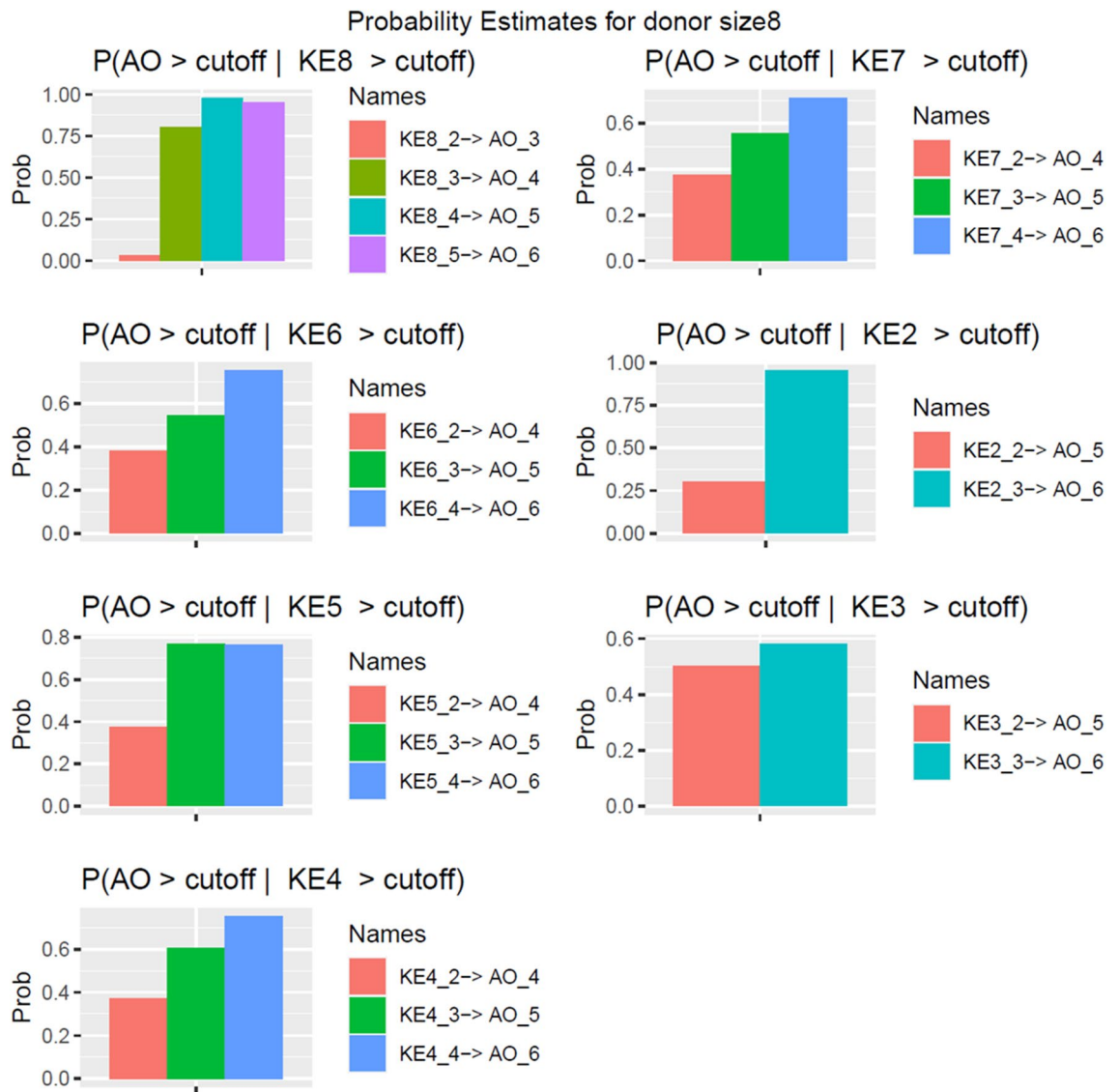
Figure 7. Transition probabilities with all dose groups combined. The transition probabilities were calculated using likelihood weighing after the dynamic Bayesian Network (DBN) was trained with data from successive exposures. The notation $KE\nu\_i \rightarrow AO\_j$ was used to denoted $P(AO^{e=j} > \Delta | KE\nu^{e-\tau=i} > \Delta)$. The probabilities were calculated when all dose groups were combined. The analysis and visualization were performed using R statistics software version 4.0.4.

