Journal of Biomedical Informatics 145 (2023) 104465

Contents lists available at ScienceDirect

Journal of Biomedical Informatics

journal homepage: www.elsevier.com/locate/yjbin

Original Research

Optimization of an adverse outcome pathway network on chemical-induced cholestasis using an artificial intelligence-assisted data collection and confidence level quantification approach

Jonas van Ertvelde [a], Anouk Verhoeven [a], Amy Maerten [a], Axelle Cooreman [a], Bruna dos Santos Rodrigues [a], Julen Sanz-Serrano [a], Milos Mihajlovic [a], Ignacio Tripodi [b], Marc Teunis [c], Ramiro Jover [d], Thomas Luechtefeld [e], Tamara Vanhaecke [a], Jian Jiang [a,1], Mathieu Vinken [a,1,*]

[a] Entity of In Vitro Toxicology and Dermato-Cosmetology, Department of Pharmaceutical and Pharmacological Sciences, Vrije Universiteit Brussel, Brussels, Belgium
[b] Sciome LLC, Durham, NC, USA
[c] Innovative Testing in Life Sciences and Chemistry, University of Applied Sciences Utrecht, Utrecht, The Netherlands
[d] Joint Research Unit in Experimental Hepatology, University of Valencia, Health Research Institute Hospital La Fe & CIBER of Hepatic and Digestive Diseases, Spain
[e] Toxtrack LLC, Baltimore, MD, USA

ARTICLE INFO

Keywords:
Cholestasis
Adverse outcome pathway
AOP network
Mechanistic toxicology
Shiny application

ABSTRACT

Background: Adverse outcome pathway (AOP) networks are versatile tools in toxicology and risk assessment that capture and visualize mechanisms driving toxicity originating from various data sources. They share a common structure consisting of a set of molecular initiating events and key events, connected by key event relationships, leading to the actual adverse outcome. AOP networks are to be considered living documents that should be frequently updated by feeding in new data. Such iterative optimization exercises are typically done manually, which not only is a time-consuming effort, but also bears the risk of overlooking critical data. The present study introduces a novel approach for AOP network optimization of a previously published AOP network on chemical-induced cholestasis using artificial intelligence to facilitate automated data collection followed by subsequent quantitative confidence assessment of molecular initiating events, key events, and key event relationships.
Methods: Artificial intelligence-assisted data collection was performed by means of the free web platform Sysrev. Confidence levels of the tailored Bradford-Hill criteria were quantified for the purpose of weight-of-evidence assessment of the optimized AOP network. Scores were calculated for biological plausibility, empirical evidence, and essentiality, and were integrated into a total key event relationship confidence value. The optimized AOP network was visualized using Cytoscape with the node size representing the incidence of the key event and the edge size indicating the total confidence in the key event relationship.
Results: This resulted in the identification of 38 and 135 unique key events and key event relationships, respectively. Transporter changes was the key event with the highest incidence, and formed the most confident key event relationship with the adverse outcome, cholestasis. Other important key events present in the AOP network include: nuclear receptor changes, intracellular bile acid accumulation, bile acid synthesis changes, oxidative stress, inflammation and apoptosis.

Abbreviations: AO, adverse outcome; AOP, adverse outcome pathway; BP_{KER}, biological plausibility score; BSEP, bile salt export pump; CAR, constitutive androstane receptor; DILI, drug-induced liver injury; EE_{KER}, empirical evidence score; ESS_{KER}, essentiality score; FXR, farnesoid X receptor; KE, key event; KEM, key event marker; KER, key event relationship; MDR, multidrug resistance protein; MRP, multidrug resistance-associated protein; NAM, new approach methodology; NTCP, sodium-taurocholate co-transporting polypeptide; OATP, organic anion transporting peptide; OECD, organization for economic co-operation and development; PXR, pregnane X receptor; ROCK, rho-associated protein kinase; ROS, reactive oxygen species; SHP, small heterodimer partner; TOT_{KER}, total score; WoE, weight-of-evidence.

* Corresponding author at: Entity of In Vitro Toxicology and Dermato-Cosmetology, Department of Pharmaceutical and Pharmacological Sciences, Vrije Universiteit Brussel, Laarbeeklaan 103, 1090 Brussels, Belgium. Tel.: +32-24774587. website: mathieuvinken.com.
E-mail address: mathieu.vinken@vub.be (M. Vinken).
[1] Share equal seniorship.

https://doi.org/10.1016/j.jbi.2023.104465
Received 7 June 2023; Received in revised form 19 July 2023; Accepted 31 July 2023
Available online 2 August 2023

Conclusions: This process led to the creation of an extensively informative AOP network focused on chemical-induced cholestasis. This optimized AOP network may serve as a mechanistic compass for the development of a battery of *in vitro* assays to reliably predict chemical-induced cholestatic injury.

1. Introduction

Statement of Significance:

Problem or Issue	The existing AOP network on chemical-induced cholestasis is outdated.
What is Already Known	Current AOP development strategies rely on manual processes, which not only is a time-consuming effort, but also bears the risk of overlooking critical data.
What this Paper Adds	The present study introduces a novel approach for AOP network optimization of a previously published AOP network on chemical-induced cholestasis using artificial intelligence to facilitate automated data collection followed by subsequent quantitative confidence assessment of molecular initiating events, key events, and key event relationships.

Cholestasis is a pathological condition that denotes any situation of impaired bile formation, excretion, or secretion with concomitant accumulation of noxious bile acids in the liver or in the blood circulation [1]. A distinction can be made between intrahepatic and extrahepatic cholestasis depending on the location of bile flow disruption. The former refers to a functional defect in bile formation in the liver, while the latter is caused by an anatomical blockage outside the liver, typically in the bile ducts [2]. Cholestatic liver injury can be induced by a plethora of factors, including chemical compounds. Chemical-induced cholestasis is mainly of intrahepatic nature and represents a subgroup of drug-induced liver injury (DILI). In fact, cholestatic liver injury underlies 20–40 % of all DILI cases and is responsible for 29 % of all drug withdrawals during premarketing and postmarketing phases of drug development [3,4]. Cessation of drug administration usually resolves DILI. Nevertheless, cholestatic DILI may cause permanent liver damage and can even be lethal [5].

Safety testing of chemicals, including pharmaceuticals, still predominantly relies on whole-animal studies focusing on apical endpoints of toxicity. These approaches not only raise ethical concerns, but also fail to generate mechanistic knowledge critical for understanding the development of adverse health effects, which results in poor human predictivity [6]. In this respect, the beginning of the 21st century has witnessed a paradigm shift in toxicity testing from animal-based approaches towards next generation risk assessment by applying new approach methodologies (NAMs) [7]. NAMs refer to non-animal approaches, including *in chemico*, *in vitro* and/or *in silico* methods. Adverse outcome pathways (AOPs) have been introduced in this context to provide a solid mechanistic basis for such NAMs [8]. AOPs allow to structure and visualize the mechanisms of toxicity starting from a molecular initiating event (MIE) towards an adverse outcome (AO) through several key events (KEs), including their links or so-called KE relationships (KERs), at a biological level relevant to risk assessment [9]. Individual AOPs sharing one or more MIE or KE are frequently merged in an AOP network, which better reflects the real-life complexity of toxicity [10,11].

