AOPERA: A Proposed Methodology and Inventory of Effective Tools to Link Chemicals to Adverse Outcome Pathways

Taylor E. Rycroft[1], Christy M. Foran[1], Adam Thrash[2], Jeff C. Cegan[1], Robert Zollinger[3], Igor Linkov[1], Edward J. Perkins[1] and Natàlia Garcia-Reyero[1]

[1]Environmental Laboratory, U.S. Army Engineer Research and Development Center, Vicksburg, MS, USA; [2]Institute for Genomics, Biocomputing & Biotechnology, Mississippi State University, Starkville, MS, USA; [3]Contractor to the Environmental Laboratory, U.S. Army Engineer Research and Development Center, Concord, MA, USA

Abstract

New approaches, like the adverse outcome pathway (AOP) framework, have been developed to describe how chemicals cause toxicity by linking *in vitro* assays to adverse health outcomes. However, approaches, tools and resources for development of AOPs have not been well described. Here we review information resources for AOP development and define a streamlined process for linking a chemical to an existing AOP. We propose a four-step process to facilitate AOP development: link the uncharacterized chemical directly to molecular initiating events (MIEs), key events (KEs), or adverse outcomes (AOs); identify analogs with toxicological information for the uncharacterized chemical; link the characterized chemical (initial chemical if characterized, a characterized analog if initial chemical is not) to MIEs, KEs or AOs; and identify AOPs that contain the MIEs, KEs or AOs that were found in Steps 1 and 3. The process and library of informational resources proposed and tested here served as the foundation for an informational online tool (AOPERA) that helps practitioners identify their current-state knowledge gaps, navigate the four-step process, and connect to relevant resources. AOPERA can be found at https://igbb.github.io/AOPERA_HTML. Additionally, we anticipate that by simplifying and standardizing the process of linking a chemical to a known AOP, we will lower the barrier to entry for this objective and increase its accessibility to new practitioners. In turn, this may increase the demand for new or improved AOPs to which practitioners can link chemicals, thereby contributing to the expansion of the library of known AOPs.

1 Introduction

Adverse outcome pathways (AOPs) describe the cascade of physiological events that link toxicant exposure to a downstream adverse health outcome (Ankley et al., 2010). By mapping connected key events (KEs) or changes in biological state that are measurable and essential to the progression of a defined biological disturbance, AOPs offer a novel alternative for mechanistically assessing a wide array of substances with limited toxicity data (Vinken et al., 2013; Villeneuve et al., 2014). In contrast to traditional toxicity investigations, which expose an organism to a chemical and seek to determine *what* happens, AOPs try to ascertain *how* it happens (Rycroft et al., 2018). AOPs are increasingly gaining support in toxicology communities of practice because they offer more information than a traditional lethal concentra-

tion (e.g., LC_{50}) value and they help expand the hazard profiles of chemicals to a broader set of acute (e.g., skin sensitization) and chronic outcomes (e.g., developmental defects). AOPs are also being assessed for their potential to inform regulatory decisions, such as setting occupational exposure limits or reference doses (Perkins et al., 2015; Wittwehr et al., 2017). As a result of these trends, many practitioners are expected to seek out methods and tools for determining whether their chemicals of interest play a role in activating existing AOPs. Unfortunately, there is currently no standard method for making this linkage, nor is there an inventory of available tools that can assist the practitioner; this study seeks to address this gap.

At present, practitioners wishing to link chemicals to AOPs face several hurdles. To begin with, there are hundreds of thousands of natural and synthetic chemicals, both with a known
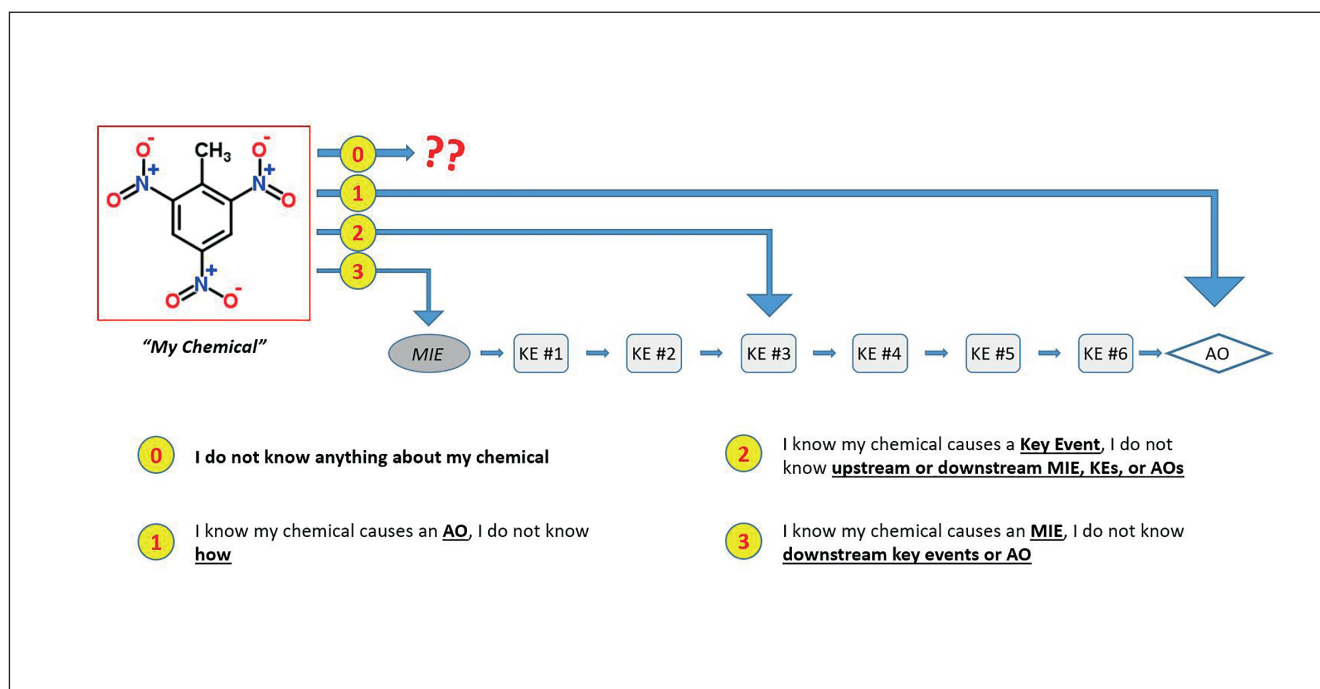
Fig. 1: Levels of information describing a practitioner's initial knowledge state pertaining to the toxicity of their chemical
Level 0: I do not know anything about my chemical; Level 1: I know my chemical causes an AO, but I do not know how; Level 2: I know my chemical causes a KE, but I do not know upstream or downstream MIE, KEs or AOs; Level 3: I know my chemical causes a MIE, but I do not know downstream key events or AO.