enabled us to reevaluate the transition probabilities and extract only the relevant conditional probabilities $(P(AO^{(e)} > \Delta | KE^{(e-\tau)} > \Delta) \neq P(AO^{(e=3)} > \Delta))$, as shown in Fig. 9B. We found that while the numerical values of transition probabilities were a little different from the DBN with whole AOP structure, their qualitative characteristics were still the same: $P(AO^{(e)} > \Delta | KE8^{(e-1)} > \Delta)$ was very high for $e \geq 4$; $P(AO^{(e)} > \Delta | KE5^{(e-2)} > \Delta)$ was $\gtrsim 0.8$ for $e = 5$ but went down to ~ 0.7 for $e = 6$; $P(AO^{(e)} > \Delta | KE4^{(e-2)} > \Delta)$ increased from 0.5 for $e = 5$ to ~ 0.8 for $e = 6$; $P(AO^{(e)} > \Delta | KE6^{(e-2)} > \Delta)$ was ~ 0.8 for $e = 6$ and it was the only instance when AO is conditionally dependent of KE6; $P(AO^{(e)} > \Delta | KE7^{(e-2)} > \Delta)$ was ~ 0.6 for $e = 6$; $P(AO^{(e)} > \Delta | KE2^{(e-3)} > \Delta) \sim 1$ for $e = 6$. A plausible reason for this high value was discussed in the previous section.