An AOP network describing chemical-induced cholestasis was introduced a decade ago [12]. This is still the only fully mature AOP network in its kind included in the AOP Wiki maintained by the Organization for Economic Co-operation and Development (OECD). This AOP network was assessed according to the tailored Bradford-Hill criteria that evaluate the biological plausibility and empirical evidence of the KERs, the essentiality of the KEs, as well as the overall confidence in the AOP network [13]. As such, 3 types of MIEs are distinguished in this AOP network, namely transporter changes, hepatocellular changes, and bile canalicular changes. The former revolves around functional inhibition or reduced expression of proteins that mediate the transport of bile acids and drugs in hepatocytes [14]. Hepatocellular changes comprise alterations in the cytoskeletal architecture of hepatocytes as well as in tight junction integrity, which ultimately affect bile flow [15]. Bile canalicular changes include dilatation and constriction of bile canaliculi that may also disrupt bile flow [16]. Collectively, transporter changes, hepatocellular changes, and bile canalicular changes result in the onset of an adverse response and an adaptive response. The accumulation of bile acids activates an adverse response accompanied by oxidative stress, mitochondrial impairment, inflammation, endoplasmic reticulum stress and cell death [2]. The adaptive response relies on the activation of specific nuclear receptors and intends to counteract the adverse response by removing bile acids through the alteration of bile acid synthesis and transportation [17,18].

A number of efforts have been undertaken over the past few years to scientifically validate and optimize the AOP network on chemical-induced cholestasis [19]. However, this AOP network still relies on manually extracted data from scientific literature, and therefore may be prone to data gaps, in particular holding for more recently described mechanisms of chemical-induced cholestasis [2]. The present study was set up to revisit and optimize the AOP network on chemical-induced cholestasis through artificial intelligence-assisted collection of relevant data available in scientific literature. The revised AOP network is subsequently thoroughly evaluated by application of the tailored Bradford-Hill criteria in compliance with OECD guidelines. As an unprecedented feature in the AOP field, the outcome of the assessment is visually implemented in the AOP network. To leverage the use of the optimized AOP network, a Shiny application, an R package to build an interactive web application straight from R, was developed [20]. This tool allows for interaction with the complete dataset in a way that resembles an AOP network. It provides options to modify the AOP network's appearance by selecting specific parameters such as chemical type or organism.

2. Methods and materials

The workflow for optimizing and assessing the AOP network on chemical-induced cholestasis consisted of 3 steps, namely (*i*) data collection and extraction, (*ii*) AOP network optimization and weight-of-evidence (WoE) scoring assessment, and (*iii*) AOP network visualization (Fig. 1).

2.1. Data collection and extraction

Data required for the optimization of the AOP network were collected from scientific papers listed in PubMed. For practical reasons, only papers published after the introduction of the initial AOP network of chemical-induced cholestasis [12] were selected for data collection. Generic terms related to cholestasis, such as liver injury, but also more specific KEs, including bile acid accumulation, relevant to chemical-induced cholestasis, were used as keywords in the "PubMed Advanced Search Builder" for the collection of relevant papers (Supplementary information, table s1). This resulted in a total of 6572 papers. Sysrev, a free web platform for data curation and systematic evidence review, was used for data collection and extraction. Sysrev combines human and machine learning algorithms for an efficient and standardized workflow in the collection and extraction of data [21]. The general Sysrev workflow consists of data source creation, label definition, user/machine recruitment/review and data export. For the present study, the abstracts of the 6572 selected papers were first loaded up into Sysrev as the data source. A simple Boolean data type ("yes/no") label was defined to

include ("yes") or exclude the paper ("no"). Three experts were assigned for the first abstract screening phase to manually screen 10 % of the article abstracts (*i.e.*, 658) in order to train the machine learning model for further automated screening. To prevent the inclusion or exclusion of false positive and false negative papers, respectively, a prediction inclusion cut-off was set at 60 %, implying that the model had to be 60 % accurate when including a paper. This resulted in 544 papers that were found eligible for the second abstract screening phase, where papers were manually screened for excluding false positive papers and review articles. This finally resulted in 148 papers that were fed into Sysrev as the new data source. A labelling system was created for the identification of MIEs, KEs and KERs. KEs identified from previous versions of the AOP network on chemical-induced cholestasis [12,19] were added as categorical values, while novel KEs were assigned as such and identified as string labels. Furthermore, labels relating to information necessary for AOP WoE assessment were introduced. Extracted data were finally exported as.csv files.

2.2. AOP network optimization and WoE scoring assessment

The labelling system was developed to standardize the depiction of KEs and KERs. MIEs and the AO are regarded as distinct types of KEs. To avoid any confusion in labelling the KEs within the AOP network, the labelling system classified MIEs, KEs, and the AO (*i.e.*, cholestasis) as KEs. Additionally, evidence supporting the incidence of identified KEs was included as KE markers (KEM). A KEM is a biomolecule or other measurable parameter used to identify a KE in a biological process, *in casu* the AOP. When multiple KEs were found in the same test system within a single paper, KEs were either assigned as upstream KE or downstream KE, both as part of a KER. Data supporting the assignment of the upstream KE and downstream KE, and their KERs, were depicted by multiple labels present in the labelling system. These data were eventually used for the WoE assessment based on the tailored Bradford-Hill criteria, representing the confidence of KEs and their KERs in the AOP network. A novel approach was introduced based on the quantification of the WoE assessment by assigning numerical values to labels utilized in the labelling system. These values were then utilized in custom-designed equations generating scores for quantifying the tailored Bradford-Hill criteria related to biological plausibility, empirical evidence, and essentiality. These 3 scores were combined in a total score representing the relative confidence of the KER in the AOP network. Data were analyzed using Python 3.6.

2.2.1. Biological plausibility score (BP_{KER})

Assessment of the KERs based on biological plausibility relied on the fundamental understanding of the biological processes and sought for