Chemical Abstracts Service Registry Number (CASRN) and without (uncharacterized) (Mitchell et al., 2013). Additionally, there are only a few hundred documented AOPs across all organisms[1], and, as a rule, they are chemical agnostic. As a result, it is unlikely that the chemical of interest to the practitioner is already linked to an existing AOP. In order to try to make this linkage *de novo*, the practitioner must perform focused empirical testing to determine whether the chemical triggers a molecular initiating event (MIE), KE or adverse outcome (AO). Such testing can be resource-intensive in terms of time, labor and funding, and must be prioritized (Tollefsen et al., 2014). Prioritization can be informed by first researching the chemical using traditional literature search methods and short-listing the molecular targets with the greatest potential. However, this strategy fails when the chemical being examined is novel or uncharacterized because toxicity data for such chemicals is limited or entirely absent. Fortunately, there are numerous informational resources and tools that improve upon the traditional literature search method and can assist a practitioner in linking both characterized and uncharacterized chemicals to AOPs. Unfortunately, locating these resources, determining their applicability, knowing when to use them for the case at hand, accessing them, and then integrating their outputs in a useful way can be a daunting task.

In this paper, we define a four-step process for linking a chemical to an existing AOP, and we highlight the available informational resources that can assist with each step of the process, il-lustrating their utility and features using a test case. We envision the process and library of informational resources proposed and tested here serving as the foundation for an informational software tool that helps practitioners identify their current-state knowledge gaps, navigate the four-step process, and connect to relevant resources. In order to aid with the process, we have developed the Adverse Outcome Pathway Exploratory Research Assistant (AOPERA), an online tool implementing the described process. Additionally, we anticipate that by simplifying and standardizing the process of linking a chemical to a known AOP, we will lower the barrier to entry for this objective and increase its accessibility to new practitioners. In turn, this may increase the demand for new or improved AOPs to which practitioners can link chemicals, thereby contributing to the expansion of known stressor-AOP linkages and potentially the expansion of the library of known AOPs.

2 Materials and methods

We developed a four-step process to assist a practitioner in identifying whether their chemical is linked to a known AOP. For each step of the process, we identified existing informational resources and tools that can assist the practitioner in completing the objective of each step. We then tested the process and resources using a representative chemical.
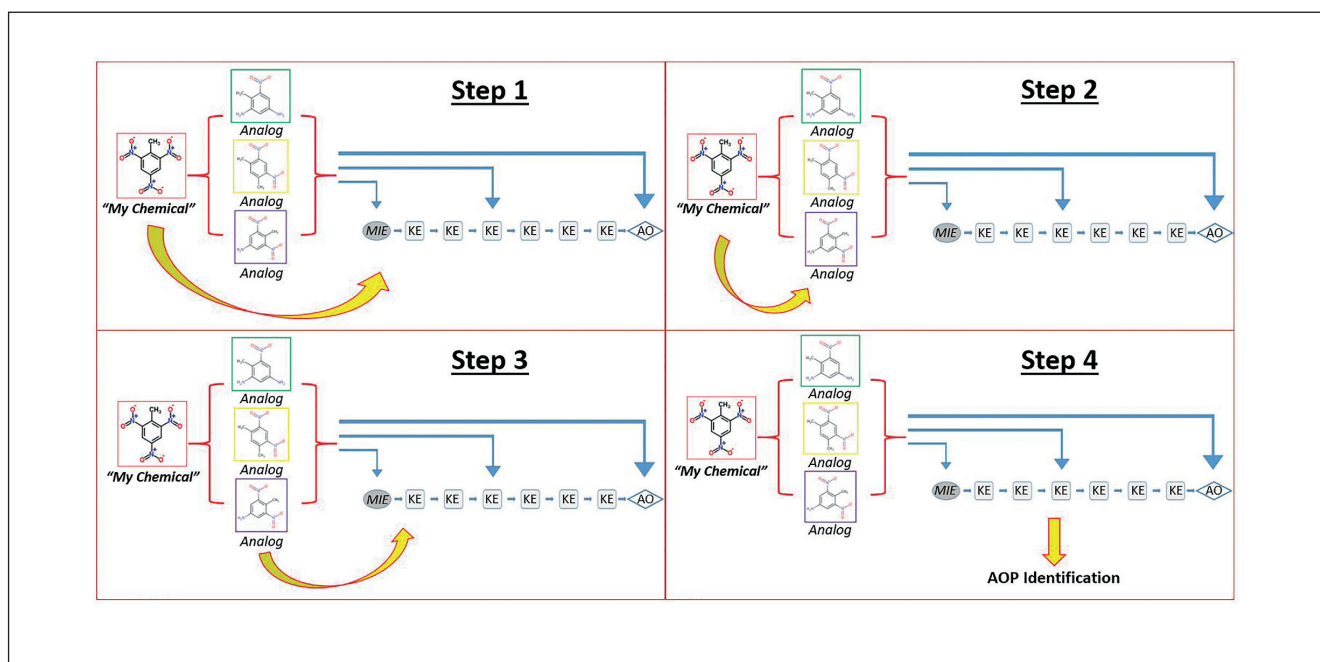
Fig. 2: The proposed four-step process of linking a chemical to a known AOP
Step 1: Link the uncharacterized chemical directly to MIEs, KEs or AOs. Step 2: Identify characterized chemical analogs for the uncharacterized chemical. Step 3: Link the characterized chemical (initial chemical if characterized, analog if initial chemical is uncharacterized) to MIEs, KEs or AOs; Step 4: Identify AOPs that contain the MIEs, KEs or AOs that were found in Steps 1 and 3.