Discussion

The goal of the AOP framework is to provide a principled approach to better understand the pathway to adverse outcome by synergistically integrating information from different layers of biological organization and advance a more informed testing strategy[3]. Owing to its flexibility and its ability to ingest information, not just from in vivo animal models but also from in vitro new approach methodologies, the AOP framework has emerged as a powerful tool for chemical testing and risk assessment. Cost effective in vitro methods enable toxicologists to collect information from different layers of biological organizations that can then be integrated into a cohesive AOP framework and used as a roadmap for an adverse outcome. Over the years, the number of expert derived
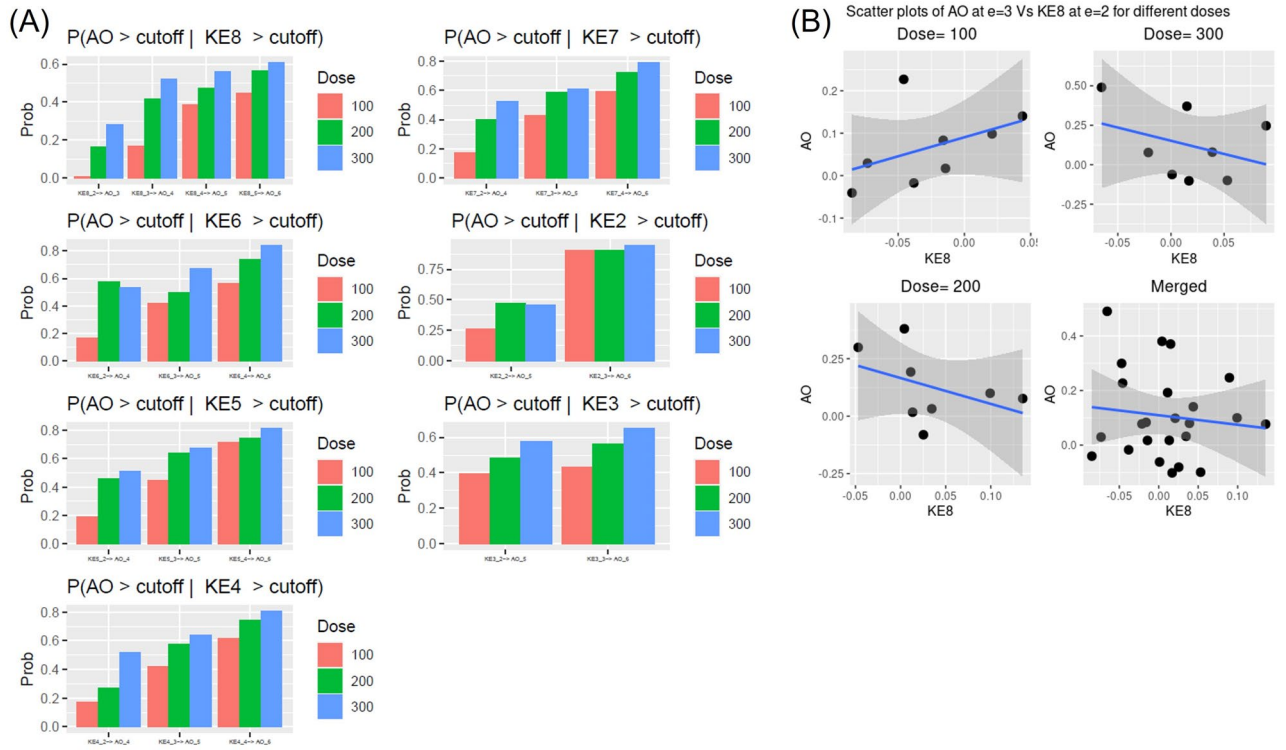
Figure 8. Transition probabilities from each dose separately. (**A**) The transition probabilities were calculated for each dose separately. The dose groups are color coded and like before, $KE\nu\_i \rightarrow AO\_j$ stood for $P(AO^{e=j,d} > \Delta | KE\nu^{e-\tau=i,d} > \Delta)$ (**B**) Scatter plots of AO at e = 3 and KE8 at e = 2 at each separate dose and pooled dose. These analyses and visualizations were performed using R statistics software version 4.0.4.
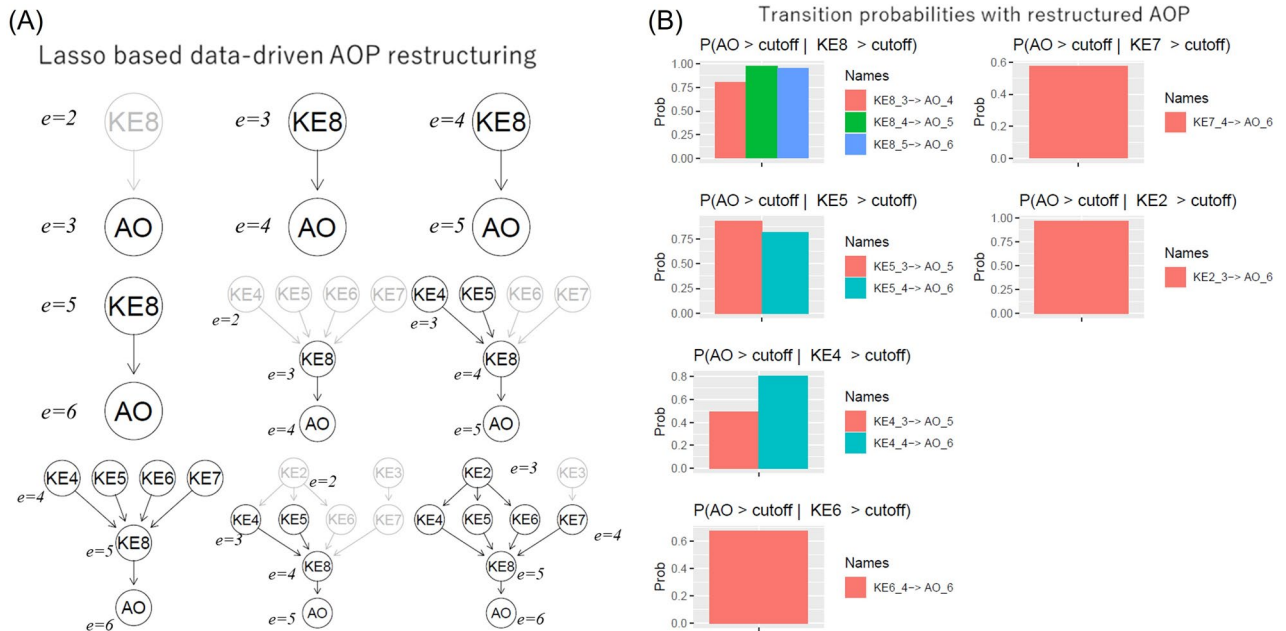