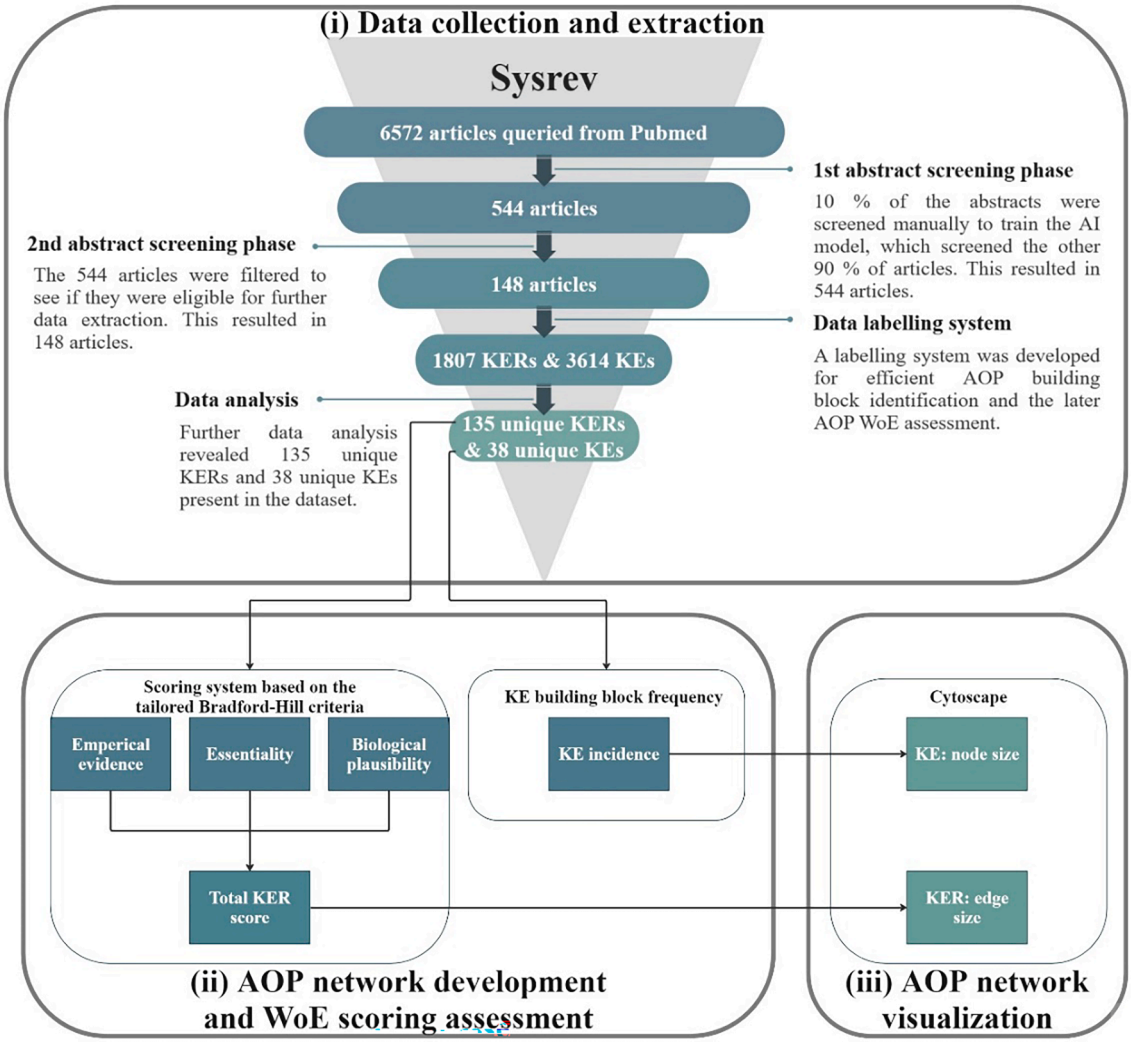

Fig. 1. Workflow for optimizing and assessing the AOP network on chemical-induced cholestasis. The workflow consisted of 3 steps, namely, (i) data collection and extraction, (ii) AOP network optimization and WoE scoring assessment, and (iii) AOP network visualization. (AOP, adverse outcome pathway; KE, key event; KER, key event relationship; WoE, weight of evidence).

any mechanistic relationship between the upstream KE and downstream KE. To ensure a clear understanding of the KEs and their underlying mechanisms, 9 experts in the field of cholestasis were assigned as Sysrev reviewers. The next step was to identify relevant KEMs that could be used to measure these events, and that have been shown to be reliable and valid in previous studies. The consistency value of the KER (C_{KER}), which represents the consistency of the correlation between KEM measurements of a KER, was calculated. The proportion of consistent measurements was multiplied by the number (#) of KERs in unique papers (#KER(unique)). High consistency of the KER indicates that KEMs show a consistent correlation (*i.e.,* positive correlation and negative correlation) between each other resulting in a high C_{KER}. In contrast, a low C_{KER} would result when mixed correlations or no correlation between KEMs of a KER were found. The biological applicability domain was integrated into the biological plausibility score. This allowed to determine if the mechanism was applicable to multiple stressors (maximum level = 4) across different taxa (maximum level = 4), test systems (maximum level = 3), and by also considering the biological level (maximum level = 4) on which the data have been generated (Table 1). The number of levels for each category of biological domain of applicability was then added to the C_{KER}, resulting in the biological plausibility score (BP_{KER}).

$$C_{KER} = \left|\frac{\#\text{corr(pos)}_{KER} - \#\text{corr(neg)}_{KER}}{\#\text{corr(pos)}_{KER} + \#\text{corr(neg)}_{KER} + \#\text{corr(no)}_{KER}}\right| \#KER(unique)$$

$$BP_{KER} = C_{KER} + \text{n(stressor)}_{KER} + \text{n(evidence)}_{KER} + \text{n(taxa)}_{KER} + \text{n(test system)}_{KER}$$

2.2.2. Empirical evidence score (EE_{KER})

Empirical evidence supporting a KER was based on toxicological data if a change in the upstream KE would lead to or is associated with the downstream KE. This assessment was done by citing evidence that shows the dose and time concordance surrounding the KER. Dose concordance was established when the upstream KE was impacted at doses that were generally equal to or lower than those affecting the downstream KE. Time concordance indicated if the upstream KE was observed at an earlier time point than the downstream KE. The labels used for assessing the empirical evidence were designed to confirm the dose-dependence and time-dependence of the upstream KE leading to the downstream KE. The number of dose-concordance and time-concordance positive data ("yes") was subtracted by the dose-concordance and time-concordance negative data ("no" or "not measured") and divided by the total amount of outcomes yielding a percentage of uncertainty. Subsequently, time-concordance (TC_{KER}) and dose-concordance (DC_{KER}) scores were calculated by multiplying the percentage of uncertainty by the number of unique KERs (#KER(unique)). TC_{KER} and DC_{KER} scores were combined to form the overall empirical evidence score of the KER (EE_{KER}). When no concordance data was found or the number of positive concordance data was equal to the amount of negative concordance data for the KER, thus displaying uncertainty in time and dose concordance of this KER, EE_{KER} was set to zero.

$$TC_{KER} = \frac{\#TC(\text{"yes"}) - \#TC(\text{"no"/"NA"})}{\#TC(\text{"yes"}) + \#TC(\text{"no"/"NA"})} \bullet \#KER(unique)$$

$$EE_{KER} = TC_{KER} + DC_{KER}$$

2.2.3. Essentiality score (Ess_{KER})

The essentiality of a KE was assessed by validating the impact of its modulation on downstream KEs within the AOP network. Experimental data that provided evidence if a downstream KE was modulated or prevented when an upstream KE was altered or blocked was utilized for calculating the essentiality score (Ess_{KER}). This indicated if the upstream KE was essential for the downstream KE to occur, representing the strength of the biological linkage between each other. The Ess_{KER} values the evidence level ("direct evidence" or "indirect evidence") and links them to the confidence uncertainty factor found for the mechanism of the KER. The term "direct evidence" was assigned to data sources utilizing knock-out models or inhibitor/inducer studies, while "indirect evidence" was applied to other types of methods. A scoring system was utilized where correlation data with direct evidence was given a score of "+2" (Dir_{pos}) for positive and "-2" (Dir_{neg}) for negative correlations, whereas data with only indirect evidence was given a score of "+1" for positive and "-1" for negative correlations. When no evidence level was found or there was no certainty for the mechanism of the KER, the essentiality score was set to zero.

Table 1
Labelling system used for data depiction in Sysrev. Labels were used to identify the upstream and downstream KEs, with their accompanying KEMs. Labels for the WoE scoring assessment were used for the different tailored Bradford Hill criteria. Labels contained different levels of information. (AOP, adverse outcome pathway; KE, key event; KEM, key event marker; OECD, Organisation for Economic Co-operation and Development; WoE, weight-of-evidence).

