Application of three approaches for quantitative AOP development to renal toxicity

Elias Zgheib, Wang Gao, Alice Limonciel, Hristo Aladjov, Huan Yang, Cleo Tebby, Ghislaine Gayraud, Paul Jennings, Magdalini Sachana, Joost Beltman, et al.

Application of three approaches for quantitative AOP development to renal toxicity

Elias Zgheib (1), Wang Gao (2,3), Alice Limonciel (4), Hristo Aladjov (5), Huan Yang (6), Cleo Tebby (2), Ghislaine Gayraud (3), Paul Jennings (4), Magdalini Sachana (5), Joost B. Beltman (6), Frederic Y. Bois (2*)

(1) Laboratoire de Biomécanique et Bio-ingénierie, CNRS UMR 7338, Sorbonne Universités – Université de Technologie de Compiègne (UTC), Compiègne, France.

(2) Models for Ecotoxicology and Toxicology Unit (DRC/VIVA/METO), Institut National de l'Environnement Industriel et des Risques (INERIS), Verneuil-en-Halatte, France.

(3) Laboratoire de Mathématiques Appliquées de Compiègne, EA 2222, Sorbonne Universités – Université de Technologie de Compiègne (UTC), Compiègne, France.

(4) Division of Molecular and Computational Toxicology, Amsterdam Institute for Molecules, Medicines and Systems, Vrije Universiteit Amsterdam, Amsterdam, The Netherlands.

(5) Environment Health and Safety Division, Organisation for Economic Co-operation and Development (OECD), Paris, France.

(6) Division of Drug Discovery and Safety, Leiden Academic Center for Drug Research (LACDR) / Leiden University, Leiden, the Netherlands.

(*) to whom correspondence should be addressed.

Corresponding author: Frederic Y. Bois, INERIS, METO unit, Parc ALATA BP2, 60550 Verneuil en Halatte, France; Frederic.Bois@ineris.fr.

Keywords: Bayesian networks, Chronic kidney disease, Potassium bromate, Predictive toxicology, Quantitative AOP, Systems biology model.

Number of words: 10,280

Number of figures: 8.

List of abbreviations:

AIC: Akaike information criterion

AO: adverse outcome

AOP: Adverse outcome pathway

BIC: Bayesian information criterion

BN: Bayesian network

cDCF: 6-Carboxy-2',7'-dichlorofluorescein

cH2DCF: 6-Carboxy-2',7'-dichlorodihydrofluorescein

cH2DCFDA: 6-Carboxy-2',7'-dichlorodihydrofluorescein diacetate

CKD: Chronic kidney disease

cytGSH: Cytosolic glutathione

cytGSSG: Cytosolic glutathione disulfide

DBN: Dynamic Bayesian network

DIC: Deviance information criterion

extGSH: Extracellular glutathione

extGSSG: Extracellular glutathione disulfide

GSD: Geometric standard deviation

GSH: Glutathione

ITS: Integrated testing strategy

$KBrO_3$: potassium bromate

KE: Key event

Keap1: Kelch-like-ECH-Associated Protein 1

KER: Key events relationship

MCMC: Markov chain Monte-Carlo

MIE: Molecular initiating event

MRP: Multidrug resistance-associated protein

NFE2L2, Nrf2: Nuclear Factor (Erythroid-derived 2)-Like 2 p

Nrf2: nuclear factor (erythroid-derived 2)-like 2, NFE2L2

OECD: Organisation for Economic Co-operation and Development

qAOP: Quantitative adverse outcome pathway

RFU: Relative fluorescence units

ROS: Reactive oxygen species

SB: Systems biology

Abstract

While hazard assessment of chemicals can make direct use of descriptive adverse outcome pathways (AOPs), risk assessment requires quantitative relationships from exposure to effect timing and magnitude. To seamlessly integrate the data generated by alternative methods or *in vivo* testing, quantitative AOPs (qAOPs) providing dose-time-response predictions are more valuable than qualitative AOPs. Here, we compare three approaches to qAOP building: empirical dose-response modeling, Bayesian network (BN) calibration, and systems biology (SB) modeling. These methods were applied to the quantification of a simplified oxidative stress induced chronic kidney disease AOP, on the basis of *in vitro* data obtained on RPTEC/TERT1 cells exposed to potassium bromate. Effectopedia was used to store the experimental data and the developed models in a unified representation so they can be compared and further analyzed. We argue that despite the fact that dose-response models give adequate fits to the data they should be accompanied by mechanistic SB modeling to gain a proper perspective on the quantification. BNs can be both more precise than dose-response models and simpler than SB models, but more experience with their usage is needed.

1 Introduction

Adverse outcome pathways (AOPs) have become an organizing framework to facilitate the development and integration of alternative test methods for assessing hazard of chemicals to human health and the environment. A dedicated program is currently running under the auspices of the Organisation for Economic Co-operation and Development (OECD). AOPs are intended "to outline and capture existing knowledge concerning the biologically plausible and empirically supported foundations for predicting apical toxicity from mechanistic data" [1]. Practically, an AOP is a chemical-independent description of a linear path from a molecular initiating event (MIE) to an eventual adverse outcome (AO) at the organism or population level [2] (see Figure 1). In between, there can be any number of intermediate critical and measurable key events (KEs) connected by key events relationships (KERs) [3,4]. In typical AOP diagrams, KEs are represented by boxes and KERs by single one-directional arrows connecting them (Figure 1). The path linking the various KEs should not form loops (feedback or feed-forward loops between two consecutive KEs can simply be indicated by a symbol and included in the KER). Thus, according to graph theory, in the absence of loops within KERs, AOPs are acyclic directed graphs [5].

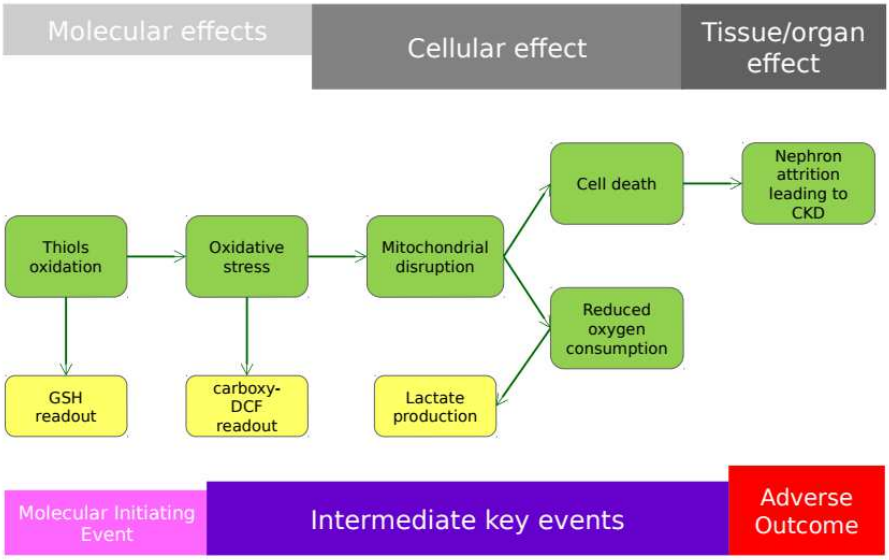

Figure 1: Chronic kidney disease AOP diagram. KEs are in green boxes, readout are in yellow boxes, KERs are represented by arrows.

AOPs can support the development of integrated testing strategies (ITS) and their application in risk assessment [7,8]. In case of ITS building, the data generated by alternative methods (*i.e.*, *in silico, in chemico, in vitro*), when combined with existing animal data, are used and assessed by means of a fixed data interpretation procedure [1,9]. For this purpose, quantitative AOPs (qAOPs) that provide dose-response and time-course predictions [10] are likely to be more valuable for ITS construction than qualitative AOPs. Parameter values for a qAOP can be either obtained from legacy data or new targeted experimental work, or by optimizing the fit of model predictions to data [2]. So far, the few published qAOPs use either empirical dose-response models to quantify KERs [*e.g.*, 11], or are based on an underlying systems biology (SB) model [*e.g.*, 10]. We propose here the use of Bayesian networks (BNs). Unlike SB models which may contain feedback and feed-forward loops, but like AOPs, BNs use acyclic directed graphs as their organizational structure [12]. The links between their nodes correspond to simple statistical dependencies. Thus, BNs can be viewed as an intermediate approach between empirical models and SB models. They have already been applied to AOPs

in the area of skin sensitization to facilitate potency assessment for classification purposes and to support hazard characterization in a semi-quantitative way [13,14]. Here, we demonstrate the application of (dynamic) BNs to AOP quantification. Moreover, we compare the results obtained by using the three modeling approaches mentioned above (*i.e.*, linked statistical dose-response relationships, dynamic BNs, and SB models) to quantify a chronic kidney disease (CKD) AOP.

CKD is a multifactorial progressive syndrome, of which aging, cardiac insufficiency, diabetes and chemical-induced nephrotoxicity are both initiating and accelerating factors. All of these factors have an oxidative stress component: Nrf2 activation and its downstream genes, including Heme Oxygenase-1, are markers of CKD [15,16]. Indeed, pharmaceutical activation of Nrf2 via bardoxolone methyl has been shown to slow CKD progression [17]. In order to develop the BN framework we have taken a relatively simple *in vitro* model, i.e. differentiated RPTEC/TERT1 cells treated with the oxidant and renal carcinogen potassium bromate ($KBrO_3$) [18,19]. Previous evidence suggests that $KBrO_3$ is a thiol reactive oxidant and a strong inducer of Nrf2 and mitochondrial injury [18,20]. Mitochondrial perturbation through increased production of reactive oxygen species is a well-described phenomenon [21]. Also we have shown using the same biological model that several compounds both activate Nrf2 and increase glycolysis rates and subsequent lactate production. These compounds include cadmium chloride, chloroacetaldehyde, cidofovir, cisplatin and cyclosporine A [22–25]. Thus enhanced lactate production is a good surrogate for decreased mitochondrial capacity in RPTEC/TERT1 cells, where HIF-1 alpha is not suspected to be activated (which would also lead to increased lactate production).

Finally, we present the implementation of the developed qAOP in Effectopedia, an OECD software tool that aims to gather experimental data and models in a unified representation, so that they can be compared, further analyzed, and used for hazard and risk assessment

purposes. Effectopedia is compliant with the OECD guidance document for the development and assessment of AOPs, according to its supplemental Users' Handbook [4,26].

2 Methods

2.1 Chronic kidney disease AOP

The proposed AOP (Figure 1) links thiol oxidation to CKD via oxidative and mitochondrial stress. Within the nephron, the proximal tubule is especially susceptible to injury from oxidative chemicals, as they can cause mitochondrial damage, which in turn can result in impairment of active and secondary transport, as well as in cell death. CKD is characterized by a progressive loss of renal function, the onset of which is initiated and/or accelerated by other factors such as diabetes, high blood pressure or exposure to nephrotoxic chemicals [22,27]. Given its high energy demand for active transport, the proximal tubule is especially susceptible to injury from oxidative chemicals and mitotoxins [28]. This AOP fulfills several, but not all, Bradford Hill criteria for weight-of-evidence assessment [8]. It should therefore be considered as preliminary, rather than definitive:

- The experimental dose-response relationships corresponding to the various KERs are consistent with the expected effects (depletion of GSH is associated with increased DCF, and increased lactate production, see the data in Figures 5-7 for example).
- There is temporal concordance among the key events and adverse outcome: GSH depletion is fast, DCF production slower, lactate response even slower (Figures 5-7), and CDK takes years to develop.
- The biological plausibility, causal coherence and consistency of experimental evidence for this AOP is strong (note that a full documentation of this evidence would extend beyond

scope and length of this paper). This AOP is being ported to the AOPWiki (https://aopwiki.org) and more supporting information should be available there soon.

- Alternative mechanisms are possible, as further discussed in this paper, but they would not exclude the mechanism postulated by the AOP.
- A major data gap is that we do not yet have data on a sufficient number of chemicals to assess the strength, consistency and specificity of the association of the AO to the MIE or the potential impact of alternative mechanisms. Consequently, uncertainties are quite high, as will be shown by the SB model, and the proposed AOP is not yet very well characterized.