Figure 9. Data-driven AOP restructuring and recalculation of transition probabilities. (**A**) The DBN was trained using lasso regression. The causal links deemed insignificant were shown in grey. (**B**) Probabilities were calculated using likelihood weighing, and only those transition probabilities where $P\left(AO^{(e)} > \Delta | KE^{(e-\tau)} > \Delta\right) \neq P\left(AO^{(e=3)} > \Delta\right)$, were shown. Like before, $KE\nu\_i \rightarrow AO\_j$ was used to denoted $P(AO^{e=j} > \Delta | KE\nu^{e-\tau=i} > \Delta)$. These analyses and visualizations were performed using R statistics software version 4.0.4.

AOP submissions have grown (https://aopwiki.org/). There are several AOPs that deal in chronic toxicity as the AO, however, AOPs that deal with repeated exposure and donor dependent onset of chronic events are still sparse. One reason for this humble showing can be difficulty in AOP to represent dynamism of biological pathway perturbation during repeated exposure to stimuli. Regardless, dearth of repeated exposure related AOPs that deal with acute phase and chronic phase KEs differently, account for donor dependent onset of chronic events, needs to be addressed.

Also, generation of the data alone is not sufficient. Data arising from molecular to phenotypic level need to be integrated and interpreted unambiguously to really render an AOP useful. Moreover, the associations between different layers of biological organization need to be quantified and made manifest. How these associations evolve over different exposure repetitions need to be quantified. Only a principled quantitative approach can achieve that.

To demonstrate the utility of a chronic qAOP framework by taking repeated exposure into consideration, we used a hypothetical AOP with acute phase responses (KEs and BMs), chronic phase KEs, and AO. As we mentioned, BMs are involved in this hypothetical AOP, because some secondary measures outside AOP could be also obtained in the real experimental situations, and they could be informative of AOP perturbation. Separation of acute phase KEs and chronic phase KEs/AO enabled us to consider the dynamism of biological perturbation. We generated a virtual dataset to quantify this AOP. In generating this dataset, we assumed that the chronic KEs were initiated only after repeated insults whereas the acute KEs were readily activated. In this study, we assumed a six-time repeated exposure study as a proof of concept, although user-defined study design is acceptable for our model. The acute KEs showed strong dose–response, the chronic KEs showed dose–response and exposure repetition-response only after they were activated. We further assumed that the onset of chronic KEs varied from donor to donor, as in vitro models with which the researcher can perform repeated exposure study is often composed of primary cells. Primary cells generally possess their original characteristics, could thus show donor-dependent variation in response against the same stimuli[27–29].

Using a GBN, we showed that several pertinent questions regarding probability of an adverse event can be answered quantitatively and qAOP model can be used prospectively in risk assessment. Specifically, using GBN model we calculated probabilities of KEs and AO to be activated above a certain threshold Δ given a dose d and an exposure repetition e. We varied Δ and d and investigated how the probability of KEs change. While what constitutes a threshold is debatable and subjective, on quick way to estimate Δ can be from the dose–response curve. The point here is, even when Δ is unknown, the BN framework gives us the flexibility to prospectively study the effect of varying Δ on probabilities of KEs and AOs. Therefore, we implemented the function to depict the changes in the probability as three-dimensional plot. The volume under the cumulative distribution (i.e., volume under the probability surface, VUS) can be then calculated as an index of the AOP perturbation with varied Δ(Fig. 6). The result of VUS calculation demonstrates that the acute KEs and chronic KEs behave very differently: VUS of acute KEs do not change significantly with exposure repetitions, whereas VUS for chronic KEs show strong exposure repetition dependence. The point to note is that this proof-of-concept study utilizes virtually generated data; real biological samples may not be as conclusive. Most important point here is that our approach enables to formulate the different types of analysis, as the model demonstrates what was intuitively clear: probability of chronic phase outcomes increase with repeated exposure.

