Published in final edited form as:
Appl In Vitro Toxicol. 2017 December 1; 3(4): 298–311. doi:10.1089/aivt.2017.0017.

Creating a Structured AOP Knowledgebase via Ontology-Based Annotations

Cataia Ives[1], **Ivana Campia**[2], **Rong-Lin Wang**[3], **Clemens Wittwehr**[2], and **Stephen Edwards**[1,*]

[1]Integrated Systems Toxicology Division, NHEERL, U.S. Environmental Protection Agency, RTP, NC, USA

[2]European Commission's Joint Research Centre, NERL, U.S. Environmental Protection Agency, Cincinnati, OH, USA

[3]Exposure Methods and Measurements Division, NERL, U.S. Environmental Protection Agency, Cincinnati, OH, USA

Abstract

Introduction—The Adverse Outcome Pathway framework is increasingly used to integrate data generated based on traditional and emerging toxicity testing paradigms. As the number of AOP descriptions has increased, so has the need to define the AOP in computable terms.

Materials and Methods—Herein, we present a comprehensive annotation of 172 AOPs housed in the AOP-Wiki as of December 4, 2016 using terms from existing biological ontologies.

Results—AOP Key Events (KEs) were assigned ontology terms using a concept called the Event Component, which consists of a Process, an Object, and an Action term, with each term originating from ontologies and other controlled vocabularies. Annotation of KEs with ontology classes from fourteen ontologies and controlled vocabularies resulted in a total of 685 KEs being annotated with a total of 809 Event Components. A set of seven conventions resulted, defining the annotation of KEs via Event Components.

Discussion—This expanded annotation of AOPs allows computational reasoners to aid in both AOP development and applications. In addition, the incorporation of explicit biological objects will reduce the time required for converting a qualitative AOP description into a conceptual model that can support computational modeling. As high throughput genomics becomes a more important part of the high throughput toxicity testing landscape, the new approaches described here for annotating key events will also promote the visualization and analysis of genomics data in an AOP context.

Keywords

Adverse Outcome Pathway; key event; ontology; knowledgebase

Correspondence address. Stephen Edwards, Systems Biologist, ISTD, NHEERL, ORD, U.S. Environmental Protection Agency, 109 TW Alexander Drive (MD B305-01), Research Triangle Park, NC 27711, Edwards.Stephen@epa.gov.

Author Disclosure Statement
No competing financial interests exist.

Introduction

The twenty-first century has seen a shift toward the use of high throughput toxicity information to keep pace with the large number of chemicals in use today (1, 2). The need to assess an ever-increasing number of chemicals with better resource efficiency and reduced animal use, requires a shift to predictive toxicology, and a data and information management plan conducive to regulatory assessment.

The Adverse Outcome Pathway (AOP) framework, originally proposed by Ankley et al. (3), has emerged as an ideal tool for integrating various types of toxicity information from many sources (4) and thereby support predictive toxicology (Figure 1). An AOP portrays existing knowledge concerning the linkage between a molecular initiating event (MIE), the initial interaction between a xenobiotic and the biological system, and an adverse outcome (AO) that impacts individual health or well-being and/or the survival of a wildlife population. The AOP development process has matured (5, 6), and recent focus has shifted to the applications for this framework. To date, applications include Integrated Approaches to Testing and Assessment (IATA) (7, 8), Quantified Structure Activity Relationships (QSARs) (9), read-across (10, 11), safety evaluation of drugs (12), and prioritization of testing strategies and screening level hazard or risk assessments (13).

Given the key role for data management in supporting the many applications associated with AOPs, the AOP community has agreed to consolidate AOP information into a central knowledgebase (AOP-KB) that allows the scientific community to share, develop, and discuss AOPs. AOP-KB provides knowledge management for information supporting all phases of development, including putative, formal and quantitative AOP (qAOP) development (5). Identifying data in a computable standardized format, facilitates data integration and information sharing, including evidence integration (14). As the AOP-KB evolves to meet the needs identified by the OECD including open access, standardized representation of data, and consistency in reporting, a shared set of chemical, biological, and toxicological ontologies has emerged as a way to unify information across the AOP-KB.

An ontology can be considered as “an area of knowledge that is formalized, such that the individual terms (or concepts) are defined by a set of assertions that connect them to other terms” (15). It provides a shared controlled vocabulary to enable knowledge domain modeling through concepts, properties, and relations. There are a variety of ways to define and envision an ontology from the perspective of computational data science. Perhaps the simplest among them is to view an ontology as a set of logically connected nodes or classes where two nodes are related to each other by a property relationship in the form of a subject-predicate-object called a triple. The most common format for describing these triples is the resource description framework (RDF). Note that all classes and their relationships in an ontology are annotated by a collection of controlled vocabulary terms.

To enable automatic information sharing and integration across knowledge domains computationally, however, these RDF data have to be semantically modeled by the RDF Schema (RDFS) and its extension, Web Ontology Language (OWL). Through a set of predefined constructs of OWL classes, OWL individuals, and OWL properties, for example

subClassOf, equivalentClass, disjointWith, subPropertyOf, equivalentProperty, etc., RDFS/OWL describe ontology classes and properties and provide them with contextual relationships. In short, an ontology represents a computable RDF document semantically encoded by RDFS/OWL. Multiple ontologies from different knowledge domains and species with shared vocabulary can be computationally integrated as a common knowledgebase.

An ontologized AOP-KB (O-AOP-KB), where individual Key Events (KEs) are annotated with ontology terms, offers several benefits. First, it enables a computer reasoner to make inferences and discover additional connections among KEs and AOPs. Second, it facilitates query of the AOP-KB beyond simple text matching at varying degrees of granularity. If a query consists of individual KEs or Key Event Relationships (KERs) with each containing only a single ontology class, the logical hierarchical structure built into an ontology graph allows accurate and complete retrieval of all its subsumed or equivalent entities. Semantically similar KEs or KERs could also be calculated. If a query contains a group of KEs belonging to a common AOP involving multiple ontology classes, semantically similar AOPs could also be retrieved. In this case, an AOP becomes equivalent to a profile of KEs, similar to the phenotypic profiles anchored by genes or diseases in the semantic analysis demonstrated in the Monarch Initiatives (16, 17). Finally, O-AOP-KB should also facilitate modular development of AOPs by encouraging reuse of KEs or KERs and reducing redundancy. Given these benefits, this study aims to provide ontology-based annotations of AOP KEs in the AOP-Wiki, and to illustrate how the O-AOP-KB contributes to improving current practices in the development, evaluation, and optimization of AOPs.

Methods

Review of Ontologies and Controlled Vocabularies

An initial review was conducted of Open Biomedical Ontologies (OBO) ontologies (18) (Supplementary Table 1), a few ontologies that were not OBO members, as well as several controlled biomedical vocabularies from the literature. From this, a minimum list (Table 1) of ontologies and vocabularies was chosen based on literature review of those containing entities at a level of biological organization applicable to AOPs; as well as characteristics and chosen vocabularies derived from existing ontologies, such as the Beta Cell Genomics project (19) and PubChem RDF project (20). The OBO Foundry ontologies selected were all reference ontologies rather than application ontologies. One controlled vocabulary, Medical Subject Headings (MeSH) (21), was selected due to its widespread use in the medical domain, even though it is not a formal ontology.

Ontologies and controlled vocabularies were both included in the minimum list (Table 1) in order to provide the most representative terms descriptive of AOP-Wiki KEs. Ontologies with infrequently used terms, as low as 1–3 classes across all KEs, were re-evaluated for inclusion in the minimum list. This selection represented the minimum number of ontologies that could accurately represent the biological space of KEs in the AOP-Wiki.

Extension of the Existing AOP- Ontology

Burgoon and colleagues previously created an Adverse Outcome Pathway Ontology (AOPO) (https://github.com/DataSciBurgoon/aop-ontology) (22) to broadly support the development of decision support tools based on AOPs. The core ontology can then be extended to support a variety of uses (23). This work extends the AOPO (Figure 2) as a basis for incorporating the terms from the selected biological ontologies to more fully describe the KEs of an AOP. The extension of this ontology is intended to be in concert with the RDF triple/RDFS/OWL framework for semantic representation of information (24).

This work relies on a new concept called the Event Component (Figure 2), which is based on terms from existing biological ontologies (Table 1). According to this concept, each Event Component was defined via a Process, an Object, and an Action term. The Process represents the dynamics of the underlying biological system (e.g. receptor signaling). The Object was the subject of the perturbation, for example, a specific biological receptor that is activated or inhibited. Ideally, the Process and the Object represent the normal biology that is perturbed as part of the AOP, and not the perturbation itself. Limitations of the existing ontologies made exceptions to this necessary, however. It is important to note that the process and object terminology refer to biological processes and objects and therefore do not necessarily correspond with a continuant/occurrent classification in ontological terms (25).

The Action represents the perturbation of the system described by the other two terms that results in this KE (e.g. ‘decreased’ in the case where a receptor is inhibited to indicate a decrease in the signaling by that receptor). Action terms were based on a short list of terms (Table 2) deemed sufficient to cover the expected cases across current and future AOPs. These terms represented a summary of qualities chosen from the Phenotypic Attribute and Trait Ontology (PATO) ontology (26); for example, “increased,” “decreased,” and “functional change.”

A fourth and separate term, the Context, was incorporated to represent the location or biological environment in which the KE took place. The Context term was tied to the level of biological organization as described by authors in the AOP-Wiki. For KEs with molecular or cellular level of biological organization, the KE was assigned a Cellular or Organ context. For KEs at the tissue or organ level of biological organization, the KE was assigned an Organ context. KEs at the individual or population levels of biological organization were not assigned a Context, as their biological context was represented by the existing "Taxonomy," "Sex," and "Life Stage" of applicability terms in the AOP-Wiki.

Annotation of AOP-Wiki Key Events

As one of the modules making up the AOP-KB, the AOP-Wiki (https://aopwiki.org) supports collaborative development of AOP descriptions in an encyclopedia-style text format, and captures the evidence supporting AOPs. Following selection of the ontologies, KEs from the AOP-Wiki were mapped to classes from the chosen ontologies in Table 1. KE titles and text on individual KE Wiki pages were manually reviewed for text phrases representative of KEs. Query text phrases were entered into three ontology browser sites: Ontobee (27), NCBO Bioportal (28), and the EMBL-EBI Ontology Lookup Service (29).

From the browser's list of output “hits,” ontology classes were reviewed in order to find the appropriate matching class for annotation. Classes were selected based on the appropriateness of the definition in capturing the biology occurring in the KE description, and choice from the ontology tree of the most specific class or target node available that matched the level of detail provided in the KE description.