Here, we quantified this AOP up to the initiation of cell death following induction of oxidative stress, since our analysis is based on *in vitro* data obtained in human proximal tubule (RPTEC/TERT1) cells exposed to $KBrO_3$. The link from cell death to kidney function impairment therefore cannot be modeled based on the available data and we will focus on a set of core early KEs leading to proximal tubule damage.

2.2 Experimental data

Thiol oxidation following exposure to various concentrations of potassium bromate ($KBrO_3$) (control, 0.375, 0.75, 1.5. 3, and 6 mM) was measured through glutathione (GSH) depletion in a cell-free environment (see Figure 5). Depletion was measured in triplicates after 1 hour, using the luminescence-based GSH-Glo kit from Promega (V6912), according to manufacturer's instructions, as described in Limonciel *et al*. [18].

Oxidative stress was quantified using the Invitrogen™ Carboxy-H2DCFDA test (catalog #C400). In brief, the cell permeant reagent 6-carboxy-2',7'-dichlorodihydrofluorescein diacetate (cH2DCFDA) is first introduced in the culture medium. After diffusion into cells, it

is deacetylated by cellular esterases to 6-carboxy-2',7'-dichlorodihydrofluorescein (cH2DCF), which remains trapped in the cell and is oxidized by hydroxyl, peroxyl radicals and other reactive oxygen species (ROS) to 6-carboxy-2',7'-dichlorofluorescein (cDCF), which is highly fluorescent. RPTEC/TERT1 cells were grown as described by Aschauer *et al*. [29] and exposed to various concentrations of $KBrO_3$ (control, 0.75, 1.5. 3, and 6 mM) as described by Limonciel *et al*. [18]. Briefly, cells were grown and matured into a mature monolayer in 96-well cell culture plates kept at 37°C / 5% CO_2 and were fed fresh medium 24 hours before chemical exposure. Cells were incubated with 40 µM cH2DCFDA 4 hours before washing out the excess extracellular dye and starting exposure to $KBrO_3$ dissolved in culture medium. cDCF production was measured over time (approximately every 15 min, up to 24 hours, in 8 replicates) as relative fluorescence units (RFU) by fluorescence spectroscopy using a Tecan Pro M200 microplate reader.

Mitochondrial injury was estimated by lactate concentration in collected RPTEC/TERT1 cell culture supernatants, measured in quadruplicates at the start of the experiments and then every 24 hours, up to 3 days, following exposure to various $KBrO_3$ concentrations (control, 0.25, 0.5, 1, 2, and 4 mM) (see Figure 6). Supernatant lactate production rate is a measure of glycolysis rate, and increased glycolysis can be due to a decrease in mitochondrial respiration [30]. The culture medium, with the given $KBrO_3$ concentrations, was changed every day after an aliquot was taken for lactate measurement using the absorbance-based assay described in Limonciel *et al*. [15].

2.3 Dose-response based qAOP

In the empirical dose-response approach, dose(-time)-response equations were fitted to data on the effect of $KBrO_3$ on GSH, cH2DCF, and lactate. With such data, linking chemical

exposures to KEs, the corresponding equations need to be mathematically inverted to obtain chemical-independent KERs. Only the exposure to MIE relationship can be used as is. For example, if we have three data sets for the activity at dose D_X of chemical X on each KE of an AOP, we need to fit three dose-response equations:

$$KE_1 = f(D_X) \tag{1}$$

$$KE_2 = g(D_X) \tag{2}$$

$$KE_3 = h(D_X) \tag{3}$$

The relationship between KE_1 and D_X is given directly by Eq. 1 However, the relationship between KE_1 and KE_2 needs to be derived from Eqs 1 and 2:

$$KE_2 = g(D_X) = g\left(f^{-1}(KE_1)\right) \tag{4}$$

where f^1 denotes the inverse function of f. Similarly, for the relationship between KE_3 and KE_2 we have:

$$KE_3 = h(D_X) = h\left(g^{-1}(KE_2)\right) \tag{5}$$

For dose-time-response relationships, the principle is the same, with time as an extra variable in the above functions. However, in some cases the function may not be monotonic and therefore will not be invertible.

The relationship between the concentration of $KBrO_3$ (C_{KBrO3}) and the percentage of GSH (Pct_{GSH}) remaining *in vitro* after one hour, representing the MIE, was modeled with a modified exponential decrease equation (Eq. 6):

$$Pct_{GSH} = 100 \times exp\left(-k \cdot C_{KBrO_3}^{b}\right) \tag{6}$$

Its parameters are the GSH degradation rate constant k, and power b (which increases the degradation rate if b>1).

The inverse of Eq. 6 is:

$$C_{KBrO_3} = \left(\frac{log(100) - log(Pct_{GSH})}{k}\right)^{1/b} \quad (7)$$

The relationship between C_{KBrO3}, time t and the quantity of cDCF formed (Q_{cDCF}, reflecting the amount of oxidative stress) was modeled empirically by Eq. 8:

$$Q_{cDCF} = A + B \cdot \left(1 + \delta - exp\left(-k_d \cdot C_{KBrO_3}\right)\right)\left(1 - exp(-k_t \cdot t)\right) \quad (8)$$

Its parameters are A (baseline response), B (maximum increase above baseline), δ (maximum increase modulation by dose), k_d (dose coefficient), k_t (time coefficient).

The solution of Eq. 8 for C_{KBrO3} is:

$$C_{KBrO3} = log\left(\left(1 + \delta - \frac{Q_{cDCF} - A}{B \cdot (1 - exp(-k_t \cdot t))}\right)^{-k_d}\right) \quad (9)$$

Replacing C_{KBrO3} in Eq. 8 by the expression given in Eq. 7, we obtain the following KER between Pct_{GSH} and Q_{cDCF}:

$$Q_{DCF} = A + B \cdot \left(1 + \delta - exp\left(-k_d \cdot \left(\frac{log(100) - log(Pct_{GSH})}{k}\right)^{1/b}\right)\right)\left(1 - exp(-k_t \cdot t)\right) \quad (10)$$

To model the C_{KBrO3} - time - lactate concentration (C_{lac}) relationship, we used a polynomial equation which adequately fitted the data:

$$C_{lac} = a + bC_{KbrO_3} + (c + eC_{KbrO3})t + (d + fC_{KbrO3})t^2 \quad (11)$$

If we replace C_{KBrO3} in Eq. 11 by the value given in Eq. 8, the relationship between Q_{cDCF}, time and C_{lac} becomes:

$$C_{Lac} = a + b \cdot log\left(\left(1 + \delta - \frac{Q_{cDCF} - A}{B \cdot (1 - exp(-k_t \cdot t))}\right)^{-k_d}\right) + \left[c + e \cdot log\left(\left(1 + \delta - \frac{Q_{cDCF} - A}{B \cdot (1 - exp(-k_t \cdot t))}\right)^{-k_d}\right)\right] t + \left[d + f \cdot log\left(\left(1 + \delta - \frac{Q_{cDCF} - A}{B \cdot (1 - exp(-k_t \cdot t))}\right)^{-k_d}\right)\right] t^2 \quad (12)$$

A relationship (even more complex) between GSH and lactate concentration could be obtained by replacing Q_{cDCF} by Pct_{GSH}, using Eq. 10.

2.4 Bayesian network qAOP

A BN is a probabilistic model whose underlying structure is a graph (equivalently, a network) where each node represents a variable of the problem (*i.e.*, for an AOP: chemical substance, MIE, KEs and AO), and each arc between two nodes represents a direct dependency (ideally, a causal relationship) [31]. The AOP shown on Figure 1 can be taken as a BN structure. Within such a BN, a probabilistic relationship (specifically, a component of a conditional distribution function) is defined by each arc linking two variables. For example, if an arc joins variables *A* and *B*, a relationship such as "*A* is distributed normally around $k \square B$, with a variance equal to s^2" has to be defined. As a result, every node of the network has a probability distribution conditioned by other network variables. This implies that a variable cannot depend upon itself, even indirectly, and therefore cycles are not allowed in BNs. Evidence on a set of nodes (for example, measurement of some KEs) affects the probability distributions of all their dependent nodes [32]. Training a BN from data means that one searches for those dependencies (and associated distributions) between variables that best explain the data. On the other hand, calibrating a BN implies estimating the parameters of the

distribution functions that link variables. In our case, we do not need to learn our BN's structure, since it is given by the AOP in Figure 1, but we need to calibrate it.

However, standard BNs do not provide a direct mechanism for representing temporal dependencies. Given that we have dose-time-response data on cDCF and lactate production, and that their time evolution is progressive rather than instantaneous, it is natural to use a dynamic Bayesian network (DBN) to integrate those data [33]. DBNs, typically, replicate an underlying structure at several (discrete) times corresponding to measurement time points. Figure 2 shows the DBN we constructed to quantify the chronic kidney disease AOP. Each node of a given time slice may depend on nodes in the previous time slice and on nodes in the same time slice [34]. In this way, the value of a node at time t_i may depend on its own value at time t_{i-1}, without introducing a loop in the graph. For example, in Figure 2, the cDCF readout at a given time point depends on its previous value (indeed, in the *in vitro* system cDCF accumulates with time). The same applies to the lactate concentration. There are also some instantaneous or constant dependencies: We considered that C_{KBrO3} was constant with time throughout the experiments (note that this is an approximation, but we have no information on the kinetics of $KBrO_3$ in the *in vitro* system). The thiol depletion readout (GSH level remaining after 1 hour) is simply an indicator of $KBrO_3$ potency and was also taken to be constant. The DBN structure being defined, we now turn to the form of the conditional distributions linking the nodes.

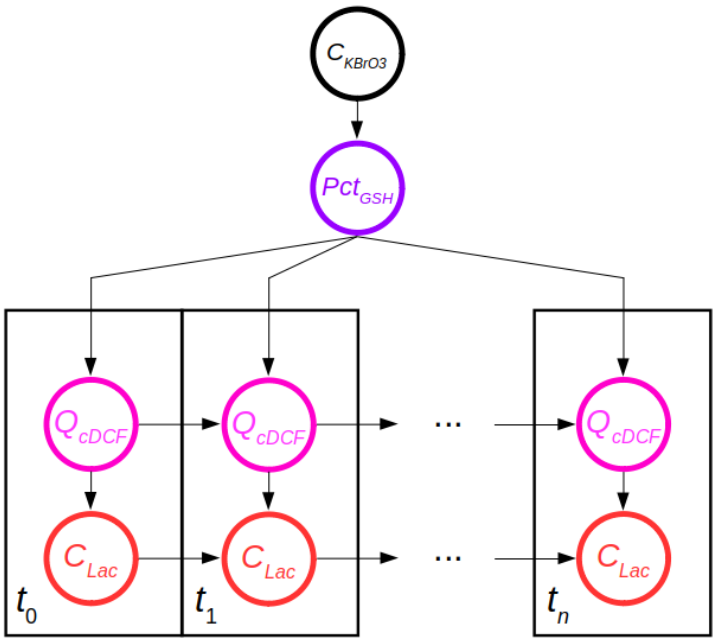

Figure 2: Structure of the dynamic Bayesian network qAOP for chronic kidney disease. $KBrO_3$ concentration and the GSH readout do not vary with time, while the cDCF and lactate readouts were observed at different time intervals. The arrows indicate probabilistic dependencies.

2.4.1 Node to node relationships

For the dependence of observed Pct_{GSH} on C_{KBrO3} we use a simplified probabilistic version the dose-response based qAOP (*cf.* Eq. 6):

$$Pct_{GSH} \sim Normal(100 \times exp(-k_{GSH} \cdot C_{KBrO3}), \sigma^2_{GSH}) \quad (13)$$

with depletion rate constant k_{GSH} and variance σ^2_{GSH}. Note that for simplicity we set parameter b to 1.