We also used the DBN and the repeated exposure data to estimate probability of AO to be activated when a certain KE was activation earlier (transition probability). We showed that activation of chronic KEs earlier can be used as a reliable proxy for AO activation later. Not surprisingly, the closer the proximity of the KE to the AO node, the better predictor it was of AO activation. We also showcased what can happen when acute phase events are poorly correlated with chronic events. Such information bottleneck, when present, can seriously limit the ability to infer likelihood of AO based on early evidence of acute phase response. As clear from Figs. 7 and 8A (P(AO > cut off | KE2 > cut off), inference of AO can be compromised when based solely on upstream events that are information bottlenecks. We further demonstrated that these transition probabilities also showed a clear dose dependence. Furthermore, we proposed data-driven restructure of AOP where uninformative KEs could be eliminated from the probability calculations. This approach would help in reducing calculation cost and complexity of AOP on repeated-exposure-oriented chronic toxicity. Using lasso-based pruning, we demonstrate that the wiring in an AOP is itself dynamic, changing over time (Fig. 9). Such flexibility is well known in biology and must be taken into account when drawing inference on AO. Here we showed clearly that the DBN method can be used in detection of early events that are strong proxies for AO. DBN formalism is extremely flexible and very pertinent for biology in general because it can model feedback loops that are ubiquitous in any biological processes. While we fully acknowledge that we did not incorporate all the potential confounders of a real dataset, we believe that our in silico approach can serve as a proof-of-concept and bring home the fact that several specific questions regarding an adverse event, including dose and early response dependent probability of adverse outcome can be answered quantitatively.

Conclusion

Although several reports attempted to model chronic toxicity[13,30], they mostly stand upon a simple sequence from MIE through AO, in which cumulative effects are not modeled. We here demonstrated BN and DBN formalisms could enable us to incorporate cumulative effect as well as repeated perturbation of biological pathway. This means that the AOP structure itself could become dynamic as KEs can be annotated as acute or chronic phase. Although it should be acknowledged that our model was developed with a virtual setup (i.e., a hypothetical AOP with a virtual dataset), we believe our proof-of-concept study was able to explore the possibility that AOP framework could be expanded for more flexible toxicity consideration. In addition, considering a part of

overlap between chronic toxicity and diseases, our concept for AOP framework could be the future application for the risk assessment of chronic diseases if relevant in vitro NAMs and AOPs are available.

The current study utilizes a hypothetical AOP with two MIEs. Usually, AOPs do not have multiple MIEs. The reason for us to have two MIEs are really two folds. First, multiple assays can be used to measure an MIE. While it is entirely feasible to subsume all these endpoints, nothing in principle prevents them from being treated separately for modeling. Secondly, having multiple MIEs in our framework enables this method to be seamlessly translated to an AOP-network scenario, where multiple MIEs can influence a downstream KE[31].

In addition, using high throughput screening assays and high content imaging, molecular and phenotypic endpoints can be captured more intensively than ever. AI and machine learning (ML) approaches with these data stream could be highly compatible with AOP framework. So far, inference of outcomes by stimulation is largely dependent on molecular level information thus limited to the early and apical outcomes. However, AOPs provide a natural framework for integration of molecular, phenotypic and organ level information. qAOPs—quantified using HTS, HC imaging, and other NAM data—can be used for estimation of AO conditioned on early evidence. Here we showed two such scenarios in Fig. 5 and Sup Fig. S2. In principle, evidence of arbitrary complexity can be modeled. For prospective study, qAOP can be used to simulate perturbation, and test alternate hypothesis. qAOP wirings can be altered in silico, and various scenarios that can potentially decrease likelihood of AO can be explored. Furthermore, and most importantly, qAOPs can be used for data poor chemical as a read-across by considering qAOPs of structural analogs.

For the practical implementation of NAMs, including qAOP, within the regulatory framework, several key points need to be addressed. It is crucial to promote the evolution of science through knowledge sharing by reusing both input and output data across different studies. Therefore, the use of a common reporting template is essential[32]. Additionally, the modeling approach should always be evaluated in terms of its domains of applicability, potential bias, and level of certainty or uncertainty[33].

When focusing on qAOP modeling, transparency in the modeling approach, dataset, and their quality and reliability are of utmost importance for decision-making[34]. To this end, guidelines and workflows have already been proposed, and researchers are encouraged to follow them when using the qAOP models including the models developed in this study. By adhering to these practices, we can ensure greater consistency and reliability in the application of qAOP models in regulatory decision-making processes.

Data availability
The data that support the findings of this study are available from the corresponding author, upon reasonable request.