Defining the initial knowledge state
Before a practitioner can link their chemical to an existing AOP, they must first identify the initial level of knowledge they have pertaining to the toxicity of their chemical (Fig. 1). At this preliminary stage, the practitioner must also confirm whether their chemical is characterized or uncharacterized.
− Level 0: The practitioner does not know anything about the toxicity of their chemical.
− Level 1: The practitioner knows that their chemical causes an AO but does not know the upstream MIE or KEs.
− Level 2: The practitioner knows that their chemical causes a KE but does not know upstream or downstream KEs or the MIE and AO.
− Level 3: The practitioner knows their chemical causes a MIE but does not know the downstream KEs or AO.

Initiating the process
If the practitioner identifies Level 0 as their initial knowledge state, then the four steps that they should follow to link their chemical to a known AOP are listed below and are shown in Figure 2. If the practitioner's chemical is uncharacterized, they should follow all four steps sequentially. If the chemical is characterized, then the practitioner need only execute Steps 3 and 4.
− Step 1: Link the uncharacterized chemical directly to MIEs, KEs or AOs.
− Step 2: Identify characterized chemical analogs for the uncharacterized chemical.

− Step 3: Link the characterized chemical (initial chemical if characterized, analog if initial chemical is uncharacterized) to MIEs, KEs or AOs.
− Step 4: Identify AOPs that contain the MIEs, KEs or AOs that were found in Steps 1 and 3.
At each step of the process, it is important to thoroughly document all MIEs, KEs, AOs and chemical analogs identified in Steps 1-3 of the process to ensure sufficient traceability of conclusions made in Step 4.

Identifying effective resources and testing the process
We performed a web search for available tools, models and informational resources that can assist a practitioner in completing the objective of each of the four steps. Resources were tested for their ability to perform the necessary function of the step, and information was recorded including the resource's ease of access and use, input requirements (e.g., drawing of chemical structure, simplified molecular-input line-entry system (SMILES) representation, etc.), and type of output (e.g., MIE/KE/AO, chemical analog, etc.). Several resources were found to be applicable to more than one step in the process (Tab. 6). Additionally, many resources were found to be inapplicable to our specific test case but to have potential for a broader set of cases.

We tested the four-step process using 2,4,6-trinitrotoluene (TNT) as our representative chemical. While TNT is a characterized chemical (CASRN: 118-96-7) and would not require that the practitioner execute Steps 1 and 2, we completed the process

Tab. 1: Resources that can assist in the process of linking an uncharacterized chemical directly to MIEs, KEs or AOs

Name	Ease of access/use	Input	Output	Available at
VEGA QSAR	++	SMILES	MIE, KE, AO	https://www.vegahub.eu/about-qsar/
OECD QSAR Toolbox	- -	Draw	AO	http://www.oecd.org/chemicalsafety/risk-assessment/oecd-qsar-toolbox.htm
Toxicity Estimation Software Tool (TEST)	+	Draw	AO	https://www.epa.gov/chemical-research/toxicity-estimation-software-tool-test
OncoLogic	+	Draw	AO	https://www.epa.gov/tsca-screening-tools/oncologictm-computer-system-evaluate-carcinogenic-potential-chemicals
EPA's New Chemical Categories	+	Manual category search	AO	https://www.epa.gov/reviewing-new-chemicals-under-toxic-substances-control-act-tsca/chemical-categories-used-review-new
Mcule	-	Draw	MIE, KE	https://mcule.com/

Very easy (++), easy (+), some difficulty (-), difficult (--)

as if TNT were uncharacterized so that we could thoroughly vet the method and verify findings using the real-world toxicity and chemical analog data available for TNT.

3 Results and discussion

Step 1
The practitioner's objective in Step 1 is to link their uncharacterized chemical directly to MIEs, KEs or AOs. We identified six resources that can potentially assist the practitioner in this step (Tab. 1) and evaluated each resource using TNT as our test case. In practice, this step would not be necessary for TNT, as most of the resources considered would recognize it as a characterized chemical. However, some of the resources make toxicity predictions based solely on chemical structure rather than documented empirical findings, so TNT still serves as a useful test of these resources' structural analysis capabilities.