The resulting annotation of KEs with ontology classes from the minimum list was compiled in Excel. AOP-Wiki authors provided a list of MIEs associated with target genes targeted in the ToxCast effort (30), annotation of these gene targets were provided by Object terms primarily from Protein Ontology (PRO) (31), MeSH (21), and Gene Ontology (GO) (32). Manual curation was conducted initially with two reviewers, followed by consultation with the original authors of all annotated AOPs. Annotated AOPs were delivered as an Excel table to AOP authors for their review and commentary. Following the incorporation of author feedback, annotated KEs were included in the AOP-Wiki on September 9, 2017.

Results

A review of OBO-compliant ontologies (18) was conducted in order to choose a minimum list of ontologies and controlled vocabularies (Table 1) from which to draw ontology terms or “classes”. A full list of the ontologies considered is provided in Supplementary Table 1. OBO Foundry ontologies were chosen as the most broadly accepted and syntactically uniform ontologies available when possible. Ontologies were reviewed based on published literature for their domain of study, their taxonomic applicability, syntax and available format for download, the level of biological organization that they cover from molecular to population, whether the ontology imported or aggregated other data sources, and their state of active development (33). The upper level Relations Ontology (RO) (34) and Basic Formal Ontology (BFO) (25) were excluded from review, as many of their terms are imported and reused in other OBO Foundry member ontologies (35).

For 172 current AOPs in the AOP-Wiki as of December 4, 2016, annotation of KEs with ontology classes (Supplementary Table 2) resulted in a total of 685 KEs being annotated for 165 of the 172 AOPs with a total of 809 Event Components, 651 of which were unique. A total of 61 unique Cellular context terms from Cell Ontology (CL) (36), and 60 Organ context terms from the Uber Anatomy Ontology (Uberon) (37), were used.

Development of the Event Component data model and annotation of existing wiki entries was an iterative process. The result was a set of conventions listed below for annotating KEs via Event Components:

1. Multiple Event Components

   KEs as currently defined based on the existing guidance (38) tend to capture a broader portion of the biological system than can be defined using a single biological process or object. As a result, to fully describe the KE using explicit biological entities, multiple Event Components may be needed with at least one required. This provides more flexibility in describing the KE at early stages of AOP development. Over time, it may be determined that some KEs should be

split while others should be merged. Having formal descriptions of Event Components should assist with this process.

2. Required terms for Event Components

Empirically, most Event Components were best described using the Process-Object-Action syntax as defined in the Event Component model. However, it was noted that in a minority of cases, the Event Component is best described using only the process or object term coupled with an action representing the perturbation. While this limitation could be overcome by extending the existing biological ontologies, such an extension will introduce external dependencies for the annotation process. As a result, the Event Component requires a Process or an Object term but does not require both. Because the Action term represents the perturbation of the system that defines the KE, this is always a required term.

3. Excluded entities

The goal of this project was to annotate the abstract KE concept with explicit biological entities to promote a computational understanding of the system. Because of this, only ontologies that explicitly defined biological objects and processes were considered. Descriptions of the experiments or assays that would be used to measure the KE will be considered separately (23). Chemical entities were only included if they are normally produced endogenously (e.g. estrogen); xenobiotics should not be an intrinsic part of the KE. Keeping the biological description of an AOP distinct from the measurements associated with each KE clearly delineates the measurements used, which may change over time for practical reasons, from the fundamental biology, which shouldn’t change except as a consequence of our expanding knowledge of the system.

4. Specificity

When a search query resulted in multiple “hits,” it was determined that the most specific or “child” ontology class should be selected to accurately represent the KE, in accordance with the accepted best practices for use of ontologies. When reviewing the level of detail recorded in existing AOPs, however, we noted that most of the time the description of the KEs is not as specific as the terms in many of the biological ontologies. For this reason, we provide a caveat to this convention that the term should be the most specific term from the ontology that matches the level of detail of the text-based KE description.

5. Consideration of taxonomic applicability

Efforts were made to match the chosen ontologies based on taxonomic applicability as defined in the KE description. When a KE included multiple taxons, an ontology that was minimally taxon-specific was chosen in order to maximize reuse of Event Component terms across AOPs. In the case of MIEs affecting a target gene, orthology was assumed and the minimally taxon-specific Object protein was chosen, unless the author specified a taxon-specific protein.

6. Action terms

The subset of Action terms selected for representation in the AOP-Wiki were: increased, decreased, morphological change, functional change, occurrence, disrupted, arrested, delayed, premature, abnormal, and pathological. Initially, reviewers used the set of all PATO terms (26) for annotation and mapped these to the smaller subset of Action terms (Table 2). These user-friendly terms were based on what would be more immediately comprehensible for AOP-Wiki users and would typically map to more than one term from PATO.

7. Cell and Organ Context

As many KE descriptions did not specify a biological context in which they occurred, and since Context was blank at the higher levels of organization, the Cell or Organ Context term was optional. Cell Context terms were restricted to those derived from the CL ontology (36) while Organ terms were restricted to those derived from Uberon (37). Molecular or Cellular level events had a Cellular Context or an Organ Context; while events at the Tissue or Organ levels of biological organization had only an Organ Context. In the case of an event occurring at a subcellular level, for example “Mitochondrial dysfunction,” the context was still Cellular but the subcellular location was specified by the Object term. Annotation of “Mitochondrial dysfunction” resulted in the Object term, ‘mitochondrion’ (GO:0005739) and the cell context remained ‘eukaryotic cell’ (CL:0000255). In general, 'eukaryotic cell' (CL:0000255) was chosen to represent a KE at the molecular or cellular level without a specified Cellular context. In some cases, the same KE occurred in multiple AOPs, but with different Cell or Organ contexts. In the future, authors will be encouraged to think about the importance of biological context when defining KEs.

While the conventions listed above were based on the comprehensive annotation of the AOPs in the AOP-Wiki at the beginning of this effort, the following case studies from the AOP-Wiki (39) highlight the rationale behind these conventions.

Case Example 1: Aromatase inhibition leading to reproductive dysfunction

A summary diagram of the annotated AOP is presented in Figure 3 with additional details in Supplementary Table 3; current information on this AOP is available online at https://aopwiki.org/aops/25 (40, 41). Where KEs represented enzyme activation or inhibition, it was determined that ‘aromatase activity’ (GO:0070330) was a more accurate representation than ‘signaling’ (GO:0023052), though both terms could be used to describe the interaction of a ligand with a receptor to result in an intracellular response. According to convention (4), the Process term chosen to best represent the MIE “Aromatase, Inhibition” was ‘aromatase activity’ (GO:0070330), a subclass of ‘catalytic activity’ (GO:0003824). Illustrating convention (*5*), while the taxonomic applicability of this AOP is fish species, the Object term 'cytochrome P450 19A1' (PR_000006100) was chosen to describe, "A protein that is a translation product of the human CYP19A1 gene or a 1:1 ortholog thereof" (31).

The KE "17beta-estradiol synthesis by ovarian granulosa cells, Reduction" illustrates convention (*3*); the Event Component includes the Object term ‘17beta-estradiol’ (CHEBI: 16469) (42). Since 17beta-estradiol is an endogenous chemical entity produced by synthesis

in granulosa cells, it was included in the annotation as the Object of 'estrogen biosynthetic process' (GO:006703). Illustrating convention (*6*), the Action term 'decreased' represents a decreased concentration of circulating 17beta-estradiol.

The KE “Vitellogenin synthesis in liver, Reduction” describes transcription of vitellogenin genes being regulated by estrogens via their action on specific nuclear receptors (40, 41). According to convention (*4*), the most specific descriptive class should be chosen to describe “vitellogenin synthesis,” however, this example illustrates a case in which a higher-order term was chosen in order to better illustrate the author’s intent. Since the author did not specify “transcription” or “translation” when naming this KE, the more generic Process term ‘gene expression’ (GO:0010467) was chosen to describe transcription of genes specific to vitellogenin, and synthesis of the Object protein 'Vitellogenins' (MESH_D014819), according to convention (*4*). In the GO hierarchy, ‘gene expression’ (GO:0010467) is a parental class for ‘transcription, DNA-templated' (GO:0006351) and for ‘translation’ (GO: 0006412), so it includes both of these processes involved in protein synthesis.

For the KE, “Cumulative fecundity and spawning, Reduction,” the Process term ‘egg quantity’ (VT:1000294) (43) was chosen as a reproductive system trait describing fecundity. Illustrating convention (*6*), the Action modifier ‘decreased’ described a reduction in the amount of the ‘egg quantity.’ Illustrating convention (*2*), the Event Component describing the KE does not have a Process term, but is adequately described by the Object term 'egg quantity' combined with the Action term 'decreased.'

For the KE “Plasma vitellogenin concentrations, Reduction,” initially the Process term 'abnormal circulating protein level' (MP:0005416) was considered, but it was determined that the combination of the Object term ‘vitellogenins’ (MESH_D014819) and the Action term ‘decreased’ was sufficient to represent the KE in conjunction with its context without the need for a Process term, illustrating convention (*2*). Similarly, for the KE “Plasma 17beta-estradiol concentrations, Reduction,” initially the Object term ‘abnormal circulating hormone level’ (MP:0005418) was considered. After consideration of this event in multiple AOPs, it was discussed that the KE could be accurately described using the terms ‘17beta-estradiol’ (CHEBI:16469) and ‘decreased’ without the use of the phenotype term from MP. The Organ Context ‘blood plasma’ (UBERON:0001969) described both KEs involving a decreased concentration of hormone or protein in plasma.

Illustrating convention (*1*), the KE "Vitellogenin accumulation into oocytes and oocyte growth/development, Reduction" illustrates that KEs can have multiple Event Components. The first Event Component, consisting of a triple of the terms ‘receptor-mediated endocytosis’ (GO:0006898), ‘vitellogenins’ (MESH_D014819) and ‘decreased’, describes the decreased uptake of vitellogenin from the blood by oocytes via receptor-mediated endocytosis (40, 41). The second Event Component, ‘oocyte growth’ (GO:0001555) and ‘decreased’, describes reduced oocyte growth during vitellogenesis. The third component, ‘oocyte development’ (GO:0048599) and ‘decreased’, describes reduced oocyte development. This cellular level KE occurs in the ‘oocyte’ (CL:0000023), so the Context was repeated for all three Event Components.

Illustrating convention (*7*), for the AO, “Population trajectory, decrease” no Cellular or Organ context is defined because this is a population-level event and the Context is fully captured by the Life Stage/Sex/Taxonomic Applicability fields. The Action term ‘Decreased’ modifies ‘population growth rate’ (PCO_0000008), describing a decreased rate of population growth.