References
1. van der Zalm, A. J. *et al.* A framework for establishing scientific confidence in new approach methodologies. *Arch. Toxicol.* **96**, 2865–2879 (2022).
2. Vinken, M. The adverse outcome pathway concept: A pragmatic tool in toxicology. *Toxicology* **312**, 158–165 (2013).
3. Ankley, G. T. *et al.* Adverse outcome pathways: A conceptual framework to support ecotoxicology research and risk assessment. *Environ. Toxicol. Chem.* **29**, 730–741 (2010).
4. Perkins, E. J. *et al.* Building and applying quantitative adverse outcome pathway models for chemical hazard and risk assessment. *Environ. Toxicol. Chem.* **38**, 1850–1865 (2019).
5. Spinu, N., Cronin, M. T. D., Enoch, S. J., Madden, J. C. & Worth, A. P. Quantitative adverse outcome pathway (qAOP) models for toxicity prediction. *Arch. Toxicol.* **94**, 1497–1510 (2020).
6. Blanchette, A. D., Burnett, S. D., Grimm, F. A., Rusyn, I. & Chiu, W. A. A Bayesian method for population-wide cardiotoxicity hazard and risk characterization using an in vitro human model. *Toxicol. Sci.* **178**, 391–403 (2020).
7. Alepee, N. *et al.* State-of-the-art of 3D cultures (organs-on-a-chip) in safety testing and pathophysiology. *ALTEX* **31**, 441–477 (2014).
8. De Blasi, R. A., Campagna, G. & Finazzi, S. A dynamic Bayesian network model for predicting organ failure associations without predefining outcomes. *PLoS One* **16**, e0250787 (2021).
9. Jeong, J. *et al.* Developing adverse outcome pathways on silver nanoparticle-induced reproductive toxicity via oxidative stress in the nematode *Caenorhabditis elegans* using a Bayesian network model. *Nanotoxicology* **12**, 1182–1197 (2018).
10. Moe, S. J. *et al.* Quantification of an adverse outcome pathway network by Bayesian regression and Bayesian network modeling. *Integr. Environ. Assess. Manag.* **17**, 147–164 (2021).
11. Spinu, N. *et al.* Probabilistic modelling of developmental neurotoxicity based on a simplified adverse outcome pathway network. *Comput. Toxicol.* **21**, 100206. https://doi.org/10.1016/j.comtox.2021.100206 (2022).
12. Kirsten, G. *et al.* The adverse outcome pathway approach in nanotoxicology. *Comput. Toxicol.* **1**, 3–11 (2017).
13. Elias, Z. *et al.* Application of three approaches for quantitative AOP development to renal toxicity. *Comput. Toxicol.* **11**, 1–13 (2019).
14. Qu, W., Tokar, E. J., Kim, A. J., Bell, M. W. & Waalkes, M. P. Chronic cadmium exposure in vitro causes acquisition of multiple tumor cell characteristics in human pancreatic epithelial cells. *Environ. Health Perspect.* **120**, 1265–1271 (2012).
15. Jennings, P. *et al.* Assessment of a new cell culture perfusion apparatus for in vitro chronic toxicity testing. Part 2: Toxicological evaluation. *ALTEX* **21**, 61–66 (2004).
16. Koppelstaetter, C. *et al.* Assessment of a new cell culture perfusion apparatus for in vitro chronic toxicity testing. Part 1: Technical description. *ALTEX* **21**, 51–60 (2004).
17. Samardzija Nenadov, D. *et al.* Long-term in vitro exposure of human granulosa cells to the mixture of endocrine disrupting chemicals found in human follicular fluid disrupts steroidogenesis. *Toxicol. In Vitro* **79**, 105302. https://doi.org/10.1016/j.tiv.2021.105302 (2022).
18. Tratnjek, L., Sibinovska, N., Kristan, K. & Kreft, M. E. In vitro ciliotoxicity and cytotoxicity testing of repeated chronic exposure to topical nasal formulations for safety studies. *Pharmaceutics* **13**, 1750. https://doi.org/10.3390/pharmaceutics13111750 (2021).
19. Haswell, L. E. *et al.* The development of an in vitro 3D model of goblet cell hyperplasia using MUC5AC expression and repeated whole aerosol exposures. *Toxicol. Lett.* **347**, 45–57 (2021).
20. Scutari, M. & Denis, J.-B. *Bayesian Networks: With Examples in R* (CRC Press, 2014).