Collectively, the six tools predicted that TNT is genotoxic, mutagenic, carcinogenic, hepatotoxic as well as a developmental toxicant and skin sensitizer. They also concluded that TNT would not bind to estrogen receptors. Thus, using this collection of resources, a practitioner can link an uncharacterized chemical directly to AOs using only a drawing of the chemical structure or a SMILES identifier. Individually, some tools were easier to use and demonstrated more utility than others for our specific test case, as described in Table 1. Additionally, some tools' output LC_{50}/LD_{50} values for various model organisms, while important to practitioners in other analyses, were not considered useful for this analysis because the health endpoint "death" does not substantially narrow the field of existing AOPs that a chemical may influence.

In our test case, VEGA QSAR (see Tab. 1) proved to be the easiest to use as well as the most informative, owing to its five non-lethal health endpoint predictions (mutagenicity, carcinoge-

nicity, developmental toxicity, skin sensitization and hepatotoxicity) and one MIE/KE prediction (estrogen receptor binding). The OECD QSAR Toolbox (see Tab. 1) proved to be a powerful tool, but it required a large file download and the user interface was quite complicated. The supporting documentation was extensive and distributed across numerous files, making it difficult for a user to leverage the tool for their needs in a timely manner and to troubleshoot issues. The output, however, was useful in our test case, revealing a potential AO (genotoxicity) and linking to literature that supported that conclusion. TEST was easy to download and use and predicted two potential non-lethal health endpoints (developmental toxicity and mutagenicity). OncoLogic (see Tab. 1) was also easy to use but it only offered one non-lethal endpoint (carcinogenicity) and it did not narrow its prediction to the *type* of cancer, likely making it too broad of a prediction to aid practitioners in linking their chemical to an AOP. The US Environmental Protection Agency's (EPA) New Chemical Categories document (see Tab. 1) was easy to use but not useful for our test case. TNT fell within the "polynitroaromatics" category of the document, which did not suggest a non-lethal health endpoint or likely mechanisms of action. Lastly, Mcule (see Tab. 1) was easy to access but difficult to use because it required the user to select from over 9,000 potential molecular targets to which the test chemical might bind, which is far too many to assess in an exploratory way. Docking scores were not benchmarked, so it was unclear whether a resultant score met an actionable threshold that indicated a likely MIE or KE. Additionally, many of the features of the tool required a fee-for-service upgrade to gain full benefits.

Step 2
The practitioner's objective in Step 2 is to identify characterized analogs for their uncharacterized chemical. We identified eleven resources that can potentially assist the practitioner in this step (Tab. 2) and evaluated each resource using TNT as our test case. The resources in this step return CASRNs of known chemicals

Tab. 2: Resources that can assist in the process of identifying analogs for an uncharacterized chemical

Name	Relevant findings and limitations	Available at
Toxicity Estimation Software Tool (TEST)	– Analogs based on toxicity similarity; were predicted to yield similar results in a fathead minnow LC_{50} (96 h), *Daphnia magna* LC_{50} (48 h), *T. pyriformis* IGC_{50} (48 h), or oral rat LD_{50} toxicity test.	https://www.epa.gov/chemical-research/toxicity-estimation-software-tool-test
Chemistry Dashboard ("CompTox")	– Only yielded results for molecular formulas that corresponded to chemicals with a CASRN. Therefore, if a practitioner with an uncharacterized chemical were to enter the formula for their chemical and it did not align with a single characterized chemical, then CompTox might not produce an analog. – Nearest-neighbor analogs predicted using OPERA v2.2 open-source models.	https://www.epa.gov/chemical-research/chemistry-dashboard
Analog Identification Methodology (AIM)	– Did not assign similarity coefficients, so there was no way to determine whether one analog had greater similarity to TNT than another.	https://www.epa.gov/tsca-screening-tools/analog-identification-methodology-aim-tool
ChemSpider	– Search times out when the Tanimoto similarity threshold is set to 0.99. – Similar to CompTox, ChemSpider's molecular formula search only yielded results for molecular formulas that corresponded to chemicals with a CASRN.	http://www.chemspider.com/AboutUs.aspx
eMolecules	– Some analogs were output as visuals without a name or CASRN; required drawing them in another software to identify them.	https://www.emolecules.com/info/aboutus-vision.html
PubChem	– No exact similarity coefficient is assigned to analogs; the tool only shows that analogs exceed a similarity threshold.	https://pubchemdocs.ncbi.nlm.nih.gov/about
ChemIDplus (part of TOXNET)	– Highest similarity coefficient threshold is 0.90; cannot narrow search further.	https://chem.nlm.nih.gov/chemidplus/jsp/chemidheavy/help.jsp
EURL ECVAM QSAR Database	– Some analogs were unlabeled and needed to be identified using another software tool.	https://eurl-ecvam.jrc.ec.europa.eu/databases/jrc-qsar-model-database
Mcule	– Unable to identify any TNT analogs with a similarity coefficient greater than 0.71.	https://mcule.com/
VEGA QSAR	– No challenges or limitations identified.	https://www.vegahub.eu/about-qsar/
ChemMaps	– Did not assign similarity coefficients, so there was no way to determine whether one analog had greater similarity to TNT than another. – The tool positions TNT in a map format in which the relationship to other chemicals is based on structural properties, however the specific chemical properties and the method that form the basis for the similarity dispositions are not specified, so the reliability is unclear.	http://www.chemmaps.com/

similar in structure or toxicity endpoints to the searched chemical. Similar to Step 1, this step would not be necessary for TNT because it addresses a need for an uncharacterized chemical and TNT is characterized, but TNT was still used in order to assess the utility of each tool for executing this step and to compare findings to real-world data.