AOP KE identification			
Upstream KE **AOP WoE assessment labels**	**Upstream KEM**	**Downstream KE**	**Downstream KEM**
Tailored Bradford Hill criteria	**Contained label**	**Level**	
Biological plausibility and domain of applicability	Stressor	Pharmaceutical/biocide/food constituent/other	
	Evidence level	functional/transcriptional/translational/other	
	Correlation between the KEs	Positive/negative	
	Species	Human/mouse/rat/large mammals	
	Test system	*In vivo/in vitro/ex vivo*	
Essentiality	Correlation between the KEs	Positive/negative/no correlation	
	Assay/method	Translational study/knock-out study/transcriptional study/recovery study/inhibitor/inducer/other	
Empirical evidence	Time concordance	No/yes/not measured	
	Dose concordance	No/yes/not measured	

$$Ess_{KER} = |(2) \bullet \%(Dir_{pos}) + (-2) \bullet \%(Dir_{neg}) + \%(Indir_{pos}) + (-1) \bullet \%(Indir_{neg})| \bullet \#KER(unique)$$

2.2.4. Total score (TOT_{KER})

The scores were normalized to demonstrate equal contributions from the 3 scores to the comprehensive evaluation of the KER. TOT_{KER} was calculated as the sum of the normalized scores of BP_{KER} (nBP_{KER}), EE_{KER} (nEE_{KER}) and Ess_{KER} ($nEss_{KER}$), representing the total confidence in the KER.

$$Tot_{KER} = nBP_{KER} + nEE_{KER} + nEss_{KER}$$

2.3. AOP network visualization

The AOP network was graphically designed in Cytoscape. MIEs, KEs and the AO are presented in blue, yellow and red, respectively. The size of the KE nodes and KER edge is proportional to the incidence of the KE in the dataset and the "total score" of TOT_{KER}, respectively. The "xploreaop" web application was developed using the Shiny package [20] within the R Language and environment for statistical computing [22]. Network visualizations in the app were created using the R packages network3D, visNetwork, tidygraph, ggraph and igraph[23–27]. The tidyverse suite of R packages [28] was utilized for data wrangling, loading, and functional programming tasks. The complete source code, additional information, reuse or extension of the web application and a complete list of dependencies can be found at: https://github.com/onto x-project/xploreaop. The application is currently hosted on a Posit Connect server (https://posit.co/products/enterprise/connect/) at: htt ps://rstudio-connect.hu.nl/xploreaop/.

3. Results

3.1. AOP network data description

The approach used for optimizing the AOP network on chemical-induced cholestasis identified a total of 3614 KEs and 1807 KERs from the data source. Of those, 38 KEs and 135 KERs were found to be unique (Supplementary information, table s2). The labelling system applied in the WoE scoring assessment enabled the collection of information regarding the nature of the stressors, organisms, test systems and assays used for measuring the KEs and KERs. This provided insight into the type of data used for the optimization of the AOP network (Fig. 2). Data used for optimizing the AOP network primarily originated from *in vivo* rodent models, while *in vitro* models were predominantly of human origin (Fig. 2a,b). KEs and KERs were mainly triggered by pharmaceuticals, followed by biocides and food additives (Fig. 2c). KEs present in the data source were identified by corresponding KEMs of 4 types, namely transcriptional, clinical, translational and functional markers (Fig. 2d). A total of 1237 KERs were identified through transcriptional level assays, 670 KERs through clinical KE marker assays, and 648 KERs were identified through translational assays. Functional assays were used to measure KE function or activity and helped to identify 622 KERs.

3.2. AOP network optimization and WoE scoring assessment

The incidence of KEs and confidence in KERs were visualized as node and edge sizes, respectively (Fig. 3). Transporter changes formed the KE with the highest incidence (*i.e.*, 1068 KEs). Transporter changes as part of the AOP network on chemical-induced cholestasis relate to changes in activity and/or expression levels and/or subcellular localization, which may both lead to cholestasis [29]. The mostly reported transporter as a KEM was the bile salt export pump (BSEP), which encompassed 22 % of the KE transporter changes. The primary function of BSEP is to facilitate enterohepatic circulation by conveying bile salts from hepatocytes to the bile [30]. Severe forms of cholestasis, which involve the accumulation of bile acids in the liver, have been associated with the impairment of BSEP function [31]. As a result, BSEP is extensively used as a marker in the assessment of cholestatic risk inflicted by chemicals [32,33]. However,

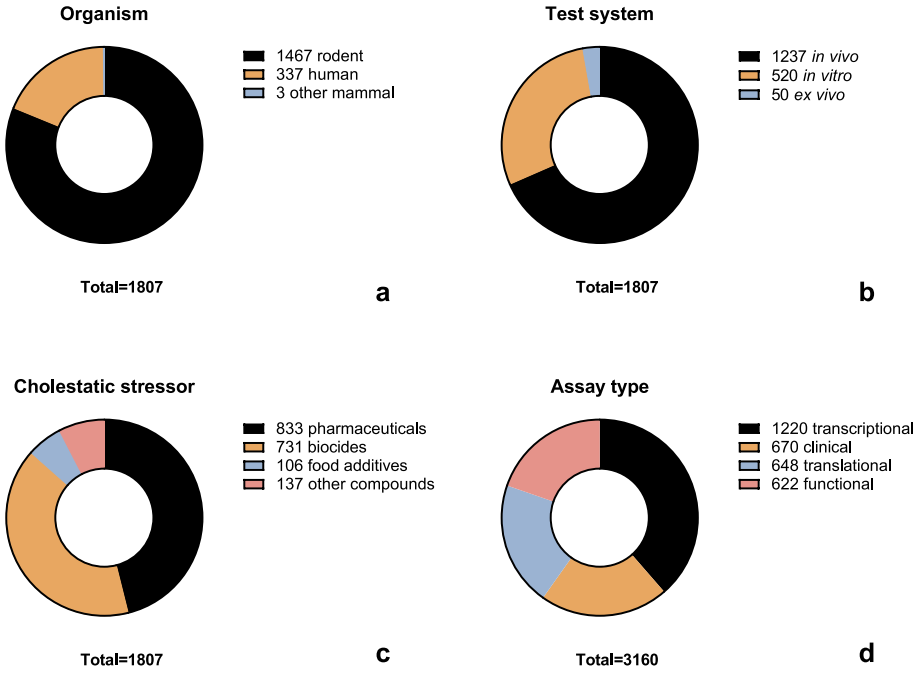

Fig. 2. AOP network data description. (a) Type and number of organisms found for each KER. (b) Type and number of test systems found for each KER. (c) Type and number of cholestatic stressors found for each KER. (d) Type and number of assays found for KEs. (KE, key event; KER, key event relationship).