Case Example 2: Chronic binding of antagonist to N-methyl-D-aspartate receptors (NMDARs) during brain development induces impairment of learning and memory abilities

A summary diagram of the annotated AOP is presented in Figure 4 (Supplementary Table 3); current information on this AOP is available online at https://aopwiki.org/aops/13 (44, 45).

Annotation of this AOP highlights multiple points for consideration. First, the MIE "NMDARs, Binding of Antagonist" and the first KE "Inhibition of NMDARs" initially resulted in identical Process terms, with ‘signaling’ (GO:0023052) encompassing both of the text terms "inhibition" and "binding." Although Event Components within an AOP may have shared Process terms, identical Event Components shouldn't occur for two different KEs within a single AOP. Discussion with AOP authors resulted in a changed annotation for these KEs, and the resultant Process terms ‘NMDA glutamate receptor activity’ (GO:0004972) and ‘signaling’(GO:0023052) for the MIE and first KE, respectively. These annotations were completed by the search of Object terms able to adequately describe the event Processes. A search query for 'NMDA receptors' retrieved several results from various ontologies: ‘NMDA selective glutamate receptor complex’ (GO:0017146), 'Receptors, N-Methyl-D-Aspartate' (MeSH_D016194), and ‘glutamate receptor ionotropic, NMDA 1’ (PR: 000008245). In this case, the author did not specify a particular subunit, so the term descriptive of the complex (GO:0017146) was preferred. Where possible, GO terms were preferred to MeSH terms because GO is an ontology, while MeSH is a controlled vocabulary, so ‘NMDA selective glutamate receptor complex’ (GO:0017146) was preferred to 'Receptors, N-Methyl-D-Aspartate' (MeSH_D016194). Finally, in accordance with convention (*4*), the term ‘NMDA selective glutamate receptor complex’ (GO:0017146) was selected for the Object term of the MIE as well as for the Object of the first KE. Illustrating convention (*6*), for the MIE, the action term ‘arrested’ was selected to modify the Process term ‘NMDA glutamate receptor activity’ (GO:0004972), describing arrested receptor activity. For the first KE, the action term ‘arrested’ was also chosen to modify ‘signaling’ (GO:0023052).

Secondly, this AOP brings up a discussion regarding the selection of terms for inter- and intracellular transport. GO defines ‘transport’ (GO:0006816) as "The directed movement of substances into, out of or within a cell, or between cells, or within a multicellular organism by means of some agent such as a transporter or pore." Illustrating convention (*4*), ‘calcium ion transport’ is a subclass of ‘ion transport,’ and both are children of the parent class ‘transport.’ In order to ensure specificity of this Process term, the more specific class ‘calcium ion transport’ was chosen along with the Object term ‘calcium ion’ (CHEBI: 39124) from ChEBI.

For the KE "Release of BDNF, Reduced," the text describes the transcription and release of BDNF from glutamatergic neurons. This was another instance in which an exception was

made in order to more accurately capture the AOP author's intent. According to convention (*4*), the most specific descriptive class should be 'brain-derived neurotrophic factor receptor activated activity' (GO:0060175); however, this example illustrates a case in which a higher-order term, 'gene expression' (GO:0010467) was chosen. 'Gene expression' (GO:0010467) is a parent class for 'transcription, DNA-templated' (GO:0006351) and for 'translation' (GO:0006412), so the subsuming class was chosen. Illustrating convention (*1*), a second Event Component used the term 'secretion' (GO:0046903) as the Process term to describe release of the protein. The KE "Dendritic morphology, Aberrant" illustrates the choice of Action terms following convention (*6*), and 'abnormal' was the chosen modifier for the Process term 'dendrite morphogenesis' (GO:0048813). 'Dendrite' (GO:0030425) was the chosen Object of the morphological change.

The annotation of KEs relating to synaptic function highlights convention (*4*). For the KE "Presynaptic release of glutamate, Reduced," it was determined that 'glutamate secretion, neurotransmission' (GO:0061535) was the most applicable process. In GO, 'glutamate secretion, neurotransmission' is a subclass of 'Neurotransmitter secretion' (GO:0010554), and it was appropriate to specify neurotransmitters in order to include 'Glutamate' (CHEBI: 14321) as the Object term, as defined by convention (*3*). For the KE "Synaptogenesis, Decreased," 'synapse assembly' (GO:0007416), which is a subclass of 'cellular component assembly' (GO:0022607), was chosen as the Process from a query for "synaptogenesis." Conversely, in the case of the event "Neuronal network function, Decreased," the query term "neuronal network communication" did not have an exact match in GO, but from the KE description the Process term was determined to be 'synaptic signaling' (GO:0099536).

The AO, "Learning and memory, Impairment" was initially annotated with the Process term 'learning or memory' (GO:0007611), defined as "The acquisition and processing of information and/or the storage and retrieval of this information over time." As 'learning or memory' (GO:0007611) is a higher-order term, with subclasses 'learning' (GO:0007612) and 'memory' (GO:0007613), two Event Components were constructed in order to specify these two processes individually according to convention (*4*).

Case Example 3: PPAR-alpha activation in utero leading to impaired fertility in males

Peroxisome Proliferator Activated Receptor α (PPARα) is a ligand-activated transcription factor which belongs to the family of nuclear receptors and acts as key regulator of lipid metabolism. The linkage between the activation of PPARα and the impairment of steroidogenesis leading to effects on reproduction is supported by evidence in both rodent and human studies and described in an AOP currently under review (https://aopwiki.org/aops/18) (46, 47). A summary diagram of the annotated AOP is presented in Figure 5 (Supplementary Table 3).

The Process underlying the MIE, "PPARα, activation," was initially annotated with the term 'signaling' (GO:0023052) defined as, "The entirety of a process in which information is transmitted within a biological system. This process begins with an active signal and ends when a cellular response has been triggered." In order to describe the underlying molecular process in its entirety, from the ligand-mediated activation of the transcription factor to its following translocation to the nucleus, the class 'peroxisome proliferator activated receptor

signaling pathway' (GO:0035357), a subclass of 'signal transduction' (GO:0023033) was chosen in accordance with convention (*4*). The directionality of the Process is defined by the Action term 'increased,' as illustrated by convention (*6*) and the Cell Context term 'eukaryotic cell' (CL:0000255) was used to represent the wide range of PPARα expression (46, 47).

For the KE, "Testosterone level, Reduction," the text describes the reduction of circulating testosterone in the bloodstream (46, 47). As a whole, annotation of this KE combines different conventions. At first, the Object term 'abnormal circulating hormone level' (MP: 0005418) was chosen to capture the imbalance of hormone levels in the blood. After considering similar events across multiple AOPs, it was decided that annotation of this KE with the Object term 'testosterone'(CHEBI:17347) and the Action term 'decreased' was sufficient to describe reduced levels of circulating testosterone. Illustrating convention (*2*), the Process term of this KE was left blank. The Organ Context, 'blood' (UBERON: 0000178) represents the biological context in which this KE occurs, illustrating convention (*7*).

The revision of this annotation following discussion with authors showcases that there are multiple ways that a KE could be described using terms from the ontologies in the minimum list.

The two AOs were annotated without a Process term, as shown by convention (*2*). The AO "Male reproductive tract, Malformation," was described by using 'male reproductive organ' (UBERON:0003135) as the Object term together with the Action modifier 'morphological change'. For this AO, which is at the Organ level of organization, the Organ Context was specified as indicated by convention (*7*) by 'Male reproductive system' (UBERON: 0000079). To represent the AO "Fertility, impaired," the term 'fertility' (MeSH_D005298) was selected as the Object along with the Action 'decreased.' In this case, since the AO is at the individual level of biological organization, the context is not captured through ontology terms but is captured elsewhere in the AOP by the sex, taxonomic, and life stage applicability fields.

Case Example 4: Inhibition of Thyroperoxidase and Subsequent Adverse Neurodevelopmental Outcomes in Mammals

A summary diagram of the annotated AOP is presented in Figure 6 (Supplementary Table 3); current information on this AOP is available online at https://aopwiki.org/aops/42 (48). The MIE, "Thyroperoxidase, Inhibition," is described by the Process term 'iodide peroxidase activity' (GO:0004447), Object term 'thyroid peroxidase' (PR:000016584), and the Action term 'decreased.' Selection of 'thyroid peroxidase' (PR:000016584) illustrates convention (*5*), as taxonomic applicability for the MIE includes multiple taxons and species orthology was assumed in the choice of terms. 'Thyroid peroxidase' (PR:000016584) is defined as "A protein that is a translation product of the human TPO gene or a 1:1 ortholog thereof" (31). 'Thyroid follicle,' (UBERON:0005305) represents the Organ context in which catalysis of thyroperoxidase was occurring, according to convention (*7*).

The KE "Thyroid hormone Synthesis, Decreased," initiated a discussion about the Process term. A query of GO terms resulted in two possible options, 'hormone metabolic process' (GO:0042445) and its subclass, 'thyroid hormone generation' (GO:0006590). Ultimately, 'thyroid hormone generation' (GO:0006590) was chosen in order to specify the formation of thyroid hormone in accordance with convention (4).

The next KEs, "Thyroxine (T4) in serum, decreased," and "Thyroxine (T4) in neuronal tissue, decreased," were the subject of further discussion with authors. Initially, this KE was described using the MP terms, 'abnormal circulating hormone level' (MP:0005418), and 'abnormal hormone level' (MP:0003953) for Process. However, the authors didn't think the use of phenotypes to describe this KE was intuitive, and suggested deprecation of MP terms. The Object term 'thyroxine' (CHEBI:30660) and the Action term 'decreased' were sufficient to describe the KE in both cases, without the need for a Process term, following convention (*2*).

Secondly, this KE illustrates convention (*7*); the Organ Context terms, 'serum' (UBERON: 0001977), and 'brain' (UBERON:0000955), specify the location in which these KEs occur.

The following KE, "Hippocampal Gene Expression, Altered," illustrates the point brought up in Case Example 1 regarding the description of gene expression and protein translation. In this situation, the Process term, 'regulation of gene expression' (GO:0010468), defined as "Any process that modulates the frequency, rate or extent of gene expression," was used to specify upregulation of genes in the developing brain. The Object term 'hippocampal formation' (UBERON:0002421) was used to denote the object of gene expression.

The KE, "Hippocampal Anatomy, Altered," illustrates convention (*6*). The Action term, 'morphological change,' is used to describe a structural change in the hippocampus. In the following KE, "Hippocampal Function, Decreased," the Action term 'abnormal' denotes altered synaptic function in the hippocampus.