Collectively, the eleven resources returned 34 distinct chemical analogs with a similarity coefficient greater than or equal to 0.90. The similarity coefficient was calculated by the eleven resources in Step 2 and presented alongside the analog outputs. Ten of the tools produced analogs they deemed similar in chemical structure to TNT, while one tool (TEST) appeared to produce analogs it considered similar in toxicity. All tools were easy to use; however, some demonstrated more utility than others for our specific test case. Key findings and limitations that we encountered are summarized in Table 2.

Step 3
The practitioner's objective in Step 3 is to link a characterized chemical to MIEs, KEs or AOs. This characterized chemical may be the practitioner's original chemical, in which case they would have skipped Steps 1 and 2, or it may be an analog identified in

Tab. 3A: Resources that can potentially assist *directly* in the process of linking a characterized chemical to MIEs, KEs or AOs

Name	Relevant findings and limitations	Available at
ToxCast (Toxicity Forecaster)	– Shows assays with target molecules and the number of active assays for that target (e.g. a positive result for interacting with the oxidoreductase enzyme in rat brain tissue).	https://www.epa.gov/chemical-research/toxcast-dashboard
VEGA QSAR	– Offers information pertaining to five AOs (mutagenicity, carcinogenicity, developmental toxicity, skin sensitization, and hepatotoxicity) and one MIE/KE (estrogen receptor binding).	https://www.vegahub.eu/about-qsar/
OECD QSAR Toolbox	– Offers information pertaining to an AO (genotoxicity) and links to supporting literature.	http://www.oecd.org/chemicalsafety/risk-assessment/oecd-qsar-toolbox.htm
Toxicity Estimation Software Tool (TEST)	– Offers information pertaining to two AOs (developmental toxicity and mutagenicity).	https://www.epa.gov/chemical-research/toxicity-estimation-software-tool-test
EPA's New Chemical Categories	– Organizes chemicals into 56 categories and presents hazard concerns for each category.	https://www.epa.gov/reviewing-new-chemicals-under-toxic-substances-control-act-tsca/chemical-categories-used-review-new
ACToR	– Offers information pertaining to four AOs (mutagenicity, carcinogenicity, genotoxicity and developmental toxicity) and lists cell types for positive results.	https://actor.epa.gov/actor/home.xhtml
ATSDR Toxic Substances Portal	– Provides detailed report on health effects that may include MIEs, KEs or AOs; lists studies that identify the effect.	https://www.atsdr.cdc.gov/substances/index.asp
Pharos Chemical and Material Library	– Lists hazards according to GHS physical, health, and environmental hazards and provides links to evidentiary literature.	https://www.pharosproject.net/
EDSP21	– Offers information pertaining to AO (endocrine disruption) and potentially MIEs/KEs. – Lists agonist, antagonist, binding (active, weak) and assay results for androgen receptor, estrogen receptor, thyroid receptor including corresponding AC_{50} values.	https://www.epa.gov/endocrine-disruption/endocrine-disruptor-screening-program-edsp-21st-century
Reactome	– Shows biological pathways leading to disease and chemicals that may influence the pathways.	https://reactome.org/what-is-reactome
WikiPathways	– Database of biological pathways; open, collaborative platform with 2700+ pathways in 20+ species. – Offers information pertaining to MIEs, KEs, and AOs.	https://www.wikipathways.org/index.php/WikiPathways
Chemical Carcinogenesis Research Information System (CCRIS)	– Offers information pertaining to AO (mutagenicity); cites study results from AMES test and tests on Chinese hamster cells.	https://toxnet.nlm.nih.gov/newtoxnet/ccris.htm
Comparative Toxicogenomics Database (CTD)	– Offers information pertaining to MIEs (gene interactions) and AOs (diseases resulting from gene interactions). – Provides inference genes and inference scores; inference scores are not benchmarked, so a user needs to determine the threshold for an actionable score.	http://ctdbase.org/about/;jsessionid=9E5314E6506A153691E3C85702016161
ChemView	– Lists possible AOs and documentation for how linkages were made. – Also links to IRIS assessments.	https://chemview.epa.gov/chemview
OncoLogic	– Offers information pertaining to AO (carcinogenicity). – Produces "Oncologic Justification Report" – Broad; does not say what type of cancer	https://www.epa.gov/tsca-screening-tools/oncologictm-computer-system-evaluate-carcinogenic-potential-chemicals

Name	Relevant findings and limitations	Available at
Mcule	– Offers information pertaining to MIEs/KEs (molecular binding). – User must select from over 9,000 potential molecular targets. – Docking scores are not benchmarked; no context for interpretation. – Additional functionality requires subscription.	https://mcule.com/
Mol-Instincts Database	– Offers information pertaining to MIEs/KEs (activity scores for GPCR ligands, ion channel modulators, kinase inhibitors, and nuclear receptor ligands). – Activity scores are not benchmarked; no context for interpretation. – License purchase required after 15-day free trial.	https://www.molinstincts.com/home/index/story/story01.html
Distributed Structure-Searchable Toxicity Database Network (DSSTox)	– Potentially redundant with EPA's ToxCast and Tox21 data. – Numerical matrix data structure is not intuitive.	https://www.epa.gov/chemical-research/distributed-structure-searchable-toxicity-dsstox-database
EURL ECVAM QSAR Database	– Offers information pertaining to AOs. – In our test case, the QMRF for "Endpoint: Mutagenicity" did not offer results of the mutagenicity test but instead linked to VEGA to run the model.	https://eurl-ecvam.jrc.ec.europa.eu/databases/jrc-qsar-model-database
EPA's Virtual Tissue Models (embryo, blood vessel development, developmental toxicity, and thyroid)	– Future potential; models not available at time of analysis.	https://www.epa.gov/chemical-research/virtual-tissue-models-predicting-how-chemicals-impact-development
OpenFoodTox (EFSA)	– Offers information pertaining to AOs as well as reference values. – Links to EFSA opinions.	https://www.efsa.europa.eu/en/microstrategy/openfoodtox
NICEATM Integrated Chemical Environment (ICE)	– Offers information pertaining to four AOs (skin sensitization, skin irritation, eye irritation, and endocrine disruption).	https://ice.ntp.niehs.nih.gov/#!