The conditional distribution of Q_{cDCF} observations at time t, given Pct_{GSH} and the Q_{DCF} observation at the previous time t - h, is given by an extension of the standard DBN model in

which $Pct_{GHS,t}$ influences the equilibrium value ($E_{cDCF,t}$) for $Q_{DCF,t}$ to which it converges over time at exponential dampening rate v:

$$Q_{DCF,t} \sim Normal\left(E_{cDCF,t} - \left[E_{cDCF,t} - Q_{cDCF,t-h}\right] \cdot e^{-v_{cDCF}h}, \sqrt{\frac{1-e^{-v_{cDCF}h}}{1-e^{-v_{cDCF}}}} \cdot \sigma^2_{cDCF}\right) \quad (14)$$

$$E_{cDCF,t} = \beta_{0,cDCF} + \beta_{cDCF} Pct_{GSH} \quad (15)$$

where $E_{cDCF,t}$ is the equilibrium value of Q_{cDCF} (a linear function of Pct_{GSH}), h is the (positive) time interval between two consecutive observations, v_{cDCF} (positive), $\beta_{0,cDCF}$, β_{cDCF}, and variance σ^2_{cDCF} are the parameters to estimate. The cDCF RFU value at time zero, $cDCF_0$, was not measured, but it should be different from zero given the 4-hour pre-treatment phase in the protocol and was therefore also estimated. Positive values of v and h ensure that e^{-vh} is bounded between 0 and 1.

A similar relationship was used for lactate by replacing $Q_{cDCF,t}$ by $C_{lac,t}$, and Pct_{GHS} by $Q_{cDCF,t}$ in Eqs 14 and 15. Given the recurrent experimental change of medium during the experiment, lactate concentration was set to zero at the start of the experiment and reset to that value every 24 hours.

2.5 Systems biology model

We used a SB model to analyze the oxidative stress (cDCF) data. The model does not describe lactate formation and hence we did not consider the lactate data in this approach. It focuses on the control of the oxidative stress by Nrf2 and glutathione, one of the major toxicity pathways studied in systems toxicology [30,35,36]. Therefore, we used it only to study the relationship between $KBrO_3$ exposure, time, and cDCF fluorescence in detail (the

model is in fact a detailed representation of the KER linking the MIE and first KE of the AOP in Figure 1).

The Nuclear Factor (Erythroid-derived 2)-Like 2 NFE2L2 pathway, commonly known as Nrf2, is an important adaptive response to oxidative stress [37]. In homeostatic conditions, Nrf2 is mostly bound to the cytoskeleton-associated Kelch-like-ECH-Associated Protein 1 (Keap1) in an inactive complex within the cytoplasm, which facilitates Nrf2 degradation. Upon oxidative stress, Keap1 is oxidized and the complex dissociates, and Nrf2 can migrate to the nucleus [38], where it activates the transcription of a set of target genes implicated in the metabolism and transport of xenobiotics, and ROS scavenging by GSH [39]. When the intra-cellular level of ROS exceeds the capacity of this defense system to replenish GSH through new synthesis, GSH depletion occurs and the remaining ROS cause extensive cellular damage, cell death, nephron attrition and CKD.

Figure 3 shows the SB model we developed to study the transcriptional regulation of the GSH pathway by the Nrf2 - Keap1 complex, which merges variants of the Hamon *et al.* model for RPTEC/TERT1 cells [40] and a model developed by Geenen *et al.* [41]. Whereas Hamon's model is more elaborate with respect to genes transcription, the Nrf2 - Keap1 interaction, and the role of ATP, Geenen's model details GSH synthesis. Coupling the two models greatly improves the description of the regulation of oxidative stress in RPTEC/TERT1 cells. When merging the two models, we simplified the description of the transcription/translation process, without loss of precision in the predictions: We replaced a cascade of differential equations operating at quasi-steady-state (for gene/receptor binding/unbinding, transcription induction by activator(s), translation and degradation of mRNA) by a single Hill equation. We also simplified the folate cycle (which was at steady-state), replacing it by a constant tetrahydrofolate concentration. We removed acetaminophen- and cyclosporine-specific reactions as these were irrelevant for the current application. We finally included the

description of the dynamics of HMOX1 and SRXN1 genes, which are often used as activation markers for the Nrf2 pathway. The simplification protocol, its validation and the resulting Hill-based equations are presented in Supplemental Material (Figure S1, Supplemental Protocol and Table S8).

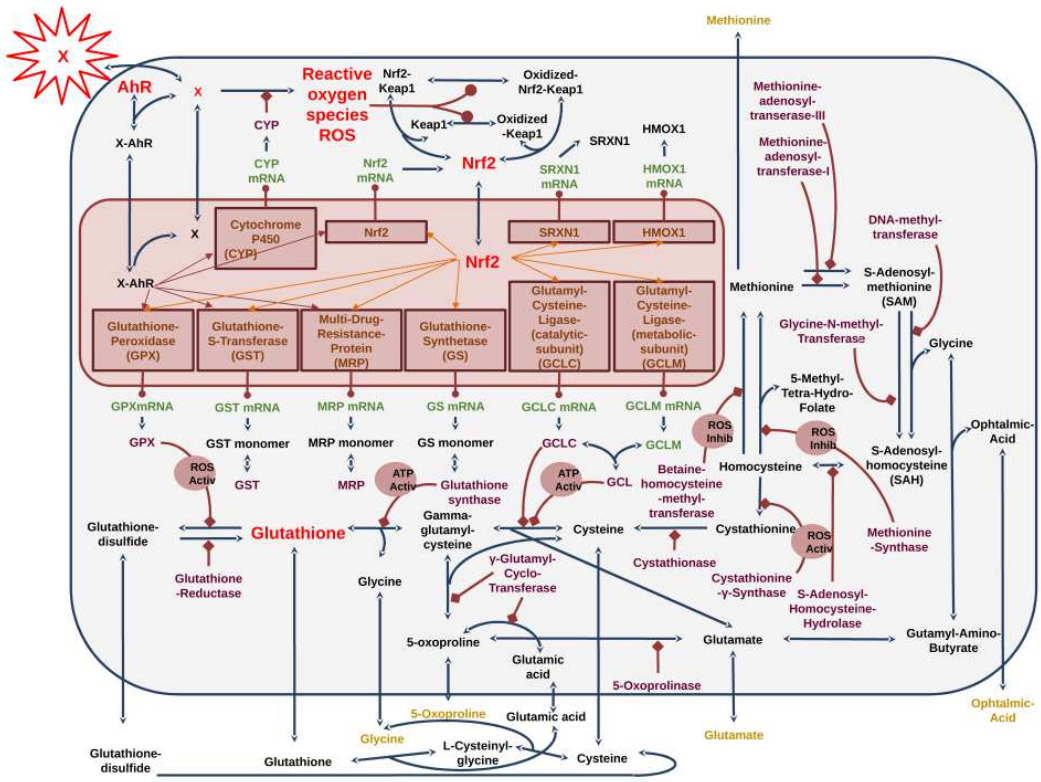

Figure 3: Schematic overview of the assembled SB model. All steps are modeled. This model covers both transcriptional and biochemical aspects of GSH synthesis and metabolism and its control by the *Nrf2-Keap1* signaling pathway. The blue compartment is cytosol and the red one is nucleus. Blue arrows show reactant(s):product(s) exchange during biochemical or transport reactions, and red arrows indicate enzymatic catalysis (diamond heads) or gene transcription (round heads). In the nucleus, red boxes represent genes and arrows indicate gene activation. Names of genes are in orange, of *mRNA* are in green, of enzymes are in purple, of other proteins and metabolites in blue and of extracellular constants in yellow.

In order to calibrate the model with our experimental data on the effect of $KBrO_3$ on GSH and cH2DCF, we added several first order reactions to the model (Figure 4):

a. Action of $KBrO_3$ on extra-cellular GSH (parameter $k_{GSHe,KBrO3}$);

b. Formation of cDCF from cH2DCF by ROS-mediated oxidation (parameter $k_{cDCF,ROS}$);

c. MRP-driven efflux [42,43] or passive bleaching of cDCF (the same parameter, k_{bl}, can describe the two processes);

d. Formation of cDCF from cH2DCF by direct action of $KBrO_3$ (parameter $k_{cDCF,KBrO3}$);

e. Action of $KBrO_3$ on intra-cellular GSH (parameter $k_{GSHc,KBrO3}$, which is multiplied by $k_{GSHe,KBrO3}$ to yield the reaction rate constant, and is in fact the ratio of the external to internal reaction rate constants).

The complete model code (with 57 differential equations and 335 parameters) is given as Supplemental Material.

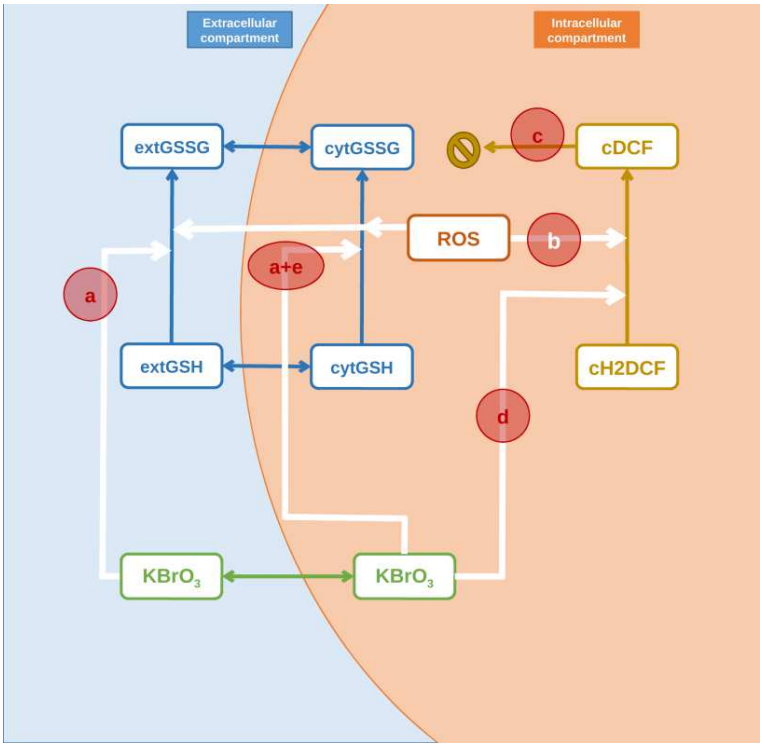

Figure 4: Potassium bromate ($KBrO_3$) and 6-carboxy-2',7'-dichlorofluorescein (cDCF) specific reactions of the SB model. Other abbreviations: extGSH is extra-cellular glutathione; cytGSH: cytosolic glutathione; extGSSG: extra-cellular oxidized glutathione; cytGSSG: cytosolic oxidized glutathione; ROS: reactive oxygen species; cH2CFD: 6-carboxy-2',7'-dichlorodihydrofluorescein. Reactions are represented by red circles: *a* is the oxidation of extGSH by $KBrO_3$; *b*: oxidation of cH2CFD by ROS; *c*: cDCF efflux or bleaching; *d*: oxidation of cH2DCF by $KbrO_3$; *e*: oxidation of cytGSH by $KBrO_3$.

2.6 Parameter estimation

Parameter calibrations for the three types of qAOP investigated were done in a Bayesian statistical framework, using Markov chain Monte Carlo (MCMC) simulations [44,45], or Hamiltonian MCMC [46]. Basically, for each parameter to calibrate, a prior distribution summarizing existing knowledge was updated on the basis of the likelihood of the current data to yield a posterior distribution. Those distributions were obtained by random sampling

from several simulated Markov chains. The convergence of the simulated chains was checked using the R_{hat} criterion of Gelman and Rubin [47].

The complexity of the various qAOP models differed and slightly different sampling strategies were used. For the dose-response based qAOP, we used a Metropolis-Hastings MCMC algorithm, as implemented in the *GNU MCSim* software [48]. Two Markov chains of 50,000 iterations were run in parallel, keeping one in four of the last 40,000 iterations. For each estimated parameter, non-informative uniform prior distributions were used (note that the boundaries of those prior distributions were never reached) (see Table S1 in Supplemental material). As usually done for measurements at different concentrations, the data were considered to be log-normally distributed with geometric means given by the corresponding model predictions and geometric standard deviations (σ_{GSH}, σ_{cDCF}, and σ_{lac}), sampled from half-normal distributions (with *a priori* about 5%, 20% and 20% precision respectively, see Table S1 in Supplemental material). Note that in this qAOP, the statistical error model (*i.e.*, the likelihood of the data) is clearly separated from the structural equations.