21. Koller, D. & Friedman, N. *Probabilistic Graphical Models: Principles and Techniques* (MIT Press, 2009).
22. Scutari, M. Learning Bayesian networks with the bnlearn R package. *J. Stat. Softw.* **35**, 3. https://doi.org/10.18637/jss.v035.i03 (2010)
23. Nagarajan, R., Scutari, M. & Lèbre, S. *Bayesian Networks in R With Applications in Systems Biology* (NYSpringer-Verlag, 2013).
24. Hastie, T., Tibshirani, R. & Friedman, J. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction* (Springer Science & Business Media, 2009).
25. Zou, H. & Hastie, T. Regularization and variable selection via the elastic net. *J. R. Stat. Soc. Ser. B (Stat. Methodol.)* **67**, 301–320 (2005).
26. Shinichi, N. & Holger, S. A general and simple method for obtaining R^2 from generalized linear mixed-effects models. *Methods Ecol. Evolut.* **4**, 133–142 (2013).
27. Jetten, M. J. *et al.* Interindividual variation in gene expression responses and metabolite formation in acetaminophen-exposed primary human hepatocytes. *Arch. Toxicol.* **90**, 1103–1115 (2016).
28. Mori, S., Ishimori, K., Matsumura, K., Ishikawa, S. & Ito, S. Donor-to-donor variability of a human three-dimensional bronchial epithelial model: A case study of cigarette smoke exposure. *Toxicol. In Vitro* **82**, 105391. https://doi.org/10.1016/j.tiv.2022.105391 (2022).
29. Bovard, D. *et al.* Comparison of the basic morphology and function of 3D lung epithelial cultures derived from several donors. *Curr. Res. Toxicol.* **1**, 56–69 (2020).
30. Conolly, R. B. *et al.* Quantitative adverse outcome pathways and their application to predictive toxicology. *Environ. Sci. Technol.* **51**(8), 4661–4672 (2017).
31. Burgoon, L. D. *et al.* Predicting the probability that a chemical causes steatosis using adverse outcome pathway Bayesian networks (AOPBNs). *Risk Anal.* **40**, 512–523 (2020).
32. Edoardo, C. *et al.* OECD harmonised template 201: Structuring and reporting mechanistic information to foster the integration of new approach methodologies for hazard and risk assessment of chemicals. *Regul. Toxicol. Pharm.* **142**, 105236. https://doi.org/10.1016/j.yrtph.2023.105426 (2023).
33. Brozek, J. L. *et al.* GRADE Guidelines 30: The GRADE approach to assessing the certainty of modeled evidence-an overview in the context of health decision-making. *J. Clin. Epidemiol.* **129**, 138–150. https://doi.org/10.1016/j.jclinepi.2020.09.018 (2020).
34. Paini, A. *et al.* Towards a qAOP framework for predictive toxicology-linking data to decisions. *Comput. Toxicol.* **21**, 100195. https://doi.org/10.1016/j.comtox.2021.100195 (2022).

Author contributions

S.I., K.Y., and D.F. conceived the idea for the study. S.I. and S.M. wrote the original manuscript. S.M. and W.W. carried out R programming of qAOP models. S.I., K.E., S.M., A.M., and S.I. performed the ideation of hypothesized AOP structure and virtual data generation with the assumptions described in the manuscript. S.I. and S.M. carried out data analysis with the qAOP models in R software. K.Y. and D.F. supervised the point of an argument of this manuscript. All authors critically reviewed, revised the original manuscript, and approved the final version to be published.

Funding

Competing interests

Shigeaki Ito, Kazuo Erami, Shugo Muratani, Akina Mori, Sakuya Ichikawa, and Kei Yoshino are employees of Japan Tobacco Inc., and Sayak Mukherjee, William White, and Dawn Fallacara are employees of Battelle. The authors declare that they have no competing interests with respect to the research, authorship, and/or publication of this article.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-024-55220-4.

Correspondence and requests for materials should be addressed to S.I.

Reprints and permissions information is available at www.nature.com/reprints.

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