The AO, "Cognitive Function, Decreased," illustrates convention (*1*), having two Event Components in order to accurately portray changes in both Processes 'learning or memory' (GO:0007611) and 'cognition' (GO:0050890) as portrayed by the KE description in the AOP-Wiki. Unlike annotation of the KE "Learning and memory, Impairment" in Case Example 2, the more generic term 'learning or memory' (GO:0007611) was chosen, since the author didn't specify 'learning' (GO:0007612) or 'memory' (GO:0007613). Since the AO is at the Individual level of biological organization, it did not have a Context term, illustrating convention (*7*).

Case Example 5: Protein Alkylation leading to Liver Fibrosis

The development of liver fibrosis is the result of a complex process involving many hepatic cell types, receptors and signaling pathways (49, 50). A summary illustration of the annotated AOP is presented in Figure 7 (Supplementary Table 3); further information on this AOP to date is available online at https://aopwiki.org/aops/38.

Annotation of the KE, "Cell injury/death," illustrates convention (*2*). The term "cell death" (GO:0008219) was selected for Process in the absence of an exact match for the query, "cell

injury," and also in accordance with the KE text description, which focuses on mechanisms of apoptosis and necrosis (49, 50). The context is described by 'eukaryotic cell' (CL: 0000255), according to convention (*7*).

According to convention (*4*), for the KE, “TGFbeta1 expression, Up Regulation,” the Process, ‘transforming growth factor beta1 production’ (GO:0032905), was chosen to represent the expression of TGFbeta1. Unlike in Case Example 2, which had two Event Components to describe the separate processes “gene expression” and “secretion” according to convention (*1*) one Event Component was adequate to describe this KE.

The KEs, "Activation and Recruitment of Hepatic macrophages (Kupffer Cells), Increased," and, "Stellate cells, Activation," both describe cellular activation. According to convention (*4*), they were annotated similarly using the Process terms, ‘macrophage activation’ (GO: 0042116) and ‘hepatic stellate cell activation’ (GO:0035733), both sub-classes of 'cell activation' (GO:0001775), defined as, "A change in the morphology or behavior of a cell resulting from exposure to an activating factor such as a cellular or soluble ligand" (32). "Kupffer cell" (CL:0000091) and, "hepatic stellate cell" (CL:0000632), respectively, were designated as both the Object term and the cellular context for these two events.

To describe the development of the AO, “Liver fibrosis,” 'liver fibrosis' (MP:0003333) was chosen as the Process term with the Object term ‘liver’ (UBERON:0002107). This annotation was completed by the Action term, ‘occurrence’, and the Context term, ‘liver’ (UBERON:0002107) according to convention (*7*), since this AO is an organ level event. In this case our object term matches the context term because the object is in fact the organ. Both terms are important, however, because we want to support queries/reasoners that are focused on one term or the other.

Discussion

Hundreds of ontologies and controlled vocabularies representing disparate domains are publicly available on the web. A comprehensive review (Supplementary Table 1) of them to inform the AOP framework at various levels of biological organization led to the selection of a minimum list (Table 1) of ontologies and vocabularies, from which individual terms (classes) were queried. The minimum list was dynamic, as ontologies that were not used were discarded and ontologies that were needed were added. Ontologies such as Population and Community Ontology (PCO) (51), and Protein-protein interaction (MI) (52) that were minimally used in the mapping, will be reconsidered in future to see if identical classes from other ontologies can adequately describe these KEs. In this way, the number of ontologies in the minimum list would be reduced, making parsing of terms more efficient. It is also likely that future AOPs may fall outside the scope of the minimum list, requiring the addition of new ontologies to represent these domains. Further discussion with authors brought up the need to consider including ontologies such as Cell Behavior Ontology (53), representing existential behaviors of cells, and eNanoMapper (54) ontology, representing common vocabulary terms used in nanosafety research, in the minimum list.

The triplicate Event Component model consisting of a Process, Object, and an Action; and the Context entity for mapping KEs to ontology classes was effective for the majority of AOPs examined. However, there were a few KEs for which a suitable set of terms could not be found in the minimum list. For example, for the KE, "Skin, sensitization" (55), the annotation curators and the AOP authors were not able to come to a consensus, as no representative Process term was found in the minimum list of ontologies. AOP authors suggested the term, 'skin sensitization' (ENM:0000034), from the eNanoMapper ontology (54). The focus for this initial round of annotation was to describe normal biological processes and their perturbation by exogenous stressors. Still to be considered is whether a separate effort should be made to explicitly define the aberrant phenotype using ontological terms, or whether these should be included in the biological description of the KE. There are several ontologies that would be suitable for this purpose.

The minimum list adequately characterized the biological space currently represented by KEs. In this annotation, the most representative ontology for the Process term was GO, representing the majority of Process terms at all levels of organization, but especially at the molecular and cellular levels. The most representative ontologies for Object terms were ChEBI and PR, representing chemical entities and proteins respectively.

The subset of eleven action terms (Table 2) was chosen based on the definition of the Action entity to provide the most comprehensible term selection for AOP-Wiki users, and can be expanded in future if needed. The original set of PATO classes chosen describes qualities (26), and it was thought that these terms would be confusing to users. In addition, using a short list of well-defined terms should increase the consistency of term usage by different authors, which is extremely important when using these objects for computational reasoning applications.

As KEs were annotated, the need became apparent for a set of conventions to define their annotation. Case examples showcased the application of these conventions, and other specific issues that came up during manual curation. One concern brought up in discussion with experts was the need to ensure that these conventions were defined in such a way that two people mapping KE terms to ontology classes are most likely to arrive at an identical annotation. Convention (*4*) was implemented in order to clarify the choice of terms based on specificity and compliance with semantic principles. As described in the Case Examples, there were a few instances where the annotation curators were not able to choose the most specific term possible because author-provided descriptions of KEs are usually in more general terms.

The annotated KEs can easily be translated later onto post-composed ontology classes using commonly adopted Entity-Quality (EQ) statement syntax (56). Based on the EQ model, Web Ontology Language (OWL) has been used for the formalization of phenotype descriptions, as a means to represent complex phenotypes by a combination of a quality and one or more entities (57). From there, conversion to post-composed ontology classes should be straightforward. MeSH was an exception; as a controlled vocabulary, it does not subscribe to OWL semantics (57), but it was included in the minimum list in order to represent disease-related terminology that wasn’t otherwise available from ontologies in the minimum list. For

the purposes of the Wiki, MeSH terms were used interchangeably with ontologies; however, consideration would need to be given were MeSH to be integrated in an extension of the original AOP-Ontology (23). In the future, if the set of ontologies chosen for this project are to be integrated into a unique ontology, and for better integration with the existing AOPO (23), the Ontology for General Medical Science (OGMS) (58) and other ontologies such as SNOMED CT (59) and ICD-11 (60) should be considered both as replacements or extensions to our biological descriptions and as more formal descriptions of the associated measurements.

The Event Component information described herein was incorporated in the AOP-Wiki on September 9, 2017 (39). In the future, authors will be encouraged to annotate new AOPs using the best practices defined here and the controlled terminology based on terms from the selected ontologies. As bioinformatics approaches supporting the AOP development process mature (14, 61), the inclusion of explicit ontology terms in the description of KEs should become automatic. The Event Component model has direct application in informing users of capabilities in the AOP-Wiki through improving query of existing KEs, and by providing more flexibility for authors in creating new KEs. By separating Event Component entities into the Process, Object, and Action terms for use in searching, authors can create a new KE when a very similar one exists to account for cases where slight differences in the perturbation of the same part of the biological system can result in disparate AOPs. Currently, the AOP-Wiki contains ontology terms from the minimum list (Table 1), allowing AOP authors to choose Event Component entities for annotation of newly created KEs.

In the broader context of the needs of the AOP-KB, O-AOP-KB provides several advantages over text-only descriptions of KEs. The ability to explicitly attach scientific evidence in support of the AOP, including assays and biomarkers linked to KEs, allows for the development of tools that integrate and display these data in the context of an AOP. As this project progresses, ontologies and controlled vocabularies representing the domains of experiments and ecology will perceivably be added in order to better describe measurements and endpoints. As high-throughput genomics becomes a major source of high-throughput toxicity testing results, this factor becomes increasingly important. For instance, defining each KE in terms of a set of Event Components allows many specific genes (mapped to the objects in our annotation scheme) to be associated with a single KE. This is a critical step in mapping genomics and proteomics information onto an AOP that is defined at a much more abstract level.

The greatest effort was made to match the author’s intent as revealed in the KE description (39), with the intention of providing associations useful for other AOP-KB modules and external applications in the future. For example, an associated AOP-Database application (AOP-DB) could use the Event Component model to provide support for relating AOPs to genes, chemicals, diseases, pathways, species orthology, and gene interaction information, via inferred associations (62). In this way, O-AOP-KB will support aggregation of data from high throughput screening assays and experiments to inform KEs. O-AOP-KB also reinforces AOP networks, which can then be used for further AOP identification and evaluation (14). By identifying explicit ontology classes as part of KEs, this greater

resolution brings in the full power of ontology-based semantic analysis to discover implicit connections among KEs and AOPs.

The nature of ontology enables information sharing and integration across species and knowledge domains computationally, which supports the evaluation of taxonomic relevance of KEs (14). Nodes (classes) in an ontology graph are both semantically and logically informative based on their positions in the hierarchy. The KEs and AOPs from various taxa, when annotated with ontology terms, can be semantically compared across species for similarities. The resultant KEs that are highly similar should be considered for consolidation. The AOPs that are highly similar, on the other hand, indicate common biology among the species involved and a wider range of applicability for such AOPs. O-AOP-KB also facilitates qAOP development backed by computational systems models (61, 63) by referencing explicit biological entities suitable for developing computational models of systems described by AOPs or AOP networks. This approach can speed up the development of qAOPs by providing explicit biological entities underlying the KEs and a better understanding of how AOPs interact with one another to form AOP networks.

In summary, application of the Event Component model to case studies of existing high-priority AOPs within the AOP-Wiki demonstrates its applicability to these use cases. Shared biological, chemical, and toxicological ontologies will be useful not only to the AOP-Wiki, but across all modules of the AOP-KB, in promoting the development of AOP-based computational decision-support tools and knowledge interoperability as a whole (64). Future directions include the possible incorporation of explicit regulatory endpoints and mapping of AOs with terminology from OECD harmonized templates (65).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The authors would like to thank the AOP developers in the OECD AOP Development Programme for their patience in translating the information they have entered into the AOP-Wiki into a formal description using ontologies. We also thank the members of the OECD EAGMST Ontology Working Group for guidance with ontology selection and feedback on the extensions made to the AOP ontology. We thank Lyle Burgoon for helpful discussions as we extended the AOP ontology. Finally, we thank Nancy Baker and David Lyons for critical review of this manuscript.