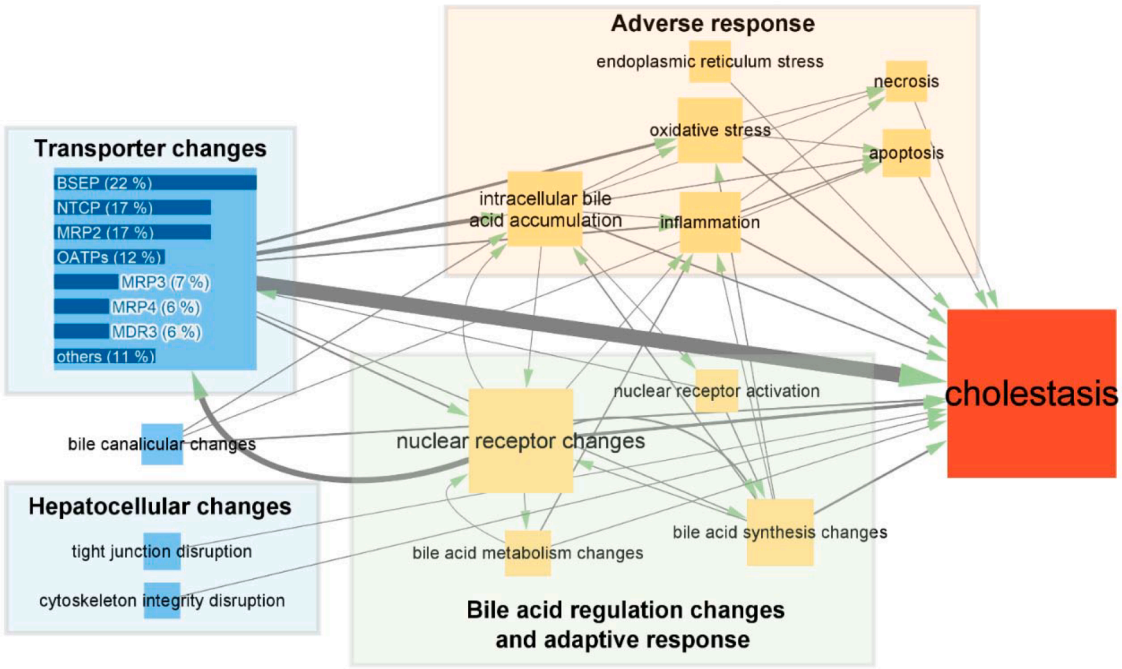

Fig. 3. Optimized and updated AOP network on chemical-induced cholestasis. The node size of the KE is proportional to the incidence of the KE in the dataset. The edge size of the KER is proportional to the total WoE score calculated for the KER (TOT_{KER}). MIEs are presented in blue, KEs in yellow and the AO in red. KEs can be subdivided into adverse response KEs (orange square) and adaptive response KEs (green square). (AO, adverse outcome; BSEP, bile salt export pump; KE, key event; KER, key event relationship; MDR3, multidrug resistance protein 3; MIE, molecular initiating event; MRP2/3/4, multidrug resistance-associated protein 2/3/4; NTCP, sodium-taurocholate co-transporting polypeptide; OATPs, organic anion transporting peptides; TOT_{KER}, total score). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

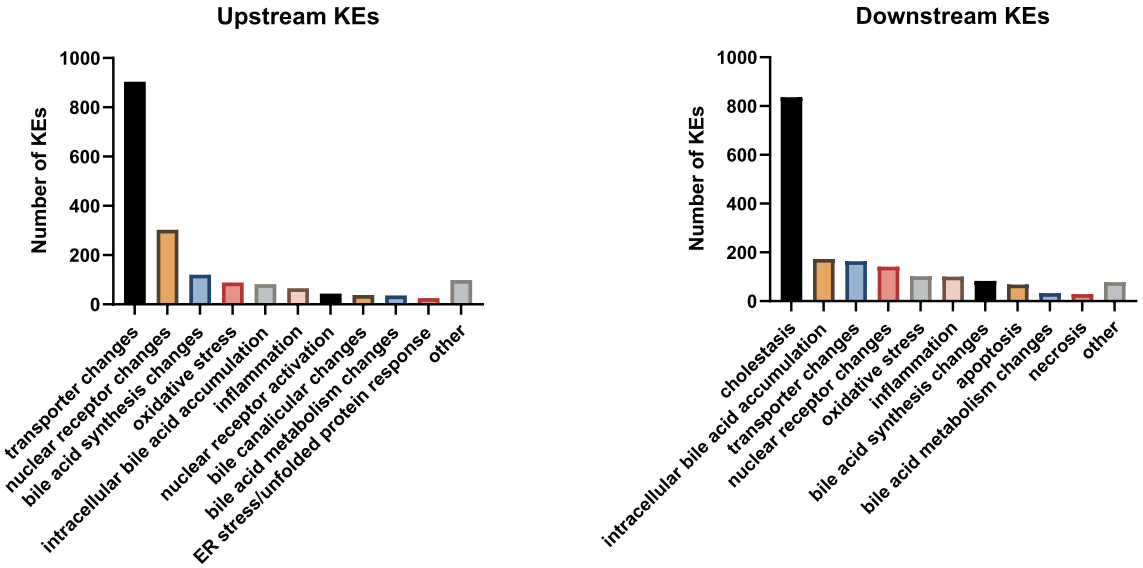

Fig. 4. Incidence of upstream and downstream KEs. The left and right bar graph show the incidence of upstream and downstream KEs, respectively. (ER, endoplasmic reticulum stress; KE, key event).

changes in activity and/or expression of other hepatobiliary transporters are equally associated with chemical-induced cholestasis. Transporters, such as sodium-taurocholate co-transporting polypeptide (NTCP), multidrug resistance-associated proteins (MRP2, MRP3 and MRP4), multidrug resistance protein 3 (MDR3), and organic anion transporting peptides (OATPs), play essential roles in the disposition of bile constituents. Numerous chemicals that cause cholestasis by inhibiting BSEP also negatively affect other hepatobiliary transporters [34–37]. The transporter changes KE incidence number is followed by the AO, namely cholestasis (*i.e.*, 853 KEs) (Fig. 4). Transporter changes as the upstream KE and cholestasis as the downstream KE were found to score the highest on all 3 tailored Bradford-Hill criteria. Accordingly, the relationship between transporter changes and cholestasis shows the highest relative confidence of all identified KERs (Fig. 5). Because the transporter changes KE was found to show the highest upstream incidence in a KER by a considerable margin (Fig. 4), it is depicted in the AOP network as one of the most important MIEs. Other KEs, which were not assigned as downstream KEs, are also shown as MIEs, and include bile canalicular

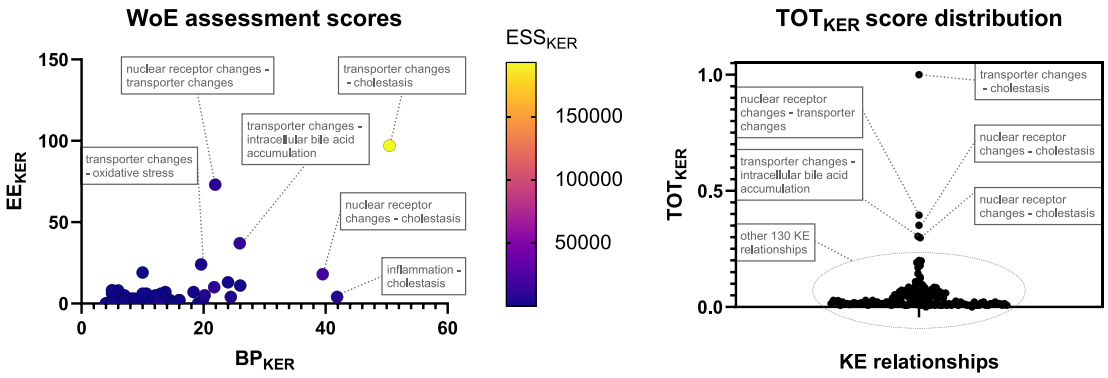

Fig. 5. WoE assessment scores and TOT_{KER} score distribution of the KE relationships. All 3 WoE assessment scores of the tailored Bradford-Hill criteria are presented, with the y-axis presenting the EE_{KER} value, the x-axis BP_{KER} value and the ESS_{KER} value visualized as a color gradient. The distribution of TOT_{KER} values is visualized as a scatterplot. (BP_{KER}, biological plausibility score; EE_{KER}, empirical evidence score; ESS_{KER}, essentiality score; KE, key event; TOT_{KER}, total score; WoE, weight-of-evidence).

and hepatocellular changes. Hepatocellular changes found in the AOP network include tight junction disruption and cytoskeleton disruption. In fact, disruption of the cytoskeletal architecture may result in the loss of hepatocellular polarity and is linked with cholestasis [15]. The AOP network illustrates that both KEs exhibit KERs with the AO, cholestasis (Fig. 3). Bile canalicular changes indicate dilatations and constriction of the bile canaliculi and may be due to interference with the rho-associated protein kinase (ROCK)/myosin light chain kinase/myosin pathway [16]. In the AOP network (Fig. 3), the bile canalicular changes KE was found to be upstream of intracellular bile acid accumulation, apoptosis, and cholestasis.