Tab. 3B: Resources that can potentially assist *indirectly* in the process of linking a characterized chemical to MIEs, KEs or AOs
These resources do not meet the objective of Step 3 independently but may be assistive when initiating a standard literature search for a characterized chemical by connecting the user to relevant databases and peer-reviewed literature.

Name	Relevant findings and limitations	Available at
Chemistry Dashboard ("CompTox")	– Notes presence in lists, rates data quality, and links to databases with bioassay data.	https://www.epa.gov/chemical-research/chemistry-dashboard
ChemSpider	– Lists articles that mention the chemical. – "Safety" feature links to GHS hazard statements.	http://www.chemspider.com/AboutUs.aspx
eChemPortal	– Links to record of chemical in participating databases where there are additional links to literature on toxicity studies.	https://www.echemportal.org/echemportal/index.action
QuickGO	– Links to literature pertaining to various proteins, enzymes, and receptors' responses to chemical.	https://www.ebi.ac.uk/QuickGO/
AmiGO	– Links to literature that references the chemical.	http://amigo.geneontology.org/amigo
International Toxicity Estimates for Risk (ITER)	– Links to numerous governmental organizations' determinations for risk values pertaining to four endpoint-routes: noncancer-oral, cancer-oral, noncancer-inhalation, and cancer-inhalation. – Links to literature that shows how risk values were derived and how the health endpoint was reached.	https://toxnet.nlm.nih.gov/newtoxnet/iter.htm

Name	Relevant findings and limitations	Available at
GENE-TOX	– Connects user to relevant peer-reviewed literature.	https://toxnet.nlm.nih.gov/cgi-bin/sis/htmlgen?GENETOX
Hazardous Substances Databank (HSDB)	– Connects user to relevant peer-reviewed literature.	https://toxnet.nlm.nih.gov/cgi-bin/sis/htmlgen?HSDB
TOXLINE	– Connects user to relevant peer-reviewed literature.	https://toxnet.nlm.nih.gov/newtoxnet/toxline.htm
LactMed (Drugs and Lactation DB)	– Connects user to relevant peer-reviewed literature.	https://toxnet.nlm.nih.gov/newtoxnet/lactmed.htm
DART (Developmental and Reproductive Toxicology DB)	– Connects user to relevant peer-reviewed literature.	https://toxnet.nlm.nih.gov/newtoxnet/dart.htm

Step 2. We identified 22 resources that can potentially assist the practitioner in this step directly (Tab. 3A), as well as 11 resources that may assist indirectly as the preliminary steps of a standard literature search (Tab. 3B). We evaluated each resource using a single TNT analog identified in Step 2. The selected analog was 4-amino-2,6-dinitrotoluene (CASRN: 19406-51-0), which eMolecules assigned a > 0.95 similarity coefficient relative to TNT. When a resource did not have a record of 4-amino-2,6-dinitrotoluene, we examined other chemicals in the resource's records to determine whether the resource might produce useful outcomes for other chemicals. Key findings and limitations that we encountered are summarized in Tables 3A and 3B.

Eight of the 22 resources listed in Table 3A yielded potential MIEs and AOs for the TNT analog 4-amino-2,6-dinitrotoluene (VEGA QSAR, OECD QSAR Toolbox, TEST, ACToR, CTD, CCRIS, OncoLogic, and Mol-Instincts database). The 4 potential MIEs were interactions with GPCR ligands, ion channel modulators, kinase inhibitors or nuclear receptor ligands, and the 8 potential AOs were genotoxicity, mutagenicity, carcinogenicity, developmental toxicity, skin sensitization, stomach neoplasms, prostatic neoplasms and liver cirrhosis (Tab. 4). In some instances, one tool would make a prediction contradictory to that of another tool, such as predicting carcinogenicity when another identified the analog as non-carcinogenic; in these cases we accepted the more conservative result (a positive finding) and documented the discrepancy. Alternatively, a practitioner not in favor of taking the more conservative result to resolve contradictory predictions may opt instead to take a weight-of-evidence approach and generate additional predictions for the same health endpoint using other characterized analogs from the list of analogs produced in Step 2.

Step 4
The practitioner's objective in Step 4 is to identify existing AOPs that contain the MIEs, KEs or AOs that were found in Steps 1 and 3 (Tab. 4). We identified three resources that can potentially assist the practitioner in this step (Tab. 5) and evaluated each resource

using the MIEs, KEs and AOs identified from using TNT and the TNT analog 4-amino-2,6-dinitrotoluene as our test cases.