The information in this document has been funded wholly (or in part) by the U. S. Environmental Protection Agency. It has been subjected to review by the National Health and Environmental Effects Research Laboratory and approved for publication. Approval does not signify that the contents reflect the views of the Agency, nor does mention of trade names or commercial products constitute endorsement or recommendation for use. The contents of this paper are the views of the authors and do not necessarily represent the views and policies of the European Commission.

References

1. Krewski D, Acosta D Jr, Andersen M, Anderson H, Bailar JC 3rd, Boekelheide K, et al. Toxicity testing in the 21st century: a vision and a strategy. Journal of toxicology and environmental health Part B, Critical reviews. 2010; 13(2–4):51–138. Epub 2010/06/25. PubMed PMID: 20574894; PubMed Central PMCID: PMCPmc4410863. DOI: 10.1080/10937404.2010.483176
2. National Research CouncilToxicity Testing in the 21st Century: A Vision and a StrategyWashington, DC: The National Academies Press; 2007216

3. Ankley GT, Bennett RS, Erickson RJ, Hoff DJ, Hornung MW, Johnson RD, et al. Adverse outcome pathways: A conceptual framework to support ecotoxicology research and risk assessment. Environmental Toxicology and Chemistry. 2010; 29(3):730–41. DOI: 10.1002/etc.34 [PubMed: 20821501]

4. Groh KJ, Carvalho RN, Chipman JK, Denslow ND, Halder M, Murphy CA, et al. Development and application of the adverse outcome pathway framework for understanding and predicting chronic toxicity: I. Challenges and research needs in ecotoxicology. Chemosphere. 2015; 120:764–77. Epub 2014/12/03. PubMed PMID: 25439131. DOI: 10.1016/j.chemosphere.2014.09.068 [PubMed: 25439131]

5. Villeneuve DL, Crump D, Garcia-Reyero N, Hecker M, Hutchinson TH, LaLone CA, et al. Adverse Outcome Pathway (AOP) Development I: Strategies and Principles. Toxicological Sciences. 2014a; 142(2):312–20. DOI: 10.1093/toxsci/kfu199 [PubMed: 25466378]

6. Villeneuve DL, Crump D, Garcia-Reyero N, Hecker M, Hutchinson TH, LaLone CA, et al. Adverse Outcome Pathway Development II: Best Practices. Toxicological Sciences. 2014b; 142(2):321–30. DOI: 10.1093/toxsci/kfu200 [PubMed: 25466379]

7. Tollefsen KE, Scholz S, Cronin MT, Edwards SW, de Knecht J, Crofton K, et al. Applying Adverse Outcome Pathways (AOPs) to support Integrated Approaches to Testing and Assessment (IATA). Regulatory toxicology and pharmacology : RTP. 2014; 70(3):629–40. Epub 2014/09/28. PubMed PMID: 25261300. DOI: 10.1016/j.yrtph.2014.09.009 [PubMed: 25261300]

8. OECDGuidance Document on Developing and Assessing Adverse Outcome PathwaysParis: OECD; 20136ENV/JM/MONO(2013)

9. OECDThe Guidance Document for Using the OECD (Q)SAR Application Toolbox to Develop Chemical Categories According to the OECD Guidance on Grouping ChemicalsParis: OECD Publishing; 2014

10. Berggren E, Amcoff P, Benigni R, Blackburn K, Carney E, Cronin M, et al. Chemical Safety Assessment Using Read-Across: Assessing the Use of Novel Testing Methods to Strengthen the Evidence Base for Decision Making. Environmental health perspectives. 2015; 123(12):1232–40. Epub 05/09. 10.1289/ehp.1409342. Epub 2015 May 8. PubMed PMID: 25956009. DOI: 10.1289/ehp.1409342 [PubMed: 25956009]

11. Patlewicz G, Ball N, Boogaard PJ, Becker RA, Hubesch B. Building scientific confidence in the development and evaluation of read-across. Regulatory toxicology and pharmacology : RTP. 2015; 72(1):117–33. Epub 2015/04/11. PubMed PMID: 25857293. DOI: 10.1016/j.yrtph.2015.03.015 [PubMed: 25857293]

12. Vinken M. Adverse Outcome Pathways as Tools to Assess Drug-Induced Toxicity. Methods in molecular biology (Clifton, NJ). 2016; 1425:325–37. Epub 2016/06/18. PubMed PMID: 27311472; PubMed Central PMCID: PMCPmc5436615. DOI: 10.1007/978-1-4939-3609-0_14

13. Perkins EJ, Antczak P, Burgoon L, Falciani F, Garcia-Reyero N, Gutsell S, et al. Adverse Outcome Pathways for Regulatory Applications: Examination of Four Case Studies With Different Degrees of Completeness and Scientific Confidence. Toxicological sciences : an official journal of the Society of Toxicology. 2015; 148(1):14–25. Epub 2015/10/27. PubMed PMID: 26500288. DOI: 10.1093/toxsci/kfv181 [PubMed: 26500288]

14. Oki NO, Nelms MD, Bell SM, Mortensen HM, Edwards SW. Accelerating Adverse Outcome Pathway Development Using Publicly Available Data Sources. Current Environmental Health Reports. 2016; :1–11. DOI: 10.1007/s40572-016-0079-y [PubMed: 26875182]

15. Bard JB, Rhee SY. Ontologies in biology: design, applications and future challenges. Nat Rev Genet. 2004; 5(3):213–22. PubMed PMID: 14970823. DOI: 10.1038/nrg1295 [PubMed: 14970823]

16. McMurry JA, Kohler S, Washington NL, Balhoff JP, Borromeo C, Brush M, et al. Navigating the Phenotype Frontier: The Monarch Initiative. Genetics. 2016; 203(4):1491–5. PubMed PMID: 27516611; PubMed Central PMCID: PMC4981258. DOI: 10.1534/genetics.116.188870 [PubMed: 27516611]

17. Mungall CJ, McMurry JA, Kohler S, Balhoff JP, Borromeo C, Brush M, et al. The Monarch Initiative: an integrative data and analytic platform connecting phenotypes to genotypes across species. Nucleic acids research. 2017; 45(D1):D712–D22. PubMed PMID: 27899636; PubMed Central PMCID: PMC5210586. DOI: 10.1093/nar/gkw1128 [PubMed: 27899636]

18. Smith B, Ashburner M, Rosse C, Bard J, Bug W, Ceusters W. The OBO Foundry: coordinated evolution of ontologies to support biomedical data integration. Nature biotechnology. 2007; 25doi: 10.1038/nbt1346
19. ZhengJ, , ManduchiE, , StoeckertC. Development of an Application Ontology for Beta Cell Genomics Based on the Ontology for Biomedical Investigations; International Conference on Biomedical Ontologies (ICBO 2013); Montreal, Canada. 201316
20. Fu G, Batchelor C, Dumontier M, Hastings J, Willighagen E, Bolton E. PubChemRDF: towards the semantic annotation of PubChem compound and substance databases. Journal of cheminformatics. 2015; 7:34. PubMed PMID: 26175801; PubMed Central PMCID: PMC4500850. doi: 10.1186/s13321-015-0084-4 [PubMed: 26175801]
21. NelsonSJ, editorMedical Terminologies That Work: The Example of MeSH; Proceedings of the 10th International Symposium on Pervasive Systems, Algorithms, and Networks (ISPAN 2009); Dec. 14–16, 2009; Kaohsiung, Taiwan.
22. Adverse Outcome Pathway Ontology [Internet]. [cited Aug 19, 2017] GitHub2017Available from: https://github.com/DataSciBurgoon/aop-ontology/blob/master/aopo2.owl
23. Burgoon LD. The AOPOntology: A Semantic Artificial Intelligence Tool for Predictive Toxicology. Applied In Vitro Toxicology. 2017; 3(Current Issue)
24. RobinsonPN, , BauerS. Introduction to Bio-ontologiesHallC, editorBoca Raton, FL: Taylor & Francis Group, LLC; 2011
25. SmithB, editorProceedings of the Third Interdisciplinary Ontology Meeting; 2012Tokyo. Tokyo: Keio University Press; 2012On Classifying Material Entities in Basic Formal Ontology.
26. Phenotypic quality [Internet]. [cited July 10, 2017] OBO Technical WG2017Available from: http://purl.obolibrary.org/obo/pato/releases/2017-07-10/pato.owl
27. Ong E, Xiang Z, Zhao B, Liu Y, Lin Y, Zheng J, et al. Ontobee: A linked ontology data server to support ontology term dereferencing, linkage, query and integration. Nucleic acids research. 2017; 45(D1):D347–D52. DOI: 10.1093/nar/gkw918 [PubMed: 27733503]
28. Whetzel PL, Noy NF, Shah NH, Alexander PR, Nyulas C, Tudorache T, et al. BioPortal: enhanced functionality via new Web services from the National Center for Biomedical Ontology to access and use ontologies in software applications. Nucleic acids research. 2011; 39(Web Server issue):W541–5. Epub 2011/06/16. PubMed PMID: 21672956; PubMed Central PMCID: PMCPmc3125807. DOI: 10.1093/nar/gkr469 [PubMed: 21672956]
29. JuppS, editorSWAT4LS International ConferenceCambridge, England: 2015A new Ontology Lookup Service at EMBL-EBI. [Dec 7–10, 2015]
30. USEPAToxicity Forecaster (ToxCast)Research Triangle Park, NC; 2013Available from: http://www.epa.gov/sites/production/files/2013-12/documents/toxcast-fact-sheet.pdf
31. Natale DA, Arighi CN, Blake JA, Bult CJ, Christie KR, Cowart J, et al. Protein Ontology: a controlled structured network of protein entities. Nucleic acids research. 2014; 42(Database issue):D415–21. Epub 2013/11/26. PubMed PMID: 24270789; PubMed Central PMCID: PMCPmc3964965. DOI: 10.1093/nar/gkt1173 [PubMed: 24270789]
32. GOC. Gene Ontology Consortium: going forward. Nucleic acids research. 2015; 43(Database issue):D1049–56. Epub 2014/11/28. PubMed PMID: 25428369; PubMed Central PMCID: PMCPMC4383973. DOI: 10.1093/nar/gku1179 [PubMed: 25428369]
33. Malone J, Stevens R, Jupp S, Hancocks T, Parkinson H, Brooksbank C. Ten Simple Rules for Selecting a Bio-ontology. PLoS computational biology. 2016; 12(2):e1004743. PubMed PMID: 26867217; PubMed Central PMCID: PMC4750991. doi: 10.1371/journal.pcbi.1004743 [PubMed: 26867217]
34. Relations Ontology [Internet]. [cited July 19, 2017] OBO Technical WG2017Available from: http://purl.obolibrary.org/obo/ro/releases/2017-07-19/ro.owl
35. Brinkman RR, Courtot M, Derom D, Fostel JM, He Y, Lord P. Modeling biomedical experimental processes with OBI. Journal of biomedical semantics. 2010; 1doi: 10.1186/2041-1480-1-s1-s7
36. Cell Ontology [Internet]. [cited July 29, 2017] OBO Technical WG2017Available from: http://purl.obolibrary.org/obo/cl/releases/2017-07-29/cl.owl
37. Mungall CJ, Torniai C, Gkoutos GV, Lewis SE, Haendel MA. Uberon, an integrative multi-species anatomy ontology. Genome Biol. 2012; 13doi: 10.1186/gb-2012-13-1-r5