For the BN qAOP, posterior parameter distributions were obtained by Hamiltonian MCMC, using the *Stan* software [49]. Three simulated Markov chains were run in parallel for 12,000 iterations, keeping the last 6,000 iterations. Non-informative uniform prior distributions were used for each parameter except for the parameters in the cDCF - time - lactate portion of the model where weakly informative Gaussian priors were used to stabilize inference (see Table S2 in Supplemental material). In this qAOP model, the data likelihood is embedded in the model formulation. There is one clear constraint for this model: time and exposure conditions *must* match for all the variables describing a particular node to node relationship. For example, lactate was measured every 24 hours and depends on cDCF, which was measured every 15 minutes, but for different $KBrO_3$ concentrations. Therefore we need to statistically

"impute" (randomly draw from their conditional distribution) the expected cDCF values at the concentrations used in the lactate experiment. Note that the alternative solution of describing the cDCF dynamics only at time points zero and 24 hours would discard most of the cDCF data and is thus unsatisfactory. To reduce the number of data points to be imputed, we chose to use only one in four cDCF data points (one per hour).

For the SB model, parameter calibration was done with Metropolis-Hastings MCMC with *GNU MCSim*. Three Markov chains of 30,000 iterations were run in parallel, keeping the last 5,000 iterations. For each estimated parameter, non-informative uniform prior distributions were used (see Table S3 in Supplemental material). The data were considered to be log-normally distributed with geometric means given by the corresponding model predictions and geometric standard deviations σ_{cDCF} (see Table S3 in Supplemental material). The data likelihood is clearly separated from the structural equations. To calibrate the model with our experimental data on the effect of $KBrO_3$ on GSH and cH2DCF, we proceeded step by step, increasing the complexity of the model by introducing reactions according to the following schedule:

1. Action of $KBrO_3$ on extra-cellular GSH (parameter $k_{GSHe,KBrO3}$), on the basis of the $KBrO_3$ - GSH cell-free experimental data; $k_{GSHe,KBrO3}$ was held at its maximum posterior value in the subsequent steps.
2. Action of $KBrO_3$ on extra-cellular GSH (parameter $k_{GSHe,KBrO3}$) and formation of cDCF by ROS-mediated oxidation ($k_{DCF,ROS}$): this represents a minimal model for explaining the $KBrO_3$ - time - cDCF data.
3. Adding efflux or bleaching of cDCF (k_{bl}).
4. Adding the direct formation of cDCF by $KBrO_3$ ($k_{cDCF,KBrO3}$) (step *4a*) or the action of $KBrO_3$ on intra-cellular GSH ($k_{GSHc,KBrO3}$) (step *4b*).
5. All of the above.

To compare the eventual improvement in fit brought about by the various model refinements we used various measures of model fit to the data: the data log-likelihood, the residual geometric standard deviation (GSD), the Akaike information criterion (AIC) (twice the difference between the number of parameters and the data log-likelihood), the Bayesian information criterion (BIC), and the Deviance information criterion (DIC) [50].

2.7 Uncertainty propagation

The output of MCMC simulations is a sample of parameter sets (or parameter vectors) drawn from their joint distribution. Those sets of parameter values were used to rerun the corresponding model to make predictions for unobserved values. This is a type of Monte Carlo simulations in which the MCMC sampler acts as a random parameter values generator. We obtained distributions of predicted values that reflects the uncertainty of all parameter values. For example, when using Eq. 12 to compute a lactate concentration, the uncertainty of all parameters entering the equation was convolved by Monte Carlo sampling and their uncertainty was fully propagated to the result. The same applies to the other models we used.

2.8 Software

The dose-response based qAOP and the SB model were simulated and calibrated with the *GNU MCSim* software, version 5.6.6 (hwww.gnu.org/software/mcsim) [48]. The BN qAOP model was simulated and calibrated using *Stan* (mc-stan.org) [49]. All plots were created with *R*, version 3.4.4 (cran.r-project.org) [51]. Effectopedia version 1.2.51 (www.effectopedia.org) [52] was used for implementation of the qAOP on the internet.

3 Results

3.1 Dose-response based qAOP

The empirical dose response models given by Eqs 6, 8, and 11 described the $KBrO_3$ - GSH, $KBrO_3$ - time - cDCF, and $KBrO_3$ - time - lactate relationships reasonably well (see Figure 5 and Figure 6, top row). Equivalent 2D representations of the time course of cDCF and lactate at the various $KBrO_3$ concentrations are given in Supplemental material Figures S2 and S3, respectively. The uncertainty of the model predictions is low for GSH (Figure 5), and it amounts to about 0.5% to 1.5% for cDCF and 5% to 12% for lactate (this cannot be usefully visualized in Figure 6 for reasons of readability). Residual uncertainty (an estimate of measurement error) is about 22% for GSH, 20% for cDCF and 30% for lactate. Table 1 summarizes the posterior distributions of the parameter values obtained by Bayesian calibration.

Table 1: Summary of the posterior parameter distributions for the dose-response based qAOP fitted to GSH, cDCF and lactate data.

Parameter	**Units**	**Maximum posterior value**	**mean (SD) [2.5pctile, 97.5pctile]**
$KBrO_3$-GSH model			
k	$1/mM^b$	1.44	1.44 ± 0.06 [1.32, 1.56]
b	-	0.73	0.73 ± 0.028 [0.68, 0.79]
σ_{GSH}	%	1.22	1.22 ± 0.022 [1.18, 1.27]
$KBrO_3$-time-cDCF model			
A	RFU	2100	2100 ± 33 [2000, 2200]
B	RFU	12500	12500 ± 210 [12200, 12800]
δ	-	0.21	$2.1\times10^{-1} \pm 5.3\times10^{-3}$ [0.2, 0.22]
k_d	1/mM	0.62	$6.2\times10^{-1} \pm 1.7\times10^{-2}$ [0.6, 0.65]
k_t	1/h	0.14	$0.14 \pm 6.7\times10^{-3}$ [0.13, 0.15]
σ_{cDCF}	RFU	1.19	1.19 ± 0.0022 [1.18, 1.19]
$KBrO_3$-time-lactate model			
a	mM	2.9	2.8 ± 0.22 [2.4, 3.2]
b	-	-6.2×10^{-2}	$-5.0\times10^{-3} \pm 0.11$ [-0.18, 0.18]
c	mM/h	-0.057	$-5.5\times10^{-2} \pm 0.015$ [-0.080, -0.030]
d	mM/h^2	1.0×10^{-3}	$0.001 \pm 2.2\times10^{-4}$ [6.5×10^{-4}, 0.0013]
e	1/h	0.041	$0.040 \pm 9.6\times10^{-3}$ [0.023, 0.056]
f	$1/h^2$	-3.8×10^{-4}	$-3.7\times10^{-4} \pm 1.5\times10^{-4}$ [-6.1×10^{-4}, -1.2×10^{-4}]
σ_{lac}	mM	1.27	1.28 ± 0.026 [1.24, 1.34]

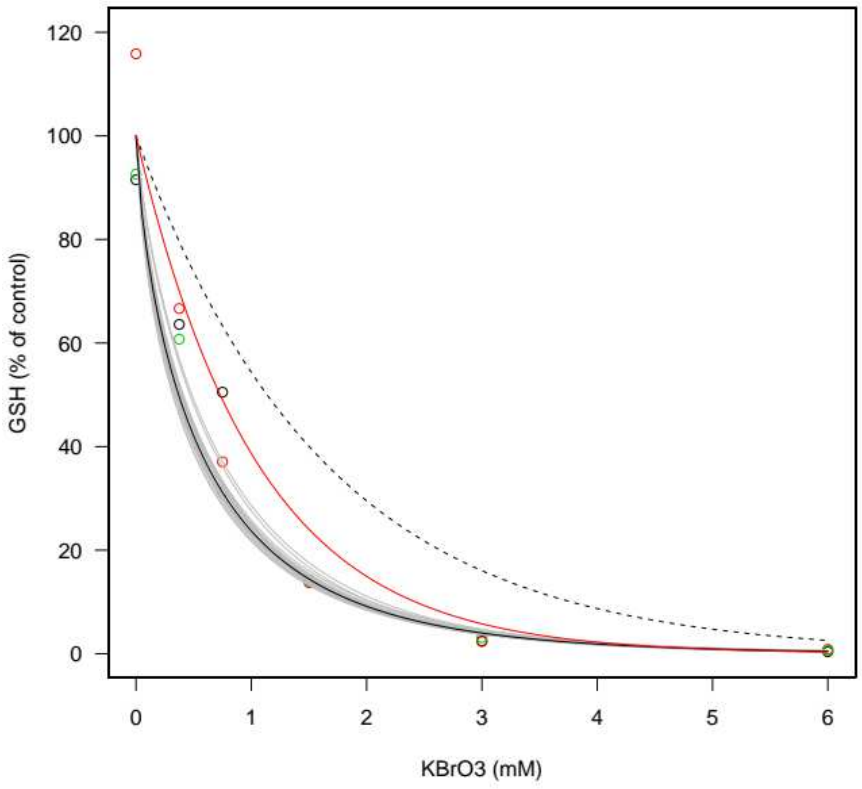

Figure 5: Fit of the $KBrO_3$ - GSH data (circles; each color represents one of the replicates) using the three qAOP models developed. The black line corresponds to the empirical model (Eq. 6). The best fit (solid line) is shown along with 20 additional random fits (gray), showing the uncertainty of the model predictions. The black dashed line represents the best fit obtained with the DBN qAOP. The red line shows the best fit for the SB model.

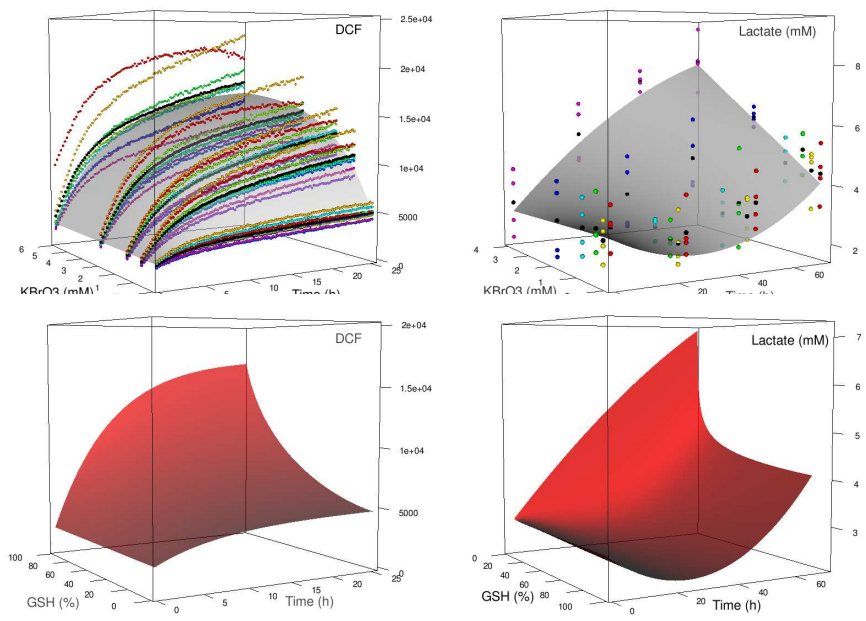

Figure 6: Fit (top row) and predictions (bottom row) of the dose-response based qAOP for the cDCF (measured in RFU) (left) and lactate (right) readouts. The best fit surfaces (gray) are plotted along with all individual data (colored dots). The predicted chemical-independent relationships (in red) for GSH - time - cDCF, or GSH - time -lactate were obtained by inversion of the qAOP equations (see text). The maximum posterior parameter values given in Table 1 were used to draw the figures.