The AOP network was organized such that the identified MIEs resulted in 2 types of downstream KEs, namely the adverse response KEs and the adaptive response KEs. KEs driving the adverse response include intracellular bile acid accumulation, endoplasmic reticulum stress, oxidative stress, inflammation, apoptosis, and necrosis (Fig. 5). The adverse response KE with the highest incidence was intracellular bile acid accumulation. The KER between the transporter changes KE leading to intracellular bile acid accumulation was among those with the highest TOT_{KER}, showing the relatively high confidence in the KER (Fig. 5). The AOP network showed intracellular bile acid accumulation also as a prominent upstream KE leading to downstream adverse response KEs, including inflammation, oxidative stress, apoptosis, and necrosis. The rise in the concentration of noxious bile acids may lead to the opening of the mitochondrial membrane permeability pore, which will subsequently lead to the formation of reactive oxygen species (ROS), in turn resulting in oxidative stress [38,39]. Oxidative stress showed to be a prominent upstream and downstream KE in the AOP network. As a downstream KE, oxidative stress showed to have a relatively high TOT_{KER} with the transporter changes KE. Apoptosis, necrosis, and cholestasis were found to be downstream of oxidative stress. The formation of ROS and/or the increase in intracellular bile acid concentration may also lead to inflammatory responses [40]. The inflammation KE showed comparative relationships with oxidative stress and were both upstream of apoptosis and necrosis. Both types of cell death have been associated with chemical-induced cholestasis [41], but apoptosis had a considerably higher incidence (Fig. 4).

A major part of the AOP network on chemical-induced cholestasis relates to bile acid regulation changes and adaptive response. Nuclear receptors are critical regulators of the synthesis and metabolism of bile acids, but also control the expression of bile acid transporters [18,42]. The data used for optimizing the AOP network show the overarching role of nuclear receptor changes as a KE. Nuclear receptor changes constituted the third mostly occurring KE in the AOP network and were found both downstream and upstream of transporter changes, intracellular bile acid accumulation, bile acid synthesis changes and bile acid metabolism changes (Fig. 4). Nuclear receptor changes as an upstream KE, leading to transporter changes as the downstream KE, had the second highest confidence score in a KER (Fig. 5). This highlights its role in counteracting the disturbed bile acid homeostasis in chemical-induced cholestasis. Farnesoid X receptor (FXR), small heterodimer partner (SHP), pregnane X receptor (PXR), and constitutive androstane receptor (CAR) are the most prominent nuclear receptors associated with disturbed bile acid homeostasis [2]. In the current AOP network, these four nuclear receptors were included as KEMs. FXR constituted a 50 % of all reported nuclear receptors, while SHP, PXR, and CAR were accounted for 10 %, 10 %, and 8 %, respectively.

4. Discussion

AOPs and their networks are to be considered dynamic tools requiring continuous updating by feeding in relevant available and newly generated data. Furthermore, in view of expediting their use for real-life applications, including industrial and regulatory testing purposes, it is important for AOPs to gain end-user confidence and acceptance. This can be achieved by developing, optimizing, and assessing AOP constructs in compliance with OECD guidance, including consideration of the tailored Bradford-Hill criteria [13,43]. However, current AOP development, optimization and assessment strategies typically rely on manual labour and therefore may be prone to bias and data gaps [44]. Recent developments in natural language processing algorithms can provide a solution to this problem. In this respect, Sysrev [21] has emerged as a user-friendly tool to collect and extract data from scientific papers both by human and artificial intelligence-assisted approaches, which can be fed into AOPs. By using Sysrev, 6572 papers were screened for eligibility in the present study, of which 148 papers were finally selected. It is important to stress that the artificial intelligence-assisted screening method may have overlooked some relevant papers, resulting in potential false negatives. In an effort to address this issue, a team of 3 experts was assigned to train the machine-learning algorithm by reviewing 10 % of the initial set of eligible papers. However, the concern of missed data still remains and may require further fine-tuning to ensure a fully thorough automated screening process. Fortunately, artificial intelligence-assisted data collection and extraction is a rapidly evolving field with new advancements being reported on a frequent basis. As more advanced natural language processing models, such as ChatGPT, are developed, artificial intelligence-extracted data may provide a more complete depiction of real-world information [45].

The WoE assessment methodology involved scoring the data extracted from relevant papers which were selected using artificial intelligence, resulting in a highly informative AOP network on chemical-induced cholestasis (Fig. 3). The AOP network presented the incidence of KEs as observed in the data and indicated which ones had been more extensively studied. The relations between KEs or KERs were

assessed by a WoE scoring strategy based on the tailored Bradford-Hill criteria, which allowed for visualizing the relative confidence levels for each KER. Researchers can interact with this data by using the "xploreaop" web application (https://rstudio-connect.hu.nl/xploreaop/) and will be able to modify the AOP network's appearance by selecting specific parameters such as chemical selection and organism type. Additionally, this web application will facilitate the process of updating the AOP network by enabling the incorporation of additional future datasets.

AOPs are versatile tools that should be developed and optimized fit-for-purpose. In this respect, AOPs and their networks can serve a plethora of purposes relevant to toxicology and risk assessment, including chemical prioritization, chemical grouping and integrated approaches to testing and assessment [9,19,44,46,47]. The AOP network optimized in the present study is in the first instance being used as the conceptual mechanistic basis for setting up a battery of *in vitro* assays to predict cholestatic liver injury induced by chemical compounds from various applicability domains. In fact, this is embedded in a 2-tiered testing approach, whereby first tier testing relies on measuring transcriptional changes indicative of cholestatic liver injury. Unlike other types of hepatotoxicity, such transcriptional signature for chemical-induced cholestatic liver insult does not have sufficient predictive value on its own, and hence cannot be used as a stand-alone method [48]. Predictive power can be considerably increased when following up with second tier testing by applying a battery of *in vitro* assays mechanistically anchored in the AOP network, in which each assay monitors a selected MIE and KE individually at the translational level, but preferably at the activity level. This 2-tiered testing approach is currently being developed as part of the European ONTOX project (https://ontox-project.eu/), which combines ontologies and artificial intelligence for the purpose of animal-free and human-centered prediction of chemical-induced adversities, including in the liver [49]. The optimized and fully assessed AOP network resulting from the present study provides an important contribution to this goal, and will thus assist in delivering safer chemicals, including pharmaceutical drugs, while using fewer animals.

5. Conclusions

This study presented a novel approach for developing and/or optimizing AOP networks with the collection of data assisted by artificial intelligence and the quantification of the tailored Bradford-Hill criteria for the WoE assessment. This approach was able to optimize the AOP network on chemical-induced cholestasis by displaying the incidence of KEs and the relative confidence of KERs (Fig. 3). The creation of the "xploreaop" web application allowed for data interaction and will ease the addition of future datasets on chemical-induced cholestasis. This optimized AOP network may serve as a mechanistic compass for the development of a battery of *in vitro* assays to reliably predict chemical-induced cholestatic injury.

Declarations.

Statement of significance.

The existing AOP network on chemical-induced cholestasis is outdated. Current AOP development strategies rely on manual processes, which not only is a time-consuming effort, but also bears the risk of overlooking critical data. The present study introduces a novel approach for AOP network optimization of a previously published AOP network on chemical-induced cholestasis using artificial intelligence to facilitate automated data collection followed by subsequent quantitative confidence assessment of molecular initiating events, key events, and key event relationships.