38. OECDUsers' Handbook supplement to the Guidance Document for developing and assessing Adverse Outcome PathwaysParis: OECD Publishing; 2016
39. [updated Mar 07, 2017] Adverse Outcome Pathway Wiki2016Available from: https://aopwiki.org/
40. VilleneuveD. [updated April 18, 2017] Aromatase inhibition leading to reproductive dysfunctionAvailable from: https://aopwiki.org/aops/25
41. VilleneuveD. Adverse Outcome Pathway on Aromatase Inhibition Leading to Reproductive Dysfunction (in Fish)Paris: OECD Publishing; 2016
42. Hastings J, de Matos P, Dekker A, Ennis M, Harsha B, Kale N, et al. The ChEBI reference database and ontology for biologically relevant chemistry: enhancements for 2013. Nucleic acids research. 2013; 41doi: 10.1093/nar/gks1146
43. Vertebrate trait [Internet]. [cited May 05, 2017] OBO Technical WG2017Available from: http://purl.obolibrary.org/obo/vt.owl
44. SachanaM, , MunnS, , Bal-PriceA. [updated May 23, 2017; May 05, 2017] Chronic binding of antagonist to N-methyl-D-aspartate receptors (NMDARs) during brain development induces impairment of learning and memory abilities2016Available from: https://aopwiki.org/wiki/index.php/Aop:13
45. SachanaM, , MunnS, , Bal-PriceA. Adverse Outcome Pathway on chronic binding of antagonist to N-methyl-D-aspartate receptors (NMDARs) during brain development induces impairment of learning and memory abilitiesParis: OECD Publishing;
46. NepelskaM, , GrignardE, , MunnS. [updated May 20, 2017May 05, 2017] PPARalpha activation in utero leading to imparied fertility in malesAvailable from: https://aopwiki.org/wiki/index.php/Aop:18
47. Nepelska M, Odum J, Munn S. Adverse Outcome Pathway (AOP): PPARα activation and reproductive toxicity. Development and application in assessment of endocrine disruptors/reproductive toxicants. Applied In vitro Toxicology. 2017 (accepted for publication April 2017) (Adverse outcome pathways as versatile tools in in vitro and in silico toxicology.).
48. CroftonKM, , GilbertMG, , FriedmanKP, , DemeneixB, , MartyMS, , ZoellerRT. [updated June 06, 2017;May 05, 2017] Inhibition of Thyroperoxidase and Subsequent Adverse Neurodevelopmental Outcomes in MammalsAvailable from: https://aopwiki.org/aops/42
49. LandesmannB. Adverse Outcome Pathway on Protein Alkylation Leading to Liver FibrosisParis: OECD Publishing; 2016
50. LandesmannB. [updated May 23, 2017;May 05, 2017] Protein Alkylation leading to Liver FibrosisAvailable from: https://aopwiki.org/aops/38
51. Population and Community Ontology [Internet]. [cited Aug 16, 2017] OBO Technical WGAvailable from: http://purl.obolibrary.org/obo/pco/releases/2017-08-16/pco.owl
52. HUPO-PSI/psi-mi-CV [Internet]. [cited May 07, 2017] OBO Technical WG2017Available from: http://purl.obolibrary.org/obo/mi.owl
53. Sluka JP, Shirinifard A, Swat M, Cosmanescu A, Heiland RW, Glazier JA. The cell behavior ontology: describing the intrinsic biological behaviors of real and model cells seen as active agents. Bioinformatics (Oxford, England). 2014; 30(16):2367–74. Epub 2014/04/24. PubMed PMID: 24755304; PubMed Central PMCID: PMCPMC4133580. DOI: 10.1093/bioinformatics/btu210
54. Hastings J, Jeliazkova N, Owen G, Tsiliki G, Munteanu CR, Steinbeck C, et al. eNanoMapper: harnessing ontologies to enable data integration for nanomaterial risk assessment. Journal of biomedical semantics. 2015; 6(1):10.doi: 10.1186/s13326-015-0005-5 [PubMed: 25815161]
55. LandesmannB. [updated Dec 03, 2016;May 05, 2017] Covalent Protein Binding Leading to Skin Sensitisation2016Available from: https://aopwiki.org/wiki/index.php/Aop:40
56. Mungall C, Gkoutos G, Smith C, Haendel M, Lewis S, Ashburner M. Integrating phenotype ontologies across multiple species. Genome Biol. 2010; 11doi: 10.1186/gb-2010-11-1-r2
57. Loebe F, Stumpf F, Hoehndorf R, Herre H. Towards improving phenotype representation in OWL. Journal of biomedical semantics. 2012; 3(Suppl 2):S5. PubMed PMID: 23046625; PubMed Central PMCID: PMC3448528. doi: 10.1186/2041-1480-3-S2-S5

58. Scheuermann RH, Ceusters W, Smith B. Toward an Ontological Treatment of Disease and Diagnosis. Summit on Translational Bioinformatics. 2009; 2009:116–20. PubMed PMID: PMC3041577. [PubMed: 21347182]
59. SNOMED. [09/09/2017] SNOMED CT2017Available from: http://www.snomed.org/snomed-ct
60. WHO. [09/09/2017] Classification of Diseases (ICD)2017Available from: http://www.who.int/classifications/icd/revision/en/
61. Wittwehr C, Aladjov H, Ankley G, Byrne HJ, de Knecht J, Heinzle E, et al. How Adverse Outcome Pathways Can Aid the Development and Use of Computational Prediction Models for Regulatory Toxicology. Toxicological sciences : an official journal of the Society of Toxicology. 2017; 155(2): 326–36. PubMed PMID: 27994170; PubMed Central PMCID: PMC5340205. DOI: 10.1093/toxsci/kfw207 [PubMed: 27994170]
62. Pittman M, Edwards S, Ives C, Mortensen H. AOP-DB: A database resource for the exploration of Adverse Outcome Pathways through integrated association networks. Toxicology and applied pharmacology. 2017 (in submission).
63. Conolly RB, Ankley GT, Cheng W, Mayo ML, Miller DH, Perkins EJ, et al. Quantitative Adverse Outcome Pathways and Their Application to Predictive Toxicology. Environmental science & technology. 2017; 51(8):4661–72. DOI: 10.1021/acs.est.6b06230 [PubMed: 28355063]
64. Mattingly CJ, Boyles R, Lawler CP, Haugen AC, Dearry A, Haendel M. Laying a Community-Based Foundation for Data-Driven Semantic Standards in Environmental Health Sciences. Environmental health perspectives. 2016; 124(8):1136–40. PubMed PMID: 26871594; PubMed Central PMCID: PMC4977056. DOI: 10.1289/ehp.1510438 [PubMed: 26871594]
65. OECD. [last accessed October 2014] OECD Harmonised Templates2014Available from http://www.oecd.org/ehs/templates/templates.htm

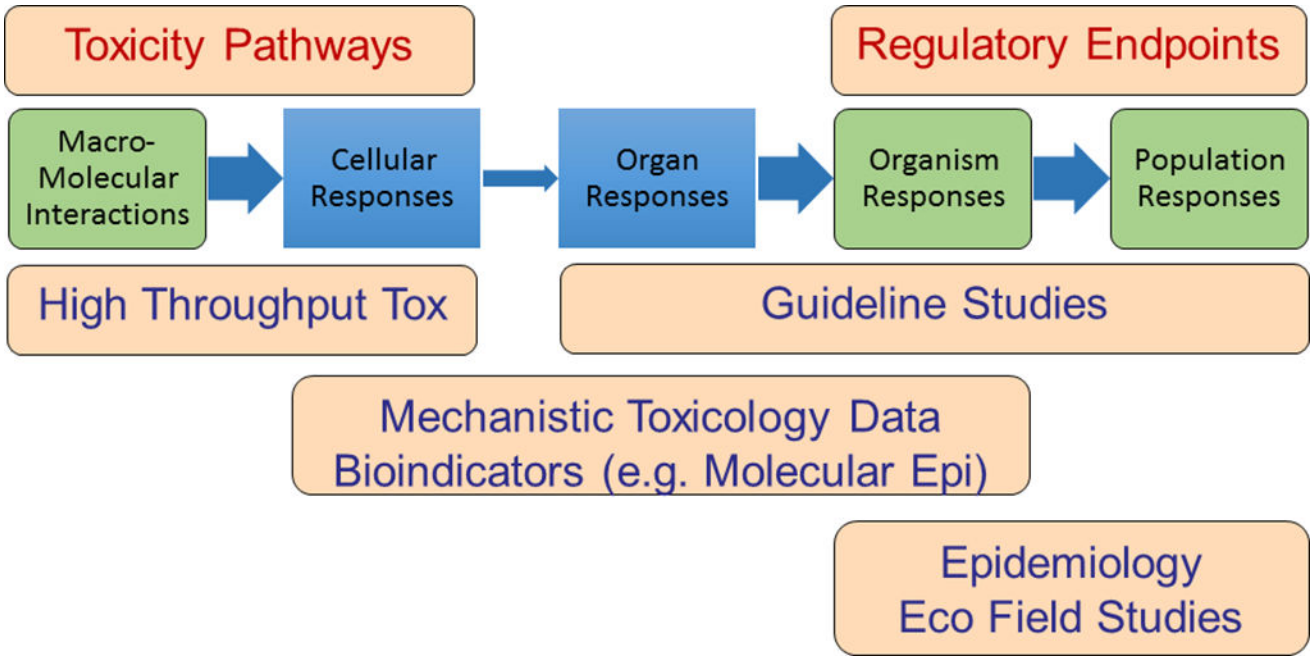


Figure 1.