By inversion of the empirical models, we can deduce the relationship between GSH, time, and cDCF or GSH, time, and lactate production (Figure 6, bottom row). These relationships should, in theory, be independent of the thiol-reactive chemical. They can be used to make predictions, including full parametric uncertainty propagation since we used a Bayesian statistical framework for parameter inference. For example, a dose causing 80% reduction of GSH after 1 hr (*i.e.*, 20% GSH left), *in the test conditions described in Methods*, should lead to a lactate concentration of 4.6 ± 0.3 [4.1, 5.1] mM (mean, SD, 5 and 95 percentiles) after 3 days of exposure.

3.2 Bayesian network qAOP

The fit of the DBN qAOP to GSH, cDCF, and lactate data is shown on Figures 5 and 7. Equivalent 2D representations are given in Supplemental material Figures S4 and S5. The fits for GSH and cDCF are less good than those of the empirical models. The fit to the lactate data (Figure 7) looks very different for the DBN model, compared to the empirical model, because the DBN model takes into account the change of medium every 24 hours. Note that all parameters of the DBN model are estimated together, so that modeling errors are spread over the overall dataset. Also, the model uses linear relationship between nodes, except for the link $KBrO_3$ - GSH. Residual uncertainty (an estimate of measurement error) is about 50% for GSH, 25% for cDCF and 10% for lactate. The error model, however, is different (normally distributed residuals, rather than log-normally distributed as in the empirical model). Table 2 summarizes the posterior distributions of the parameter estimates obtained. The model parameters have some physical interpretation: Parameter v controls the speed at which plateaus are reached in Figure 7. The β parameters condition the height of the plateaus. However, there is a subtle interplay between convergence speed, plateau level, time and dose, as can be seen in Figure S5. All parameters are significantly different from zero.

Table 2: Summary of posterior parameter distributions of the dynamic Bayesian network qAOP fitted to GSH, cDCF and lactate data.

Parameter	**Units**	**Maximum posterior value**	**mean (SD) [5pctile, 95pctile]**
$KBrO_3$-GSH link			
k_{GSH}	1/mM	0.61	0.75 ± 0.18 [0.48, 1.1]
σ_{GSH}	%	18	15 ± 6.3 [6.9, 27]
GSH-cDCF link			
$cDCF_0$	RFU	2160	2160 ± 20 [2130, 2190]
$\beta_{0,cDCF}$	RFU	1.89×10^4	$1.9\times10^4 \pm 1.1\times10^3$ [1.7×10^4, 2.1×10^4]
β_{cDCF}	RFU/%	-117	-130 ± 9.8 [-148, -117]
ν_{cDCF}	1/h	0.0783	0.10 ± 0.011 [8.2×10^{-2}, 0.12]
σ_{cDCF}	RFU	906	890 ± 10 [880, 910]
cDCF-lactate link			
$\beta_{0,lac}$	mM	9.68×10^{-3}	$1.7 \pm 3.9\times10^{-1}$ [1.05, 2.3]
β_{lac}	RFU/mM	4.05×10^{-4}	$2.5\times10^{-4} \pm 3.7\times10^{-5}$ [1.95×10^{-4}, 3.2×10^{-4}]
ν_{lac}	1/h	0.267	0.35 ± 0.064 [0.25, 0.46]
σ_{lac}	mM	0.185	0.64 ± 0.097 [0.48, 0.78]

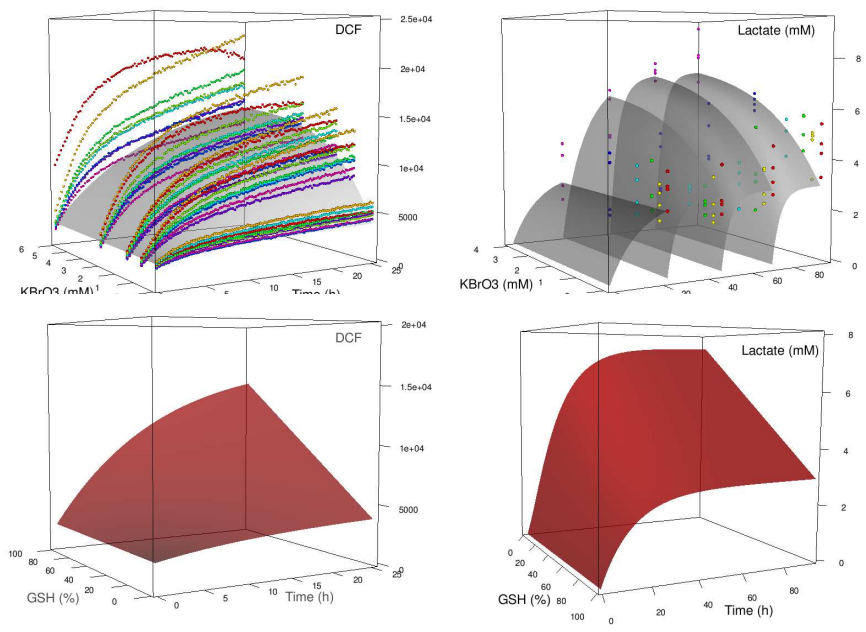

Figure 7: Fit (top row) and predictions (bottom row) of the dynamic BN qAOP for the cDCF (measured in RFU) (left) and lactate (right) readouts. The best fit surfaces (gray) are plotted along with the data mean (black dots) and all individual data (colored dots). The predicted chemical-independent relationships (in red) are shown for GSH - time - cDCF and GSH - time - lactate. The maximum posterior parameter values given in Table 2 were used to draw the figures.

The DBN qAOP model does not need mathematical inversion to produce chemical-independent predictions of the levels of cDCF and lactate as a function of GSH depletion and time, because they can be directly simulated (Figure 7, bottom row). The resulting relationship for cDCF is quite similar to that obtained with the previous qAOP, except for the linearity of the GSH - cDCF relationship. However, the GSH - lactate relationship is very different, even though constant exposures to $KBrO_3$ are simulated in both cases (the simulation is now considering a single medium change at time point zero). Note that lactate starts at zero to reach a plateau in about three days. The relationship between GSH and lactate is predicted to be linear by the DBN model, instead of being strongly nonlinear in the

empirical qAOP. As before, predictions with uncertainty estimates can be easily made. For example, the DBN qAOP predicts that a chemical dose causing 80% reduction of GSH after 1 hour (*i.e.*, 20% GSH left) leads to a lactate concentration of 5.8 ± 0.4 [5.2, 6.5] mM (mean, SD, 5 and 95 percentiles) after 3 days of exposure. This is significantly different from the prediction of the empirical qAOP.

3.3 Systems biology model

The fit of the SB model to the GSH data (calibration step 1) is shown in Figure 5 (red line). It is better than the fit of the DBN qAOP (residual uncertainty for the GSH data is about 40%), despite the fact that both use the same decreasing exponential relationship between $KBrO_3$ and GSH. However, $k_{GSHe,KBrO3}$ was calibrated to the data independently of the others parameters and its fit is not constrained by the cDCF data. The fits obtained for the $KBrO_3$ - time - cDCF data at the various model calibration steps (parameters were re-calibrated at each step) are shown on Figure 8. Equivalent 2D representations are presented in Supplemental material Figures S6 to S9. Measures of the quality of fit are given in Supplemental Material Table S4. Note that the model takes into account the 4 hours of cells pre-incubation with cH2DCFDA, and the simulation therefore starts already before exposure to $KBrO_3$ (which is defined to occur at time point zero). During that period of time, ROS already starts forming cDCF, explaining the relatively high level of fluorescence at time point zero. At step 2, with just a depletion of extra-cellular GSH by $KBrO_3$ and the formation of cDCF by ROS the model is unable to explain the data (Figure 8A). The depletion of extra-cellular GSH has only a minor effect on the intra-cellular GSH level (Supplemental Material Figure S6). Therefore, only background cellular ROS produces cDCF, at a constant rate, and the accumulation of cDCF is predicted to be linear (according to the experimental protocol cH2DCFDA is expected to be in excess, and not depleted). Allowing cDCF efflux or

bleaching offers an explanation for the leveling off of its fluorescence, yet the effect of $KBrO_3$ is still not explained satisfactorily and the data fit is very poor (Step 3, Figure 8B), despite the fact that efflux is mediated by MRP, which is under Nrf2 control (the MRP level stays at its background level and is not affected by $KBrO_3$, since intra-cellular GSH itself is not significantly lowered). Adding the possibility that $KBrO_3$ directly oxidizes cH2DCF improves the fit markedly (Step 4a, Figure 8C), and the residual error σ_{cDCF} goes down to about 20% (see Table 3). However, the effect of $KBrO_3$ is linear, which is not what the data show. In this case, also, the effect of $KBrO_3$ on ROS is close to zero. Instead of a direct oxidation of cH2DCF by $KBrO_3$, we also tested the possibility that $KBrO_3$ acts on intra-cellular GSH (Step 4b, Figure 8D). This has a clear effect on cDCF production, but it is extremely nonlinear and does not lead to a reasonable fit to the data. Finally, in step 5, we put all the above parameters in the model and re-calibrated them. This did not lead to an improvement compared to step 4a (see Supplemental Material Table 4) and the effect of $KBrO_3$ on intra-cellular GSH was estimated to be nearly absent (data not shown).

Table 3: Summary of the posterior distribution of the five SB model parameters describing the action of $KBrO_3$ on the formation of cDCF. The best parameterization (setting $k_{GSHc,KBrO3}$ at zero) is presented.

Parameter	**Units**	**Maximum posterior**	**mean (SD) [2.5pctile, 97.5pctile]**
$k_{GSHe,KBrO3}$	$(\mu M.s)^{-1}$	2.65×10-7	2.65×10-7 ± 8.45×10-9 [2.48×10-7, 2.81×10-7]
$k_{cDCF,ROS}$	$(zmol.s)^{-1}$	1.20×10-7	1.21×10-7 ± 3.2×10-9 [1.14×10-7, 1.27×10-7]
k_{bl}	s^{-1}	3.50×10-5	3.50×10-5 ± 1.4×10-6 [3.23×10-5, 3.77×10-5]
$k_{cDCF,KBrO3}$	$(\mu M.s)^{-1}$	1.22×10-9	1.22×10-9 ± 4.5×10-11 [1.13×10-9, 1.30×10-9]
$k_{GSHc,KBrO3}$	-	0	0
σ_{cDCF}	RFU	1.20	1.20 ± 6.8×10-3 [1.18, 1.21]

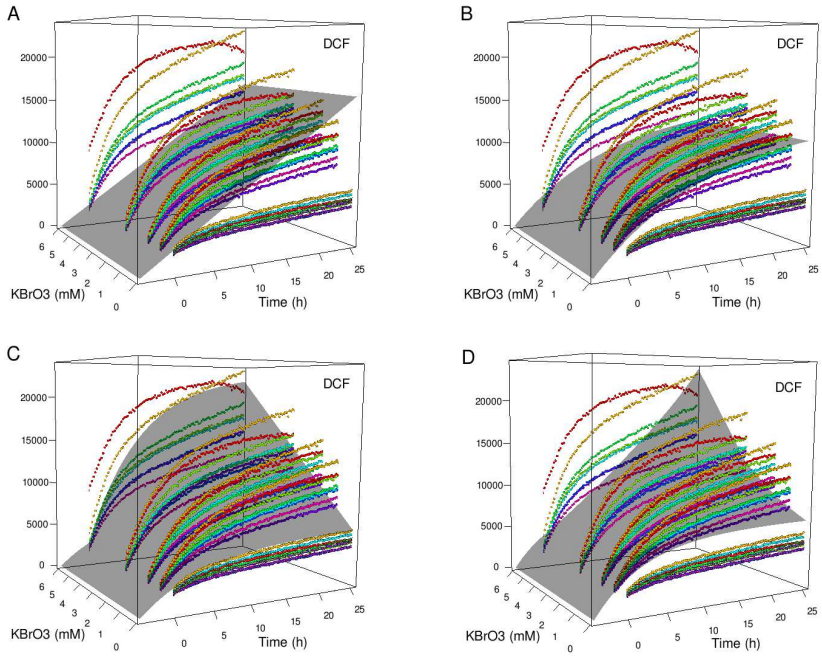

Figure 8: Best fits of SB model (gray surfaces) to the cDCF RFU data (colored dots), for different levels of model complexity: (A) action of $KBrO_3$ on external GSH and formation of cDCF by ROS; (B) same as A but with the addition of cDCF efflux or bleaching; (C) same as

B but with the addition of a direct formation of cDCF by $KBrO_3$; (D) same as B, but with the addition of an action of $KBrO_3$ on internal GSH.