CRediT authorship contribution statement

Jonas van Ertvelde: Conceptualization, Visualization, Writing – original draft, Writing – review & editing. **Anouk Verhoeven:** Conceptualization, Writing – review & editing. **Amy Maerten:** Writing – review & editing. **Axelle Cooreman:** Writing – review & editing. **Bruna dos Santos Rodrigues:** Writing – review & editing. **Julen Sanz-Serrano:** Writing – review & editing. **Milos Mihajlovic:** Writing – review & editing. **Ignacio Tripodi:** Writing – review & editing. **Marc Teunis:** Software. **Ramiro Jover:** Writing – review & editing. **Thomas Luechtefeld:** Software. **Tamara Vanhaecke:** Writing – review & editing. **Jian Jiang:** Conceptualization. **Mathieu Vinken:** Conceptualization, Writing – review & editing.

Declaration of Competing Interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Mathieu Vinken reports financial support was provided by European Commission. Mathieu Vinken reports financial support was provided by Flemish government. Mathieu Vinken reports financial support was provided by Research Foundation Flanders. Mathieu Vinken reports financial support was provided by Scientific Fund Willy Gepts. Mathieu Vinken reports financial support was provided by Center for Alternatives to Animal Testing. Mathieu Vinken reports financial support was provided by Alternatives Research and Development Foundation.

Acknowledgements

The development of the Shiny application was supported by the work in the Bachelor thesis performed by Sander van Tuijl under supervision of M. Teunis, at the University of Applied Sciences, Utrecht, The Netherlands.

Funding

This work was financially supported by the European Commission under the Horizon2020 Research and Innovation Framework program (grant number 963845 "ONTOX"), the Flemish government (Methusalem program), the Research Foundation Flanders, the Scientific Fund Willy Gepts, the Center for Alternatives to Animal Testing and the Alternatives Research and Development Foundation.

Ethics approval and consent to participate:

Not applicable.

Consent for publication

Not applicable.

Availability of data and materials

The datasets generated and/or analysed during the current study are available in the GitHub repository, https://github.com/ontox-project/xploreaop/blob/main/data-raw/D020.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jbi.2023.104465.

References

[1] F.Q. Onofrio, G.M. Hirschfield, The Pathophysiology of Cholestasis and Its Relevance to Clinical Practice, Clin. Liver Disease 15 (3) (2020) 110–114.
[2] E. Gijbels, V. Vilas-Boas, N. Deferm, L. Devisscher, H. Jaeschke, P. Annaert, M. Vinken, Mechanisms and In Vitro Models of Drug-Induced Cholestasis, Arch. Toxicol. 93 (5) (2019) 1169–1186.
[3] W.M. Lee, Drug-Induced Acute Liver Failure, Clin. Liver Dis. 17 (4) (2013) 575–586.

[4] W.F.P.M. Van den Hof, A. Ruiz-Aracama, A. Van Summeren, D.G.J. Jennen, S. Gaj, M.L.J. Coonen, K. Brauers, W.K.W.H. Wodzig, J.H.M. van Delft, J.C.S. Kleinjans, Integrating Multiple Omics to Unravel Mechanisms of Cyclosporin A Induced Hepatotoxicity In Vitro, Toxicol. In Vitro 29 (3) (2015) 489–501.

[5] A. Pieters, E. Gijbels, B. Cogliati, P. Annaert, L. Devisscher, M. Vinken, Biomarkers of cholestasis, *Biomarkers Med.* 15 (6) (2021) 437–454.

[6] E. Gijbels, A. Pieters, K. De Muynck, M. Vinken, L. Devisscher, Rodent Models of Cholestatic Liver Disease: A Practical Guide for Translational Research, Liver Int.: Off. J. Int. Assoc. Study Liver 41 (4) (2021) 656–682.

[7] National Research Council. (2007). Toxicity Testing in the 21st Century: A Vision and a Strategy. *The National Academies Press*.

[8] G.T. Ankley, S.W. Edwards, The Adverse Outcome Pathway: A Multifaceted Framework Supporting 21(st) Century Toxicology, Curr. Opin. Toxicol. 9 (2018) 1–7.

[9] M. Vinken, The Adverse Outcome Pathway Concept: A Pragmatic Tool in Toxicology, Toxicology 312 (2013) 158–165.

[10] D. Knapen, M.M. Angrish, M.C. Fortin, I. Katsiadaki, M. Leonard, L. Margiotta-Casaluci, S. Munn, J.M. O'Brien, N. Pollesch, L.C. Smith, X. Zhang, D.L. Villeneuve, Adverse Outcome Pathway Networks I: Development and Applications, Environ. Toxicol. Chem. 37 (6) (2018) 1723–1733.

[11] D.L. Villeneuve, M.M. Angrish, M.C. Fortin, I. Katsiadaki, M. Leonard, L. Margiotta-Casaluci, S. Munn, J.M. O'Brien, N.L. Pollesch, L.C. Smith, X. Zhang, D. Knapen, Adverse Outcome Pathway Networks II: Network Analytics, Environ. Toxicol. Chem. 37 (6) (2018) 1734–1748.

[12] M. Vinken, B. Landesmann, M. Goumenou, S. Vinken, I. Shah, H. Jaeschke, C. Willett, M. Whelan, V. Rogiers, Development of an Adverse Outcome Pathway From Drug-Mediated Bile Salt Export Pump Inhibition to Cholestatic Liver Injury, Toxicol. Sci. 136 (2013) 97–106.

[13] R.A. Becker, G.T. Ankley, S.W. Edwards, S.W. Kennedy, I. Linkov, B. Meek, M. Sachana, H. Segner, B. Van Der Burg, D.L. Villeneuve, H. Watanabe, T.S. Barton-Maclaren, Increasing Scientific Confidence in Adverse Outcome Pathways: Application of Tailored Bradford-Hill Considerations for Evaluating Weight of Evidence, Regul. Toxicol. Pharm. 72 (3) (2015) 514–537.

[14] M. de Lima Toccafondo Vieira, C.A. Tagliati, Hepatobiliary Transporters in Drug-Induced Cholestasis: a Perspective on the Current Identifying Tools, Expert Opin. Drug Metab. Toxicol. 10 (4) (2014) 581–597.

[15] R. Sonoi, Y. Hagihara, Tight Junction Stabilization Prevents HepaRG Cell Death in Drug-Induced Intrahepatic Cholestasis, Biology Open 10 (2021) bio058606.

[16] M.G. Burbank, A. Burban, A. Sharanek, R.J. Weaver, C. Guguen-Guillouzo, A. Guillouzo, Early Alterations of Bile Canaliculi Dynamics and the Rho Kinase/Myosin Light Chain Kinase Pathway Are Characteristics of Drug-Induced Intrahepatic Cholestasis, Drug Metab. Dispos. 44 (11) (2016) 1780–1793.

[17] J.L. Boyer, Adaptive Regulation of Hepatocyte Transporters in Cholestasis, in: The Liver, Wiley, 2020, pp. 378–389.

[18] G. Zollner, H.-U. Marschall, M. Wagner, M. Trauner, Role of Nuclear Receptors in the Adaptive Response to Bile Acids and Cholestasis: Pathogenetic and Therapeutic Considerations, Mol. Pharm. 3 (2006) 231–251.

[19] E. Gijbels, V. Vilas-Boas, P. Annaert, T. Vanhaecke, L. Devisscher, M. Vinken, Robustness Testing and Optimization of an Adverse Outcome Pathway on Cholestatic Liver Injury, Arch. Toxicol. 94 (4) (2020) 1151–1172.

[20] W. Chang, J. Cheng, J. Allaire, C. Sievert, B. Schloerke, Y. Xie, J. Allen, J. McPherson, A. Dipert, B. Borges, Shiny: Web Application Framework for R. R package version 1.7.4 (2022).