The first resource, Reactome, is an open-source, open access, manually curated and peer-reviewed biomolecular pathway database with tools for the visualization, interpretation and analysis of pathway knowledge. The second resource, AOP Knowledge Base (AOP-KB), is a project launched by the OECD to enable the scientific community, in one central location, to share, develop and discuss their AOP related knowledge. It allows all stakeholders to build AOPs by entering and then linking information about MIEs, KEs, AOs and chemical initiators[2]. The AOP-KB consists of four connected modules. The first module, e.AOP.Portal, enables search functionality within the AOP-KB. The second module, AOP-Wiki, provides a system that organizes the available knowledge and published research into a verbal description of individual pathways, using a user-friendly wiki interface. It maintains a database of AOPs, KEs, KE relationships, and stressors. The third module, Effectopedia, is a modeling platform that offers the ability to visually design and explore AOPs as well as display quantitative information about them. And the fourth original module, now the third-party tool AOPXplorer, is a computational tool that enables automated graphical representation of AOPs and the networks among them. The third resource, WikiPathways, is an open, collaborative platform dedicated to the curation of biological pathways. It was established to facilitate the contribution and maintenance of pathway information by the biology community and to reduce the barrier to participate in pathway curation.

A search within Reactome, AOP-KB, and WikiPathways for the MIEs and AOs shown in Table 4 yielded 39 unique AOPs. Of these, 15 had a development status that exceeded our screening threshold. For the AOP-KB, this threshold was a status of development that had progressed beyond "Under Development," which meant that each identified AOP must have undergone some degree of review, approval or endorsement by the OECD's Task Force for Hazard Assessment / Working Group of the National

---
[2] https://aopkb.oecd.org/index.html

Tab. 4: The MIEs, KEs and AOs predicted in Steps 1 and 3

Step	MIEs	KEs	AOs
Step 1 (Test chemical: 2,4,6-trinitrotoluene, TNT)	*N/A*	*N/A*	– Genotoxicity – Mutagenicity – Carcinogenicity – Hepatotoxicity – Developmental toxicity – Skin sensitization
Step 3 (Test analog: 4-amino-2,6-dinitrotoluene)	Interactions with: – GPCR ligands – Ion channel modulators – Kinase inhibitors – Nuclear receptor ligands	*N/A*	– Genotoxicity – Mutagenicity – Carcinogenicity – Developmental toxicity – Skin sensitization – Stomach neoplasms – Prostatic neoplasms – Liver cirrhosis

N/A, not applicable

Tab. 5: Resources that can assist in the process of identifying existing AOPs that contain the MIEs, KEs or AOs that were found in Steps 1 and 3

Name	Relevant findings and limitations	Available at
Reactome	– Open-source, open access, manually curated and peer-reviewed pathway database with tools for the visualization, interpretation and analysis of pathway knowledge.	https://reactome.org/what-is-reactome
AOP Knowledge Base (AOP-KB)	– Consists of four modules: e.AOP.Portal, AOP-Wiki, Effectopedia and AOPXplorer	https://aopkb.oecd.org/index.html
WikiPathways	– Open, collaborative platform that includes a graphical pathway editing tool and integrated databases covering major gene, protein, and small-molecule systems.	https://www.wikipathways.org/index.php/WikiPathways

Coordinators of the Test Guideline Programme (TFHA/WNT) or Extended Advisory Group on Molecular Screening and Toxicogenomics (EAGMST). For Reactome, the acceptance threshold was publication in the database.

With the short-list of 15 potential AOPs generated in Step 4, the practitioner in our test case would proceed with determining which AOPs applied to their chemical (TNT) by performing empirical tests. The most resource-effective method would be to test MIEs or KEs that overlap with multiple AOPs so that a single test may rule in/out more than one AOP.

Outcomes and limitations of the process
The four-step process proposed and tested herein can assist a practitioner in identifying whether their chemical is linked to a known AOP. We identified 30 existing resources that may be directly assistive in this effort, some of which provide utility to multiple steps (Tab. 6). The outcome of the process is a list of potential AOPs that may be influenced by the chemical under investigation. This priority list of AOPs and associated MIEs, KEs and AOs can then be used to inform more targeted *in vitro* or *in vivo* testing for confirmation of significance, such as with the EPA's "six-pack" of acute toxicity studies (U.S. EPA, 2018). Thus, the four-step *in*

silico process effectively serves as a screening method that can be performed prior to expenditure of limited laboratory resources. Like any lead-generation process, however, it is important to emphasize that the resulting list of potential AOPs is not exhaustive or conclusive, as many biomolecular pathways will be missing from the list, and conclusions can only be drawn after thorough toxicity testing.

The demonstrated process, while effective at generating a short-list of existing AOPs that the practitioner's chemical may play a role in activating, does have a few important limitations. One issue that we encountered when searching within the resources in Step 4 is that many of the MIE and AO outputs from Steps 1 and 3 were quite broad and resulted in a large number of AOP hits. For example, an interaction with "ion channel modulators" does not inform which ion channel (sodium, potassium, etc.) may be modulated and how the interaction modulates the channel (activate, inhibit, etc.). As a result, we conservatively accepted the 14 resultant AOPs that had referenced any form of modulation to any ion channel. Conversely, some MIEs or AOs in our list, e.g., prostatic neoplasms, were not linked to any potential AOPs. This is more a limitation of the specificity of the supporting resources in Steps 1, 3 and 4 than of the four-step process, but it results in

Tab. 6: Summary of resources that support the proposed four-step process and their applicability to each step