Table 3 lists the best value (maximum posterior), the mean, the standard deviation and the confidence interval [2.5 percentile, 9.75 percentile] of each of the four parameters calibrated at step 4a (yielding the best and most parsimonious model). The values of the parameters directly related to cDCF do not have an explicit biological interpretation because cDCF is measured in RFU (which should be proportional to concentration, but with an unknown proportionality constant). Note that the cDCF efflux/bleaching rate constant corresponds to a half-life of about 6 hours. The SB model can also be used to make predictions, with full uncertainty propagation. For example, a 4 mM concentration of $KBrO_3$ is predicted to lead to a cDCF fluorescence of 16600 ± 250 [16200, 17100] RFU (mean, SD, 5 and 95 percentiles) after 24 hours.

3.4 Effectopedia implementation

Effectopedia provides a graphical user interface to build an AOP diagram, which in turn gives easy access to relevant descriptions, data and models. In addition to a qualitative description of the AOP, Effectopedia provides structure for representation of test methods, collected data and executable models implemented in the supported programming languages (R, MATLAB, Java). Effectopedia was used to create several iterations of the AOP diagram. Initially, the sequence of KEs included relevant feedback mechanisms or parallel processes (branches). However, in the following step of identification of measurement methods, some of these events did not have a separate method of observation and were therefore combined into a single KE. Other events were determined to be modification factors rather than being causally related to the AO and were removed from the pathway diagram. The current version of the

AOP diagram implemented in Effectopedia is shown in Supplemental Material Figure S10. Each of the elements in the diagram can be expanded and details can be added to their description. Models were implemented in R and their source code contributed to the description of the *in silico* models, allowing other users to execute them with the same and/or different data and model parameters.

4 Discussion

In this paper, we explored various options for quantifying an AOP and deriving chemical independent KERs. Quantitative AOPs have been described previously [10,11], but we strove for a rigorous statistical treatment of the models, which is particularly important for quantifying uncertainties associated with predictions and extrapolations. For that purpose, we used MCMC simulations in a Bayesian framework [44]. Dealing with dose-time-response data significantly complicates the problem and very few off-the-shelf software provide adequate tools and models for such data, despite the fact that time is a key variable in disease progression. While spatial structure is evident in AOP representations (from molecules to cells, tissues *etc*.), time is implicit, masking large time-scale differences: Molecular reactions typically take seconds, cells respond in hours, tissues in a matter of days, and the whole body can take years to be significantly affected. A qAOP considering only dose and assuming instantaneous or fixed-delay effects would be of limited usefulness for risk assessment. This is particularly true for chronic renal disease, as humans have a large renal functional reserve and ill health is only apparent when that reserve is breached. The time-course of exposure to stressors is also important and should be considered during qAOP calibration, because *in vitro* cellular concentrations of test chemicals are usually different from the nominal medium concentrations and change with time [53]. To that purpose, qAOPs can be linked with

pharmacokinetic models, but only if they are time-consistent. Nevertheless, in the absence of kinetic data on $KBrO_3$ concentrations *in vitro*, we considered here the nominal $KBrO_3$ concentrations to be an adequate measure of exposure. Note also that the AOP we used is illustrative and not OECD approved. We deliberately focused on a short sequence of KEs to demonstrate what can be achieved with different modeling approaches. The link to cell death and the subsequent link to kidney function impairment have not been included in our models given the absence of data on these downstream KEs. A final general comment is that for a complete AOP quantification, data on the effects of several chemicals on the KEs should be studied, in order to make sure that the KERs derived are fully chemical-independent. Such data are currently not available and our qAOP thus only serves to demonstrate quantification methods.

Table 4 summarizes the principle, as well as the pros and cons of the three approaches taken. All three require proper statistical treatment to propagate the uncertainty implied by imprecise data measurements through the AOP. An excellent way to propagate uncertainty, and translate it to risk assessment in the form of Monte-Carlo samples, is to use Bayesian model calibration [44,45,50]. In any case, relatively complex and specific software is required. It will be interesting to follow the development of Effectopedia, as it offers a user-friendly and toxicology-specific AOP quantification environment. Conditional on proper statistical treatment and user-defined modeling assumptions, all three methods can describe the data well, albeit with different constraints (discussed below in more detail). Note that consistent dose-time-response relationships found by any of the three methods do support causality and concur a posteriori to other Bradford-Hill criteria. Therefore, qAOP modeling could provide further validation of those criteria.

Dose-response based qAOPs may seem the easiest to develop, as most toxicologists understand what is dose-response modeling. However, such modeling is less user-friendly

than it seems. It requires either modeling skills (to find "good" dose-time-response equations) or black box curve-fitting approaches. Given the immense number of possible choices, finding the "best" model given the data is a very difficult task, and the question of structural model uncertainty is acute. This is compounded by the fact that the data are taken at "face value" and cannot be critically evaluated (except through residual analysis such as outlier detection, but these depend on the model adopted, which is still arbitrary). All this makes the domain of application of empirical qAOPs strictly limited to the time and dose range of the data, and strongly dependent on how relevant the experimental protocol is towards the actual disease process, without providing any indication of that relevance. Furthermore, for correct statistical inference and chemical-independent KERs, some or all dose-time-response relationships fitted must be mathematically inverted. Simply chaining such relationships (that is, using the best predictions for one KE as input to the next KER, as it is often done) does not account for uncertainties in the "independent" variable at each step and does not correctly propagate uncertainty through the AOP. The result would be a largely over-optimistic precision for predictions.

DBN qAOPs offer an automatic or standardized way to develop semi-empirical qAOPs, while tuning simply the complexity of the KERs. They can nicely describe complex time dependencies in the data, *e.g.*, they successfully modeled a fairly complex time-dose-relationship for the lactate readout (*cf.* Figure 7). The end-results differ visually from those of the dose-response qAOP, because in our DBN the KER links for cDCF and lactate are linearly related to GSH levels (we are currently working on nonlinear extensions of the DBN model). That DBN qAOP is, to our knowledge, the first attempt to use such a model for a continuous dose-time-response predictive model. To accommodate the time-dependency of the data, we used a special formulation of the DBN where time enters the KERs. The work of Jaworksa *et al.* [13,32,54] pioneered the application of BNs for qualitative (*i.e.*, hazard)

assessment of chemicals, and we extend it here to qAOPs and risk assessment. The largest constraint for (D)BNs lies in the design of the experiments providing the data. The same doses and observation times should be used as much as possible. Otherwise, statistical imputation has to be used to obtain uniform dose and time schedules across experiments, and the statistical estimation problem may become overwhelming. From an experimental point of view, however, it might not be feasible to observe the different KEs with the same time frame and on the same time scale, even though it might be possible to simplify time dependencies by considering some effects to be instantaneous in comparison to others. Finally, it is possible to couple PK models with DBN models, either by pre-computing the value of the dose nodes in the DBN with a pharmacokinetic model, or by extending the DBN to simulate the pharmacokinetic data available.

Table 4: Summary characteristics of the three quantification methods used.

Method	Principle	Pros	Cons
Dose-response modeling	Find empirical equations that fit the data and, if needed, mathematically invert the models to link KEs.	Mildly simple and fast to obtain. Can describe the data well.	Mathematical sophistication required. Inversion arbitrarily constrains the KERs equations. Complex KERs may be modeled in an overly simple manner given the underlying biology. Parameter values can only be obtained by fitting. Resulting qAOP should not be used outside the time and dose range of the data. Linking with PK is difficult.
Dynamic Bayesian networks	KERs are modeled with simple equations and the whole set of KERs is modeled by a causal network. Time is built in the network structure.	Equations are simple and the network structure is dictated automatically by the AOP. Ability to describe complex behavior comes from structure rather than from complicated KER equations. Can describe the data well, more flexibly than dose-response models. Linking with PK is feasible.	Statistical sophistication needed. Parameter values can only be obtained by fitting. Unbalanced experimental design requires heavy statistical calculations. Resulting qAOP should not be used outside the time and dose range of the data.
Systems biology modeling	A set of differential equations is used to represent the KERs and time evolution of the nodes of the KEs.	Complex KEs and feed-back loops or modifiers can be modeled in detail. Forces mechanistic questioning of the data and allow formal testing of mechanistic hypotheses. Parameter values can be obtained from various sources in addition to fitting. Time, dose, and spatial organization (at the organelle, cell, or tissue level) can be seamlessly integrated. Linking with PK is easy. Resulting qAOP can be used outside the time and dose range of the data, if the structure is trusted.	Mathematical and statistical sophistication needed. Complex to develop and check, many parameters make the approach data-hungry.

The SB model we developed addresses only part of the CKD AOP: It does not include mitochondrial injury. But, for chemicals causing oxidative injury in the renal proximal tubule, it describes the links between GSH, the control of oxidative stress by Nrf2 and the formation of fluorescent cDCF in a very detailed way. The model is complex though, with 57 differential equations and 335 parameters. However, since it has been already parameterized for RPTEC/TERT1 cells, only the five parameters specific to $KBrO_3$ and cDCF reactions needed to be calibrated with the current data. We essentially found that a reasonable fit could be obtained if $KBrO_3$ acts directly on cH2DCF, and that cDCF is transported out of the cells or bleaches significantly with time. We also found that modeling the pre-incubation period gives important information about the cellular background rate of oxidative stress. Such informative modeling is easy to do with a mechanistic model and impossible to do with the previous two approaches. The non-linearity of the effect of $KBrO_3$ on cH2DCF is not well explained by a first-order reaction, but we did not introduce *ad hoc* equations or further hypotheses, because the mismatch already leads to the following point of discussion: According to our SB model, neither action on extra-cellular nor on intra-cellular GSH can entirely explain the cDCF data. This questions the application of the GSH readout as a measure of $KBrO_3$ effect in this AOP. While it is well accepted that thiol depletion can induce oxidative stress, the model suggests that this may not be the main mechanism of action of $KBrO_3$ in the readout test. Thus, $KBrO_3$ may *not* be well suited to quantify our AOP, which also calls into question the results obtained with the other two models. SB models force us to think mechanistically about the data, asking which biochemical reactions could explain them. They can also simulate particular details of the experimental protocols and background cellular processes, improving our understanding of the biology and of the tests themselves. However, we cannot entirely exclude that the model may be misleading, because its many parameters have not all been calibrated perfectly. SB models can also naturally integrate

pharmacokinetic models, since those are based on the same principles and same mathematical objects. Therefore, complicated SB models should be seen as investment for the future rather than as quick answer to urgent questions.

An Effectopedia implementation of both BN and SB models faces challenges, of which the most important is matching the internal structure of the models to the conceptual structure provided by the AOP. Currently, Effectopedia allows “global models” in which one BN or SB model can cover several KEs. Such models need to have specific outputs matching the AOP KEs. A problem in that approach is the derivation of reusable KERs. If the global model contains complex time or variable dependencies between non-adjacent KEs, they need to be explicitly represented in the AOP as feedbacks, feed-forwards or modifying factors. However, extracting such dependencies is non-trivial. Alternatively, the AOP can be re-designed if the global model indicates that some tightly coupled KEs can be merged.