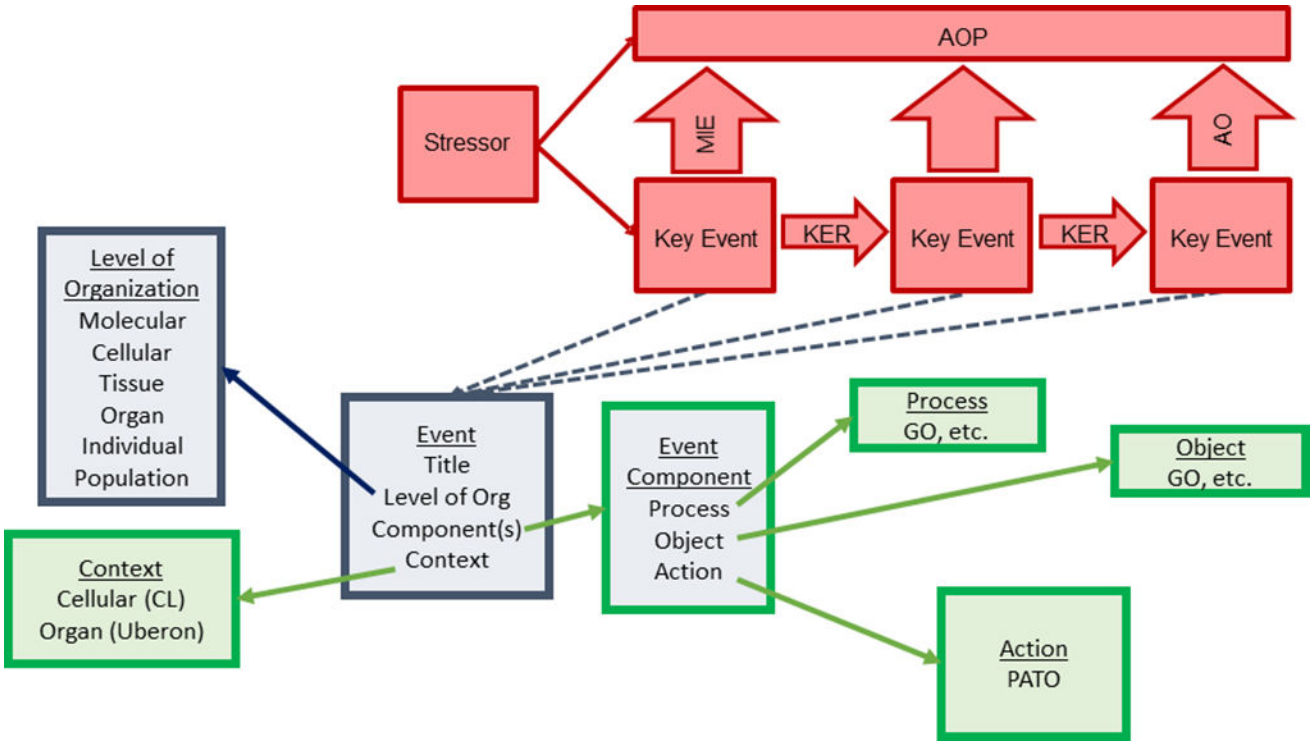


Figure 2.

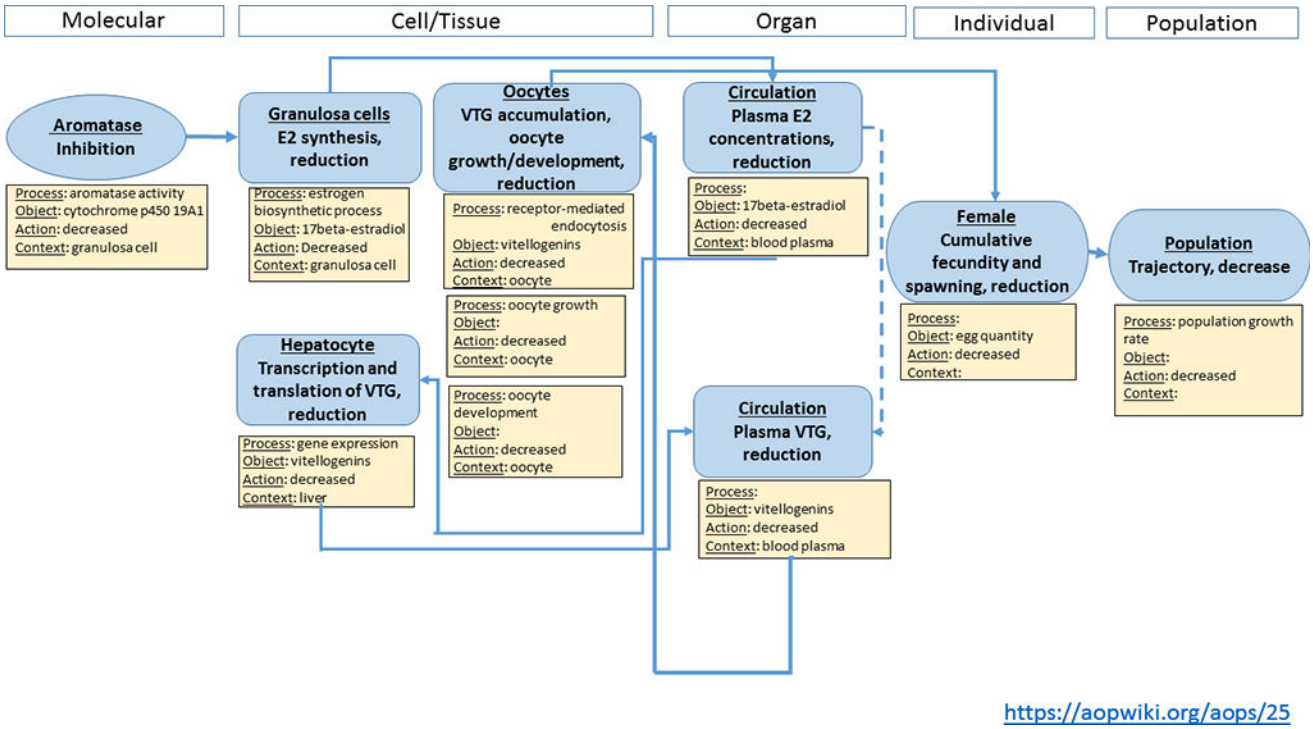


Figure 3.

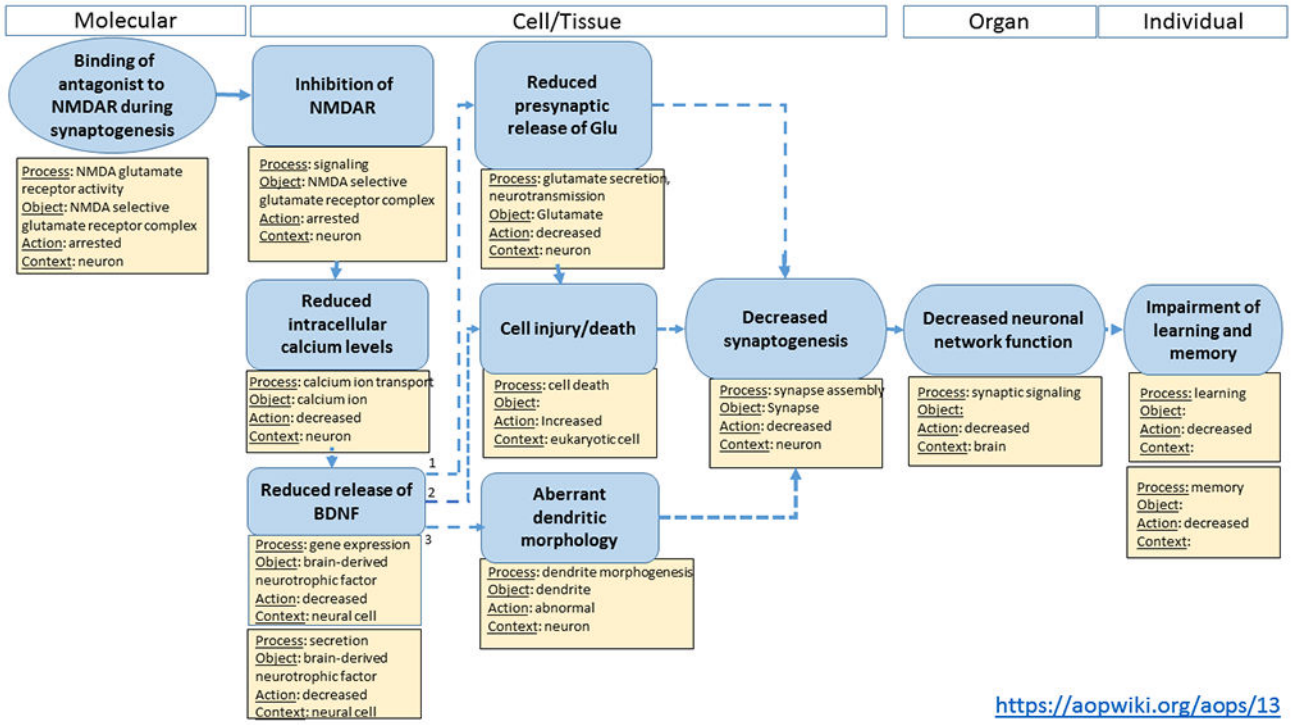


Figure 4.