[21] T. Bozada, J. Borden, J. Workman, M. Del Cid, J. Malinowski, T. Luechtefeld, Sysrev: A FAIR Platform for Data Curation and Systematic Evidence Review, Front. Artif. Intell. 4 (2021) 105.

[22] R Core Team, R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing (2023).

[23] J. Allaire, C. Gandrud, K. Russell, C. Yetman, networkD3: D3 JavaScript Network Graphs from R. R package version 0.4 (2017).

[24] B.V. Almende, Contributors, B. Thieurmel, visNetwork: Network Visualization Using 'vis.js' Library. R package version 2.1.2 (2022).

[25] G. Csardi, T. Nepusz, The igraph Software Package for Complex Network Research. *InterJournal, Complex Systems, 1695* (2006).

[26] T. Pedersen, ggraph: An Implementation of Grammar of Graphics for Graphs and Networks. R package version 2.1.0 (2022).

[27] T. Pedersen, tidygraph: A Tidy API for Graph Manipulation. R package version 1.2.3, (2023).

[28] H. Wickham, M. Averick, J. Bryan, W. Chang, L.D. McGowan, R. François, G. Grolemund, A. Hayes, L. Henry, J. Hester, M. Kuhn, T.L. Pedersen, E. Miller, S. M. Bache, K. Müller, J. Ooms, D. Robinson, D.P. Seidel, V. Spinu, K. Takahashi, D. Vaughan, C. Wilke, K. Woo, H. Yutani, Welcome to the tidyverse, J. Open Source Software 4 (2019) 1686.

[29] M.L. Fernández-Murga, P.D. Petrov, I. Conde, J.V. Castell, M.J. Goméz-Lechón, R. Jover, Advances in Drug-Induced Cholestasis: Clinical Perspectives, Potential Mechanisms and In Vitro Systems, Food Chem. Toxicol. 120 (2018) 196–212.

[30] S. Childs, R.L. Yeh, E. Georges, V. Ling, Identification of a Sister Gene to P-glycoprotein, Cancer Res. 55 (1995) 2029–2034.

[31] E. Ogimura, S. Sekine, T. Horie, Bile Salt Export Pump Inhibitors are Associated with Bile Acid-Dependent Drug-Induced Toxicity in Sandwich-Cultured Hepatocytes, Biochem. Biophys. Res. Commun. 416 (3-4) (2011) 313–317.

[32] E. Kis, E. Ioja, Z. Rajnai, M. Jani, D. Méhn, K. Herédi-Szabó, P. Krajcsi, BSEP Inhibition - In Vitro Screens to Assess Cholestatic Potential of Drugs, Toxicol. In Vitro 26 (8) (2012) 1294–1299.

[33] A. Tabernilla, B. Dos Santos Rodrigues, A. Pieters, A. Caufriez, K. Leroy, R. Van Campenhout, A. Cooreman, A.R. Gomes, E. Arnesdotter, E. Gijbels, M. Vinken, In Vitro Liver Toxicity Testing of Chemicals: A Pragmatic Approach, Int. J. Mol. Sci. 22 (2021) 5038.

[34] E.-I. Lepist, H. Gillies, W. Smith, J. Hao, C. Hubert, R.L. St. Claire, K.R. Brouwer, A. S. Ray, H.W. van Veen, Evaluation of the Endothelin Receptor Antagonists Ambrisentan, Bosentan, Macitentan, and Sitaxsentan as Hepatobiliary Transporter Inhibitors and Substrates in Sandwich-Cultured Human Hepatocytes, PLoS One 9 (1) (2014) e87548.

[35] E.M. Leslie, P.B. Watkins, R.B. Kim, K.L.R. Brouwer, Differential Inhibition of Rat and Human Na+-Dependent Taurocholate Cotransporting Polypeptide (NTCP/SLC10A1) by Bosentan: a Mechanism for Species Differences in Hepatotoxicity, J. Pharmacol. Exp. Ther. 321 (3) (2007) 1170–1178.

[36] R.E. Morgan, C.J. van Staden, Y. Chen, N. Kalyanaraman, J. Kalanzi, R. T. Dunn 2nd, C.A. Afshari, H.K. Hamadeh, A Multifactorial Approach to Hepatobiliary Transporter Assessment Enables Improved Therapeutic Compound Development, Toxicol. Sci.: Off. J. Soc. Toxicol. 136 (2013) 216–241.

[37] R.W. Yucha, K. He, Q. Shi, L. Cai, Y. Nakashita, C.Q. Xia, M. Liao, In Vitro Drug-Induced Liver Injury Prediction: Criteria Optimization of Efflux Transporter IC50 and Physicochemical Properties, Toxicol. Sci.: Off. J. Soc. Toxicol. 157 (2017) 487–499.

[38] M.J. Perez, O. Briz, Bile-Acid-Induced Cell Injury and Protection, World J. Gastroenterol. 15 (14) (2009) 1677.

[39] M.-M. Tiao, T.-K. Lin, P.-W. Wang, J.-B. Chen, C.-W. Liou, The Role of Mitochondria in Cholestatic Liver Injury, Chang Gung Med. J. 32 (2009) 346–353.

[40] B.L. Woolbright, H. Jaeschke, Inflammation and Cell Death During Cholestasis: The Evolving Role of Bile Acids, Gene Expr. 19 (3) (2019) 215–228.

[41] B.L. Woolbright, H. Jaeschke, Measuring Apoptosis and Necrosis in Cholestatic Liver Injury, Methods Mol. Biol. 1981 (2019) 133–147.

[42] M. Garcia, L. Thirouard, L. Sedès, M. Monrose, H. Holota, F. Caira, D.H. Volle, C. Beaudoin, Nuclear Receptor Metabolism of Bile Acids and Xenobiotics: A Coordinated Detoxification System with Impact on Health and Diseases, Int. J. Mol. Sci. 19 (2018) 3630.

[43] OECD. Draft Guidance Document for the Scientific Review of Adverse Outcome Pathways. Paris, France (2021).

[44] 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, Arch. Toxicol. 91 (11) (2017) 3477–3505.

[45] M.P. Polak, D. Morgan, Extracting Accurate Materials Data from Research Papers with Conversational Language Models and Prompt Engineering – Example of ChatGPT. *arXiv* (2023).

[46] E. Arnesdotter, N. Spinu, J. Firman, D. Ebbrell, M.T.D. Cronin, T. Vanhaecke, M. Vinken, Derivation, Characterisation and Analysis of an Adverse Outcome Pathway Network for Human Hepatotoxicity, Toxicology 459 (2021) 152856.

[47] M. Vinken, A Putative AOP for Pneumonia Related to COVID-19, Arch. Toxicol. 94 (9) (2020) 3343–3345.

[48] C. Parmentier, P. Couttet, A. Wolf, T. Zaccharias, B. Heyd, P. Bachellier, M. Uteng, L. Richert, Evaluation of Transcriptomic Signature as a Valuable Tool to Study Drug-Induced Cholestasis in Primary Human Hepatocytes, Arch. Toxicol. 91 (8) (2017) 2879–2893.

[49] M. Vinken, E. Benfenati, F. Busquet, J. Castell, D.-A. Clevert, T.M. de Kok, H. Dirven, E. Fritsche, L. Geris, R. Gozalbes, T. Hartung, D. Jennen, R. Jover, H. Kandarova, N. Kramer, C. Krul, T. Luechtefeld, R. Masereeuw, E. Roggen, S. Schaller, T. Vanhaecke, C. Yang, A.H. Piersma, Safer Chemicals Using Less Animals: Kick-Off of the European ONTOX Project, Toxicology 458 (2021) 152846.