Name	Step 1	Step 2	Step 3	Step 4
VEGA QSAR	X	X	X	
OECD QSAR Toolbox	X		X	
Toxicity Estimation Software Tool (TEST)	X	X	X	
OncoLogic	X		X	
EPA's New Chemical Categories	X		X	
Mcule	X	X	X	
Chemistry Dashboard ("CompTox")		X		
Analog Identification Methodology (AIM)		X		
ChemSpider		X		
eMolecules		X		
PubChem		X		
ChemIDplus (part of TOXNET)		X		
EURL ECVAM QSAR Database		X	X	
ChemMaps		X		
ToxCast (Toxicity Forecaster)			X	
ACToR			X	
ATSDR Toxic Substances Portal			X	
Pharos Chemical and Material Library			X	
EDSP21			X	
Reactome			X	X
WikiPathways			X	X
Chemical Carcinogenesis Research Information System (CCRIS)			X	
Comparative Toxicogenomics Database (CTD)			X	
ChemView			X	
Mol-Instincts Database			X	
Distributed Structure-Searchable Toxicity Database Network (DSSTox)			X	
EPA's Virtual Tissue Models (Embryo, Blood Vessel Development, Developmental Toxicity, and Thyroid)			X	
OpenFoodTox (EFSA)			X	
NICEATM Integrated Chemical Environment (ICE)			X	
AOP Knowledge Base (AOP-KB): e.AOP.Portal, AOP-Wiki, Effectopedia and AOPXplorer				X

some risk that a practitioner may conclude the *in silico* screening process with too many or too few potential AOPs to prioritize for further investigation.

Another key limitation of the process is that its success is dependent on the still developing library of existing AOPs. At the time of writing, this library was maintained within only two resources, Reactome and AOP-KB, and AOP-KB contained a total of 244 submitted AOPs, of which only 26 had progressed beyond "Under Development" status. Granted, there are significantly more known AOPs when these 244 AOPs are considered in a net-

work with overlapping KEs, but even after factoring in these permutations the library of AOPs is still quite restricted. If, by using the four-step process, a practitioner finds that their chemical may cause cancer, they are then limited to the six AOPs in AOP-KB that may lead to cancer, which is a gross underrepresentation of the numerous biomolecular mechanisms from which cancer can result. In order for this process to be truly impactful, the number of existing AOPs must be substantially increased.

We anticipate that the process and library of informational resources described in this manuscript would be particularly use-

ful to practitioners if converted to an informational software tool that helps users identify their current-state knowledge gaps, navigate the four-step process, and connect to relevant resources. To this end, we have started development of a web-based tool called the Adverse Outcome Pathway Exploratory Research Assistant (AOPERA). We have developed a pilot version of AOPERA[3]. AOPERA guides the practitioner throughout the four-step process and provides linkages and background information to each of the suggested tools.

4 Conclusion

At present, there is currently no standard method or inventory of tools that can assist a practitioner in determining whether a chemical of interest plays a role in activating an existing AOP. The four-step process that we propose and test here is a step toward closing this gap. By enabling practitioners to generate short-lists of AOPs with which their chemicals may interact, this process streamlines AOP lead-generation using *in silico* methods prior to requiring the use of resources for *in vitro* and *in vivo* testing. We anticipate that this process and library of informational resources would be particularly useful to practitioners if converted to an informational software tool that helps users identify their current-state knowledge gaps, navigate the four-step process, and connect to relevant resources. We acknowledge that the utility of this process would be greater if the supporting informational resources made more specific predictions (e.g., "prostate cancer" rather than "cancer") and the library of existing AOPs was significantly larger, and we expect that the limitations of the process will be improved as more practitioners utilize it and as the demand for new and more specific AOPs continues to grow.

References

Ankley, G. T., Bennett, R. S., Erickson, R. J. et al. (2010). Adverse outcome pathways: A conceptual framework to support ecotoxicology research and risk assessment. *Environ Toxicol Chem 29*, 730-741. doi:10.1002/etc.34

Mitchell, J., Pabon, N., Collier, Z. A. et al. (2013). A decision analytic approach to exposure-based chemical prioritiza-tion. *PLoS One 8*, e70911. doi:10.1371/journal.pone.0070911

Perkins, E. J., Antczak, P., Burgoon, L. et al. (2015). Adverse outcome pathways for regulatory applications: Examination of four case studies with different degrees of completeness and scientific confidence. *Toxicol Sci 148*, 14-25. doi:10.1093/toxsci/kfv181

Rycroft, T., Massey, O., Foran, C. M. et al. (2018). Weight of evidence frameworks in evaluation of adverse outcome pathways (303-316). In N. Garcia-Reyero and C. Murphy (eds.), *A Systems Biology Approach to Advancing Adverse Outcome Pathways for Risk Assessment*. Cham, Switzerland: Springer. doi:10.1007/978-3-319-66084-4

Tollefsen, K. E., Scholz, S., Cronin, M. T. et al. (2014). Applying adverse outcome pathways (AOPs) to support integrated approaches to testing and assessment (IATA). *Regul Toxicol Pharmacol 70*, 629-640. doi:10.1016/j.yrtph.2014.09.009

U.S. EPA – U.S. Environmental Protection Agency (2018). Data Requirements for Pesticide Registration. EPA Office of Pesticide Programs. https://www.epa.gov/pesticide-registration/data-requirements-pesticide-registration (accessed 27.09.2018).

Villeneuve, D., Crump, D., Garcia-Reyero, N. et al. (2014). Adverse outcome pathway (AOP) development I: Strategies and principles. *Toxicol Sci 142*, 312-320. doi:10.1093/toxsci/kfu199

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

Wittwehr, C., Aladjov, H., Ankley, G. et al. (2017). How adverse outcome pathways can aid the development and use of computational prediction models for regulatory toxicology. *Toxicol Sci 155*, 326-336. doi:10.1093/toxsci/kfw207

Competing financial interest declaration
The authors declare they have no actual or potential competing financial interests.

---

[3] https://igbb.github.io/AOPERA_HTML