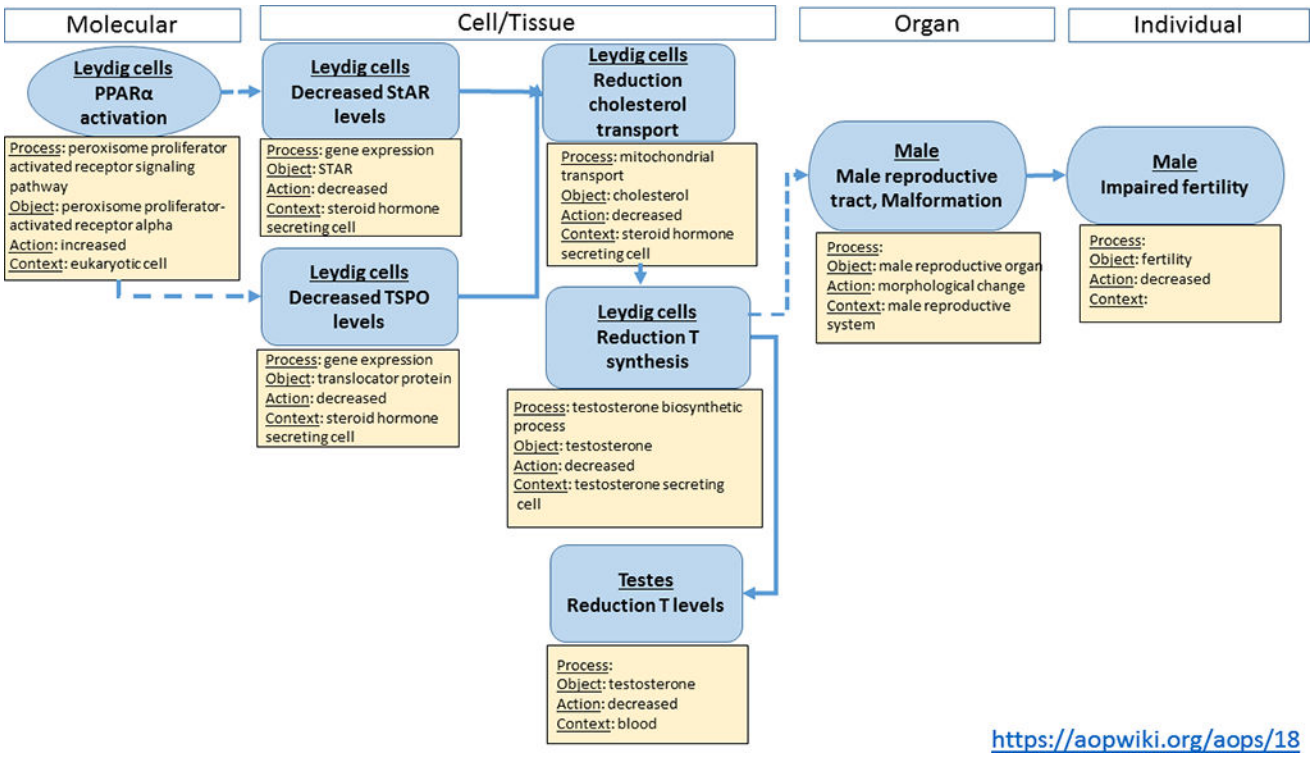


Figure 5.

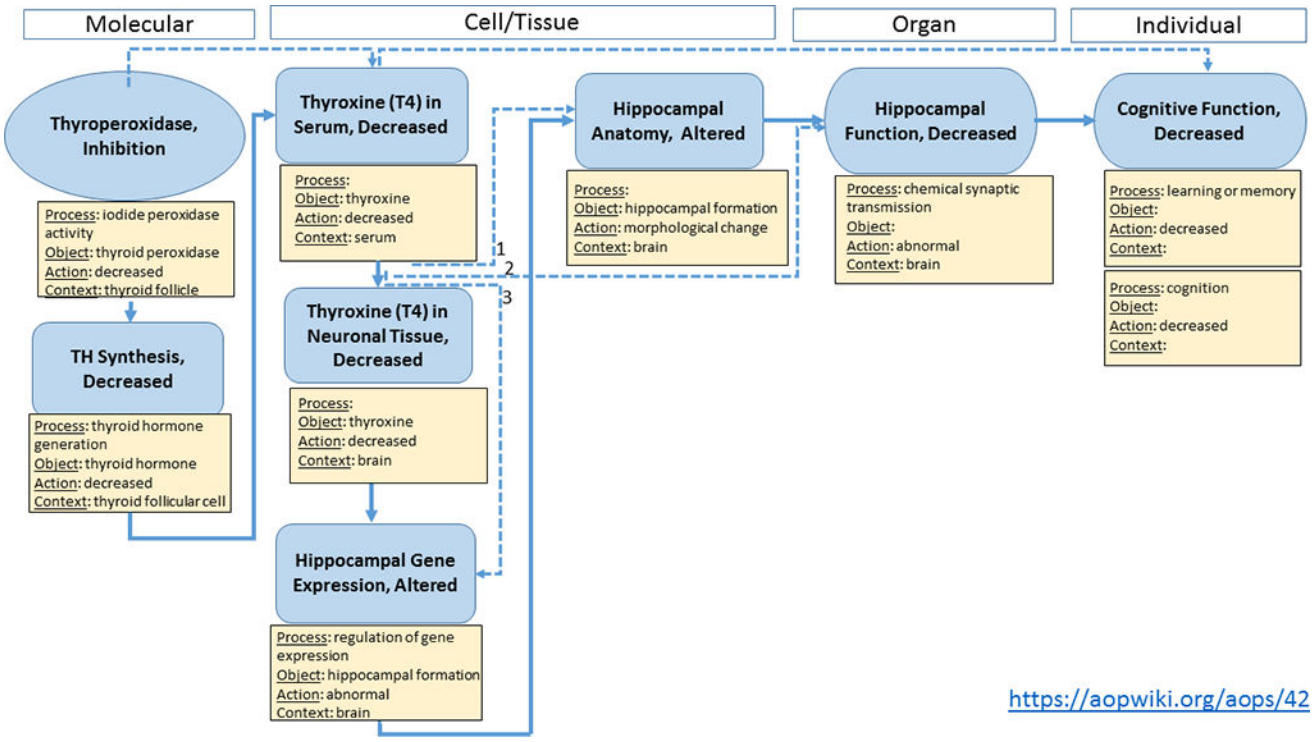
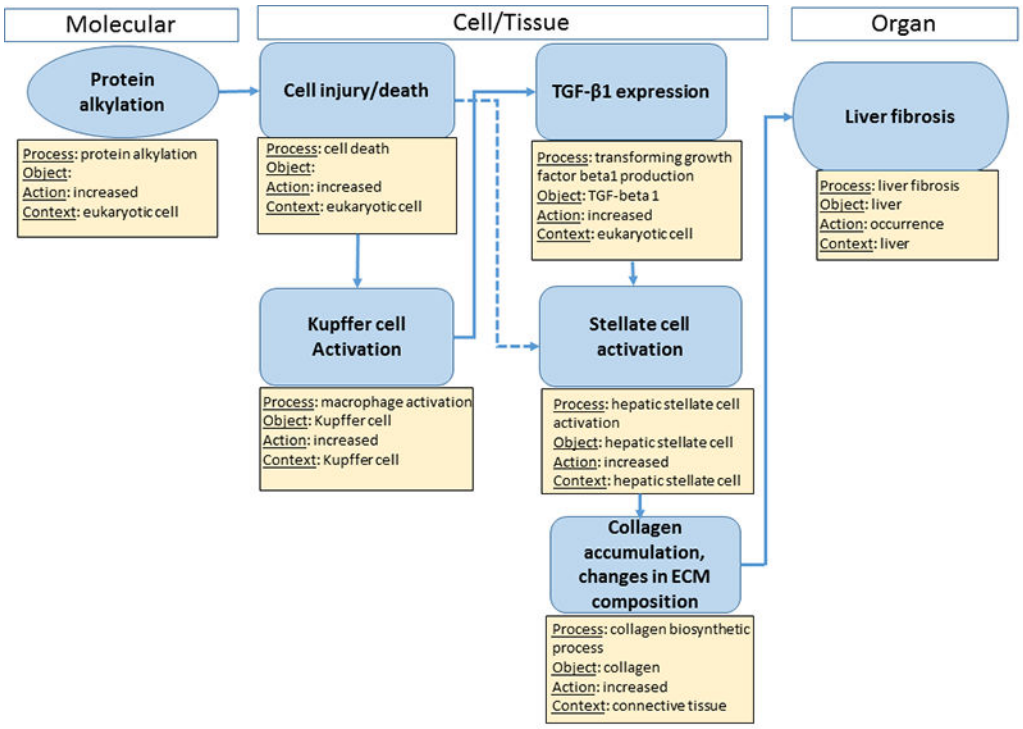


Figure 6.



https://aopwiki.org/aops/38

Figure 7.

Table 1

List of ontologies used for annotating key events

Data Source	Domain	Level of Biological Organization*						# of classes
		M	C	T	O	I	P	
	OBO Foundry							
Gene Ontology (GO)	biology	Y	Y	Y	Y	Y	?	**260**
Chemical Entities of Biological Interest (CHEBI)	biochemistry	Y	X	X	X	X	X	**67**
Protein Ontology (PRO)	proteins	Y	X	X	X	X	X	**140**
Protein-protein interaction (MI)	experiments	Y	X	X	X	X	X	**5**
Cell Ontology (CL)	cells	X	Y	X	X	X	X	**70**
Uber Anatomy Ontology (UBERON)	anatomy	?	Y	Y	Y	?	?	**80**
Foundational Model of Anatomy (FMA)	anatomy	Y	Y	Y	Y	X	X	**7**
Vertebrate Trait (VT)	vertebrate trait	X	X	Y	Y	Y	X	**7**
Human Phenotype Ontology (HP)	phenotype	X	Y	Y	Y	Y	X	**5**
Mammalian Phenotype Ontology (MP)	phenotype	X	Y	Y	Y	Y	X	**55**
Neuro Behavior Ontology (NBO)	behavioral phenotypes	X	X	X	X	Y	X	**7**
Phenotypic Quality Ontology (PATO)	phenotype	Y	Y	Y	Y	Y	Y	**N/A**
Population and Community Ontology (PCO)	populations and communities	X	X	X	X	X	Y	**3**
	Controlled Vocabulary							
UMLS/Medical Subject Headings (MeSH)	biomedical information	Y	Y	Y	Y	Y	Y	**71**
		Total						**777**

*Molecular, Cellular, Tissue, Organ, Individual,
Y = definitely covers this level of organization.
X = definitely doesn't covers this level of organ
?=maybe

Table 2

List of action terms derived from PATO

Action	PATO term	ID
increased	increased process quality	PATO:0002304
	increased object quality	PATO:0002305
	increased amount	PATO:0000470
	increased concentration	PATO:0001162
	increased duration	PATO:0000498
	increased behavioral activity	PATO:0000760
decreased	decreased process quality	PATO:0002302
	decreased quality	PATO:0002301
	decreased object quality	PATO:0002303
	decreased concentration	PATO:0001163
	decreased fecundity	PATO:0001696
	decreased amount	PATO:0001997
	decreased rate	PATO:0000911
	decreased thickness	PATO:0000592
	decreased speed	PATO:0000304
morphological change	deformed	PATO:0001617
	accumulation	PATO:0002269
	malformed	PATO:0000646
functional change	insufficient	PATO:0001628
	non-functional	PATO:0001511
	damage	PATO:0001020
occurrence	occurrence	PATO:0000057
	temporal distribution quality	PATO:0002323
disrupted	disrupted	PATO:0001507
arrested	arrested	PATO:0000297
delayed	delayed	PATO:0000502
premature	acceleration	PATO:0001028
abnormal	abnormal	PATO:0000460
	deviation(from_normal)	PATO:0000069
pathological	pathological	PATO:0001869