5 Conclusion

The three approaches tested have different advantages. Dose-response based qAOPs may seem the easiest to develop at first sight, but they have very limited extrapolation and explanatory power. Bayesian networks are in fact easier to develop, once the technology is mastered, but they impose either strong constraints on experimental design (fixed dosing and observation schedules) or require complex statistical treatment (imputation). Systems biology models are more complex to develop, but one can strive for parsimony, as when we simplified the gene regulation part of our model. Importantly, they offer insight in the data relevance and biology that the other approaches cannot afford. In any case, the three approaches we presented can all fully propagate uncertainty about qAOP predictions, which is essential for

proper risk assessment. The contrasted results we obtained demonstrate that the choice of approach is not neutral. They also emphasize the importance of data collection on:

- *In vitro* kinetics, to understand and take into account the fate of the chemicals in the test system;
- Baseline behavior of the cells, in the absence of chemical exposure. For that, raw experimental data should be delivered to modelers without pre-processing such as normalization to background values. If such normalization had been applied to our cDCF data, for example, we would have lost important information on background ROS production. Correcting for background may impair essential mechanistic understanding of AOPs, which are as much about the underlying biology as about the effects of stressors;
- Different readouts, to select the most relevant one for the underlying KE or to better understand a complex KE (such as oxidative stress);
- Other chemicals to check whether the parameterized KERs are robust and really chemical-independent.

To avoid pitfalls in qAOP development, we suggest to take at least two approaches in parallel: First, a mechanistic modeling path, able to help test hypotheses, design experiments and deeply understand the results; Second, because we cannot always wait to have a fully mechanistic model developed, a lighter statistical approach. At the moment dose-response based modeling is the simplest, but we hope that we can contribute to a more wide-spread dissemination of dynamic Bayesian networks in this area. In this spirit, one of the goals of the Effectopedia platform is to facilitate the creation of qAOPs by integrating and comparing the results brought by various modeling approaches.

Conflict of Interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Funding

This work was supported by the EU-ToxRisk project (An Integrated European "Flagship" Program Driving Mechanism-Based Toxicity Testing and Risk Assessment for the 21st Century) funded by the European Commission under the Horizon 2020 programme (Grant Agreement No. 681002). AL was also partly funded by the 2015 Long Range Initiative Innovative Science Award of the European Chemical Industry Council.

Disclaimer

The opinions expressed and arguments employed herein are those of the authors and do not necessarily reflect the official views of the OECD or the governments of its member countries.

Data Availability Statement

The data used for this analysis are given in Supplemental material Tables S5, S6, and S7.

Author Contributions

AL and PJ did the *in vitro* work; CT worked on dose-response modeling; WG and GG worked on Bayesian networks; EZ worked on systems biology modeling; HA and MS worked on Effectopedia, FYB instigated the work and coordinated all activities. AL, PJ, CT, WG, GG, EZ, HY, HA, MS, JBB and FYB participated to the writing of the manuscript.

References

[1] Organisation for Economic Co-operation and Development (OECD), Guidance Document for the Use of Adverse Outcome Pathways in Developing Integrated Approaches to Testing and Assessment (IATA). Series on Testing and Assessment, No. 260, ENV/JM/MONO(2016)67, OECD Environment, Health and Safety Publications, Paris, France, 2016. http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=env/jm/mono(2016)67&doclanguage=en.

[2] D.L. Villeneuve, D. Crump, N. Garcia-Reyero, M. Hecker, T.H. Hutchinson, C.A. LaLone, B. Landesmann, T. Lettieri, S. Munn, M. Nepelska, M.A. Ottinger, L. Vergauwen, M. Whelan, Adverse outcome pathway (AOP) development I: strategies and principles, Toxicological Sciences. 142 (2014) 312–320. doi:10.1093/toxsci/kfu199.

[3] C.A. LaLone, G.T. Ankley, S.E. Belanger, M.R. Embry, G. Hodges, D. Knapen, S. Munn, E.J. Perkins, M.A. Rudd, D.L. Villeneuve, M. Whelan, C. Willett, X. Zhang, M. Hecker, Advancing the adverse outcome pathway framework-An international horizon scanning approach, Environmental Toxicology and Chemistry. 36 (2017) 1411–1421. doi:10.1002/etc.3805.

[4] Organisation for Economic Co-operation and Development (OECD), Users' Handbook Supplement to the Guidance Document for Developing and Assessing AOPs. Second Edition. Series on Testing and Assessment, No. 233, ENV/JM/MONO(2016)12, OECD Environment, Health and Safety Publications, Paris, France, 2018. https://one.oecd.org/document/ENV/JM/MONO(2016)12/en/pdf.

[5] G.A. Pavlopoulos, M. Secrier, C.N. Moschopoulos, T.G. Soldatos, S. Kossida, J. Aerts, R. Schneider, P.G. Bagos, Using graph theory to analyze biological networks, BioData Mining. 4 (2011). doi:10.1186/1756-0381-4-10.

[6] M. Vinken, The adverse outcome pathway concept: A pragmatic tool in toxicology, Toxicology. 312 (2013) 158–165. doi:10.1016/j.tox.2013.08.011.

[7] M. Leist, A. Ghallab, R. Graepel, R. Marchan, R. Hassan, S.H. Bennekou, A. Limonciel, M. Vinken, S. Schildknecht, T. Waldmann, E. Danen, B. van Ravenzwaay, H. Kamp, I. Gardner, P. Godoy, F.Y. Bois, A. Braeuning, R. Reif, F. Oesch, D. Drasdo, S. Höhme,

M. Schwarz, T. Hartung, T. Braunbeck, J. Beltman, H. Vrieling, F. Sanz, A. Forsby, D. Gadaleta, C. Fisher, J. Kelm, D. Fluri, G. Ecker, B. Zdrazil, A. Terron, P. Jennings, B. van der Burg, S. Dooley, A.H. Meijer, E. Willighagen, M. Martens, C. Evelo, E. Mombelli, O. Taboureau, A. Mantovani, B. Hardy, B. Koch, S. Escher, C. van Thriel, C. Cadenas, D. Kroese, B. van de Water, J.G. Hengstler, Adverse outcome pathways: opportunities, limitations and open questions, Archives of Toxicology. 91 (2017) 3477–3505. doi:10.1007/s00204-017-2045-3.

[8] M. Vinken, The adverse outcome pathway concept: a pragmatic tool in toxicology, Toxicology. 312 (2013) 158–165. doi:10.1016/j.tox.2013.08.011.

[9] M. Sachana, E. Leinala, Approaching chemical safety assessment through application of integrated approaches to testing and assessment: combining mechanistic information derived from adverse outcome pathways and alternative methods, Applied In Vitro Toxicology. 3 (2017) 227–233. doi:10.1089/aivt.2017.0013.

[10] R.B. Conolly, G.T. Ankley, W. Cheng, M.L. Mayo, D.H. Miller, E.J. Perkins, D.L. Villeneuve, K.H. Watanabe, Quantitative adverse outcome pathways and their application to predictive toxicology, Environmental Science & Technology. 51 (2017) 4661–4672. doi:10.1021/acs.est.6b06230.

[11] I. Hassan, H. El-Masri, P.A. Kosian, J. Ford, S.J. Degitz, M.E. Gilbert, Neurodevelopment and thyroid hormone synthesis inhibition in the rat: quantitative understanding within the adverse outcome pathway framework, Toxicological Sciences. 160 (2017) 57–73. doi:10.1093/toxsci/kfx163.

[12] C.J. Oates, S. Mukherjee, Network inference and biological dynamics, The Annals of Applied Statistics. 6 (2012) 1209–1235. doi:10.1214/11-AOAS532.

[13] J.S. Jaworska, A. Natsch, C. Ryan, J. Strickland, T. Ashikaga, M. Miyazawa, Bayesian integrated testing strategy (ITS) for skin sensitization potency assessment: a decision support system for quantitative weight of evidence and adaptive testing strategy, Archives of Toxicology. 89 (2015) 2355–2383. doi:10.1007/s00204-015-1634-2.

[14] C. Rovida, N. Alépée, A.M. Api, D.A. Basketter, F.Y. Bois, F. Caloni, E. Corsini, M. Daneshian, C. Eskes, J. Ezendam, H. Fuchs, P. Hayden, C. Hegele-Hartung, S. Hoffmann, B. Hubesch, M.N. Jacobs, J. Jaworska, A. Kleensang, N. Kleinstreuer, J. Lalko, R. Landsiedel, F. Lebreux, T. Luechtefeld, M. Locatelli, A. Mehling, A. Natsch, J.W. Pitchford, D. Prater, P. Prieto, A. Schepky, G. Schüürmann, L. Smirnova, C. Toole, E. van Vliet, D. Weisensee, T. Hartung, Integrated testing strategies (ITS) for safety assessment, Alternatives to Laboratory Animals. 32 (2015) 25–40. doi:http://dx.doi.org/10.14573/altex.1411011.

[15] A. Limonciel, L. Aschauer, A. Wilmes, S. Prajczer, M.O. Leonard, W. Pfaller, P. Jennings, Lactate is an ideal non-invasive marker for evaluating temporal alterations in cell stress and toxicity in repeat dose testing regimes, Toxicology in Vitro. 25 (2011) 1855–1862. doi:10.1016/j.tiv.2011.05.018.

[16] J.M. Lever, R. Boddu, J.F. George, A. Agarwal, Heme oxygenase-1 in kidney health and disease, Antioxidants and Redox Signaling. 25 (2016) 165–183. doi:10.1089/ars.2016.6659.

[17] M.P. Chin, G.L. Bakris, G.A. Block, G.M. Chertow, A. Goldsberry, L.A. Inker, H.J.L. Heerspink, M. O'Grady, P.E. Pergola, C. Wanner, D.G. Warnock, C.J. Meyer, Bardoxolone methyl improves kidney function in patients with chronic kidney disease stage 4 and type 2 diabetes: post-hoc analyses from bardoxolone methyl evaluation in patients with chronic kidney disease and type 2 diabetes study, American Journal of Nephrology. 47 (2018) 40–47. doi:10.1159/000486398.

[18] A. Limonciel, A. Wilmes, L. Aschauer, R. Radford, K.M. Bloch, T. McMorrow, W. Pfaller, J.H. van Delft, C. Slattery, M.P. Ryan, E.A. Lock, P. Jennings, Oxidative stress induced by potassium bromate exposure results in altered tight junction protein expression in renal proximal tubule cells, Archives of Toxicology. 86 (2012) 1741–1751. doi:10.1007/s00204-012-0897-0.

[19] M. Wieser, G. Stadler, P. Jennings, B. Streubel, W. Pfaller, P. Ambros, C. Riedl, H. Katinger, J. Grillari, R. Grillari-Voglauer, hTERT alone immortalizes epithelial cells of renal proximal tubules without changing their functional characteristics, American Journal of Physiology-Renal Physiology. 295 (2008) F1365–F1375. doi:10.1152/ajprenal.90405.2008.

[20] K. Yamada, Y. Hiradate, M. Goto, C. Nishiyama, K. Hara, H. Yoshida, K. Tanemura, Potassium bromate disrupts mitochondrial distribution within murine oocytes during in vitro maturation, Reproductive Medicine and Biology. 17 (2018) 143–148. doi:10.1002/rmb2.12079.

[21] J.A. Dykens, Y. Will, The significance of mitochondrial toxicity testing in drug development, Drug Discovery Today. 12 (2007) 777–785. doi:10.1016/j.drudis.2007.07.013.

[22] L. Aschauer, A. Limonciel, A. Wilmes, S. Stanzel, A. Kopp-Schneider, P. Hewitt, A. Lukas, M.O. Leonard, W. Pfaller, P. Jennings, Application of RPTEC/TERT1 cells for investigation of repeat dose nephrotoxicity: A transcriptomic study, Toxicology in Vitro. 30 (2015) 106–116. doi:10.1016/j.tiv.2014.10.005.

[23] P. Jennings, D. Crean, L. Aschauer, A. Limonciel, K. Moenks, G. Kern, P. Hewitt, K. Lhotta, A. Lukas, A. Wilmes, M.O. Leonard, Interleukin-19 as a translational indicator of renal injury, Archives of Toxicology. 89 (2015) 101–106. doi:10.1007/s00204-014-1237-3.

[24] A. Wilmes, A. Limonciel, L. Aschauer, K. Moenks, C. Bielow, M.O. Leonard, J. Hamon, D. Carpi, S. Ruzek, A. Handler, O. Schmal, K. Herrgen, P. Bellwon, C. Burek, G.L. Truisi, P. Hewitt, E. Di Consiglio, E. Testai, B.J. Blaauboer, C. Guillou, C.G. Huber, A. Lukas, W. Pfaller, S.O. Mueller, F.Y. Bois, W. Dekant, P. Jennings, Integrated omic profiling reveals novel insights of cyclosporine A induced cellular stress, Journal of Proteomics. 79 (2013) 180–194. doi:10.1016/j.jprot.2012.11.022.

[25] A. Wilmes, C. Bielow, C. Ranninger, P. Bellwon, L. Aschauer, A. Limonciel, H. Chassaigne, T. Kristl, S. Aiche, C.G. Huber, C. Guillou, P. Hewitt, M.O. Leonard, W. Dekant, F. Bois, P. Jennings, Mechanism of cisplatin proximal tubule toxicity revealed by integrating transcriptomics, proteomics, metabolomics and biokinetics, Toxicology in Vitro. 30 (2015) 117–127. doi:10.1016/j.tiv.2014.10.006.

[26] Organisation for Economic Co-operation and Development (OECD), Guidance Document on Developing and Assessing Adverse Outcome Pathways. Second Edition. Series on Testing and Assessment, No. 184, ENV/JM/MONO(2013)6, OECD Environment, Health and Safety Publications, Paris, France, 2017. http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=env/jm/mono(2013)6&doclanguage=en.

[27] A. Staples, C. Wong, Risk factors for progression of chronic kidney disease, Current Opinion in Pediatrics. 22 (2010) 161–169. doi:10.1097/MOP.0b013e328336ebb0.

[28] Y. Kong, S.E. Trabucco, H. Zhang, Oxidative Stress, Mitochondrial Dysfunction and the Mitochondria Theory of Aging, in: L. Robert, T. Fulop (Eds.), Interdisciplinary Topics in Gerontology, S. Karger AG, Basel, 2014: pp. 86–107. http://www.karger.com?doi=10.1159/000358901.

[29] L. Aschauer, L.N. Gruber, W. Pfaller, A. Limonciel, T.J. Athersuch, R. Cavill, A. Khan, G. Gstraunthaler, J. Grillari, R. Grillari, P. Hewitt, M.O. Leonard, A. Wilmes, P. Jennings, Delineation of the key aspects in the regulation of epithelial monolayer formation, Molecular and Cellular Biology. 33 (2013) 2535–2550. doi:10.1128/MCB.01435-12.

[30] P. Jennings, A. Limonciel, L. Felice, M.O. Leonard, An overview of transcriptional regulation in response to toxicological insult, Archives of Toxicology. 87 (2013) 49–72. doi:10.1007/s00204-012-0919-y.

[31] J. Pearl, Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference, Morgan Kaufmann Publishers, 1988.

[32] J. Jaworska, Y. Dancik, P. Kern, F. Gerberick, A. Natsch, Bayesian integrated testing strategy to assess skin sensitization potency: from theory to practice: Integrated testing strategy to assess skin sensitization potency, Journal of Applied Toxicology. (2013). doi:10.1002/jat.2869.

[33] U.B. Kjærulff, A.L. Madsen, Bayesian Networks and Influence Diagrams, Springer New York, New York, NY, 2008. doi:10.1007/978-0-387-74101-7.

[34] V.I. Pavlovic, Dynamic Bayesian Networks for Information Fusion with Applications to Human-computer Interfaces, PhD Thesis, University of Illinois at Urbana-Champaign, 1999.

[35] S. Geenen, J.W. Yates, J.G. Kenna, F.Y. Bois, I.D. Wilson, H.V. Westerhoff, Multiscale modelling approach combining a kinetic model of glutathione metabolism with PBPK models of paracetamol and the potential glutathione-depletion biomarkers ophthalmic acid and 5-oxoproline in humans and rats, Integrative Biology. 5 (2013) 877–888.

[36] J. Hamon, M. Renner, M. Jamei, A. Lukas, A. Kopp-Schneider, F.Y. Bois, Quantitative in vitro to in vivo extrapolation of tissues toxicity, Toxicology in Vitro. (2015). doi:10.1016/j.tiv.2015.01.011.

[37] C.H. He, P. Gong, B. Hu, D. Stewart, M.E. Choi, A.M. Choi, J. Alam, Identification of activating transcription factor 4 (ATF4) as an Nrf2-interacting protein. Implication for

heme oxygenase-1 gene regulation, The Journal of Biological Chemistry. 276 (2001) 20858–20865. doi:10.1074/jbc.M101198200.

[38] H.-C. Huang, T. Nguyen, C.B. Pickett, Regulation of the antioxidant response element by protein kinase C-mediated phosphorylation of NF-E2-related factor 2, Proceedings of the National Academy of Sciences, USA. 97 (2000) 12475–12480. doi:10.1073/pnas.220418997.

[39] N.C. Andrews, H. Erdjument-Bromage, M.B. Davidson, P. Tempst, S.H. Orkin, Erythroid transcription factor NF-E2 is a haematopoietic-specific basic-leucine zipper protein, Nature. 362 (1993) 722–728. doi:10.1038/362722a0.

[40] J. Hamon, P. Jennings, F.Y. Bois, Systems biology modeling of omics data: effect of cyclosporine a on the Nrf2 pathway in human renal cells, BMC Systems Biology. 8 (2014) 76. doi:10.1186/1752-0509-8-76.

[41] S. Geenen, F.B. du Preez, M. Reed, H. Frederik Nijhout, J. Gerry Kenna, I.D. Wilson, H.V. Westerhoff, J. Snoep, A mathematical modelling approach to assessing the reliability of biomarkers of glutathione metabolism, European Journal of Pharmaceutical Sciences. 46 (2012) 233–243. doi:10.1016/j.ejps.2011.08.017.

[42] B. Laupeze, L. Amiot, A. Courtois, L. Vernhet, B. Drenou, R. Fauchet, O. Fardel, Use of the anionic dye carboxy-2',7'-dichlorofluorescein for sensitive flow cytometric detection of multidrug resistance-associated protein activity, International Journal of Oncology. 15 (1999) 571–576.

[43] C. Drennen, E. Gorse, R.E. Stratford, Cellular pharmacokinetic model-based analysis of genistein, glyceollin, and MK-571 effects on 5 (and 6)-carboxy-2′,7′-dichloroflourescein disposition in Caco-2 cells, Journal of Pharmaceutical Sciences. 107 (2018) 1194–1203. doi:10.1016/j.xphs.2017.12.004.

[44] P. Bernillon, F.Y. Bois, Statistical issues in toxicokinetic modeling: a Bayesian perspective, Environmental Health Perspectives. 108 (suppl. 5) (2000) 883–893.

[45] F.Y. Bois, Bayesian inference, in: B. Reisfeld, A.N. Mayeno (Eds.), Computational Toxicology Vol. II, Humana Press, New-York, 2012: pp. 597–636.

[46] M. Girolami, B. Calderhead, Riemann manifold Langevin and Hamiltonian Monte Carlo methods, Journal of the Royal Statistical Society Series B. 73 (2011) 123–214. doi:10.1111/j.1467-9868.2010.00765.x.

[47] A. Gelman, D.B. Rubin, Inference from iterative simulation using multiple sequences (with discussion), Statistical Science. 7 (1992) 457–511.

[48] F.Y. Bois, GNU MCSim: Bayesian statistical inference for SBML-coded systems biology models, Bioinformatics. 25 (2009) 1453–1454. doi:10.1093/bioinformatics/btp162.

[49] B. Carpenter, A. Gelman, M.D. Hoffman, D. Lee, B. Goodrich, M. Betancourt, M. Brubaker, J. Guo, P. Li, A. Riddell, *Stan*: a probabilistic programming language, Journal of Statistical Software. 76 (2017). doi:10.18637/jss.v076.i01.

[50] A. Gelman, J.B. Carlin, H.S. Stern, D.B. Rubin, Bayesian Data Analysis, 2nd Edition, Chapman & Hall, London, 2004.

[51] R Development Core Team, R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria, 2013. http://www.R-project.org.

[52] Organisation for Economic Co-operation and Development (OECD), Effectopedia - The Online Encyclopedia of Adverse Outcome Pathways, (2016). https://www.effectopedia.org/ (accessed September 27, 2018).

[53] C. Fisher, S. Siméon, M. Jamei, I. Gardner, F.Y. Bois, VIVD: virtual *in vitro* distribution model for the mechanistic prediction of intracellular concentrations of chemicals in *in vitro* toxicity assays, Toxicology in Vitro. (in press).

[54] J. Jaworska, S. Gabbert, T. Aldenberg, Towards optimization of chemical testing under REACH: a Bayesian network approach to Integrated Testing Strategies, Regulatory Toxicology and Pharmacology. 57 (2010) 157–167. doi:10.1016/j.yrtph.2010.02.003.

Figure captions

Figure 1: Chronic kidney disease AOP diagram. KEs are in green boxes, readout are in yellow boxes, KERs are represented by arrows.

Figure 2: Structure of the dynamic Bayesian network qAOP for chronic kidney disease. $KBrO_3$ concentration and the GSH readout do not vary with time, while the cDCF and lactate readouts were observed at different time intervals. The arrows indicate probabilistic dependencies.

Figure 3: Schematic overview of the assembled SB model. This model covers both transcriptional and biochemical aspects of GSH synthesis and metabolism and its control by the *Nrf2-Keap1* signaling pathway. The blue compartment is cytosol and the red one is nucleus. Blue arrows show reactant(s):product(s) exchange during biochemical or transport reactions, and red arrows indicate enzymatic catalysis (diamond heads) or gene transcription (round heads). In the nucleus, red boxes represent genes and arrows indicate gene activation. Names of genes are in orange, of *mRNA* are in green, of enzymes are in purple, of other proteins and metabolites in blue and of extracellular constants in yellow.

Figure 4: Potassium bromate ($KBrO_3$) and 6-carboxy-2',7'-dichlorofluorescein (cDCF) specific reactions of the SB model. Other abbreviations: extGSH is extra-cellular glutathione; cytGSH: cytosolic glutathione; extGSSG: extra-cellular oxidized glutathione; cytGSSG: cytosolic oxidized glutathione; ROS: reactive oxygen species; cH2CFD: 6-carboxy-2',7'-dichlorodihydrofluorescein. Reactions are represented by red circles: *a* is the oxidation of extGSH by $KBrO_3$; *b*: oxidation of cH2CFD by ROS; *c*: cDCF efflux or bleaching; *d*: oxidation of cH2DCF by $KbrO_3$; *e*: oxidation of cytGSH by $KBrO_3$.

Figure 5: Fit of the $KBrO_3$ - GSH data (circles; each color represents one of the replicates) using the three qAOP models developed. The black line corresponds to the empirical model (Eq. 6). The best fit (solid line) is shown along with 20 additional random fits (gray), showing the uncertainty of the model predictions. The black dashed line represents the best fit obtained with the DBN qAOP. The red line shows the best fit for the SB model.

Figure 6: Fit (top row) and predictions (bottom row) of the dose-response based qAOP for the cDCF (measured in RFU) (left) and lactate (right) readouts. The best fit surfaces (gray) are plotted along with all individual data (colored dots). The predicted chemical-independent relationships (in red) for GSH - time - cDCF, or GSH - time -lactate were obtained by inversion of the qAOP equations (see text). The maximum posterior parameter values given in Table 1 were used to draw the figures.

Figure 7: Fit (top row) and predictions (bottom row) of the dynamic BN qAOP for the cDCF (measured in RFU) (left) and lactate (right) readouts. The best fit surfaces (gray) are plotted along with the data mean (black dots) and all individual data (colored dots). The predicted chemical-independent relationships (in red) are shown for GSH - time -cDCF and GSH - time - lactate. The maximum posterior parameter values given in Table 2 were used to draw the figures.

Figure 8: Best fits of SB model (gray surfaces) to the cDCF RFU data (colored dots), for different levels of model complexity: (A) action of $KBrO_3$ on external GSH and formation of cDCF by ROS; (B) same as A but with the addition of cDCF efflux or bleaching; (C) same as B but with the addition of a direct formation of cDCF by $KBrO_3$; (D) same as B, but with the addition of an action of $KBrO_3$ on internal GSH.